Machine Learning in Cybersecurity

Keith Hall

CYBR 650: Current Trends in Cybersecurity

Bellevue University

February 18, 2024

Dr. Robert Flowers

Introduction

The escalating frequency of cyberattacks on a global scale surpasses the availability and need for technological solutions. The frequency and sophistication of cyberattacks targeting critical infrastructure systems are growing at an alarming rate. Eian et al. (2020) projected the fiscal impact of cybercrime to amount to $10.5 trillion by 2025. The clandestine nature of these attacks transpires without the knowledge of their intended victims, inflicting devastation and forfeiting of financial resources, intellectual property, and organizational standing. Further complicating matters is the fact that these threats originate not only externally but also internally. Even though 80% of security breaches are caused by external threat actors, insider threats can still cause significant harm to a company and its reputation, according to the Verizon 2022 Data Breach Investigations Report. To mitigate the potential risks stemming from external sources and internal adversaries, cybersecurity professionals must possess an extensive understanding of cutting-edge technologies that can be utilized to counter emergent threats.

In recent years, the integration of artificial intelligence (AI) and machine learning (ML) across all US industry sectors has resulted in significant advancements within many cybersecurity domains. Scholars are using machine learning models to drive innovation in defensive measures that could revolutionize how organizations protect critical infrastructure and information assets. However, despite the considerable progress made possible by machine learning, there remains a discrepancy between research and practice (Apruzzese et al., 2023). At the same time, traditional security solutions are insufficient to address contemporary threats due to the exponential growth of novice attack vectors. This study investigates the impact of machine learning by identifying specific areas within six cybersecurity domains where machine learning could lead to substantial breakthroughs in resolving the complex challenges confronting the industry: 1) information and

physical security policies, standards, and guidelines, 2) IT governance, 3) risk management, 4) compliance, 5) control strategies for risk mitigation, and 6) incident response management. In addition to introducing machine learning fundamentals to the novice, the study will describe two formidable obstacles impeding its complete integration into the field of cybersecurity.

Background

Over four decades have elapsed since AI and MLs earliest applications in the field of cybersecurity. During the late 1980s, initial efforts focused on rule-based systems to detect anomalies. The advent of "big data" at the beginning of the 21st century prompted a transformation of cybersecurity toward security solutions powered by AI, which incorporated machine learning algorithms for pattern recognition and intrusion detection (Jada & Mayayise, 2023). Machine learning witnessed considerable advancement with the arrival of new and more sophisticated learning models (Gupta et al., 2023). Recent advancements in machine learning have been propelled by the proliferation of internet data accessibility, the development of "deep" learning algorithms, and low-cost computation.

Transforming Cybersecurity with Machine Learning

Machine learning enables learning and predictive capabilities in computer systems through the analysis of data, as opposed to explicit programming instructions. It has enhanced overall operational capability by leveraging the processing power of computers to efficiently handle vast quantities of data (Apruzzese et al., 2023). ML is facilitating the identification of trends and patterns in data for investors and financial organizations, thereby delivering valuable insights that enhance the governance and risk management of information technology. By streamlining

and automating their compliance processes, institutions are identifying potential risks earlier and fulfilling regulatory obligations more efficiently and affordably.

ML's potential to transform cybersecurity cannot be overstated. Machine learning has the potential to effectively tackle long-standing challenges that traditional security solutions have struggled to adequately address by providing defensive controls that are dynamically enhanced, automated, and continuously updated to adapt to shifting environments and evolving threats (Sarker, 2023). The expanding field of research relating to machine learning applications in cybersecurity now encompasses threat and anomaly detection, risk assessment, threat intelligence, behavioral analysis, biometrics, and multi-factor authentication. Nevertheless, Apruzzese et al. (2023) assert that cybersecurity, when compared to other industries, has not yet completely embraced machine learning. Despite the extensive body of research on more effective, supervised ML alternatives, Apruzzese discovered that over 90% of businesses that employ AI/ML in their defensive tools, still utilize unsupervised methods for anomaly detection. Their findings exposed a significant disparity between theoretical understanding and real-world implementation, which was ascribed to a lack of optimism among decision makers regarding the benefits and drawbacks of machine learning applications in the field. Muser and Garriott (2021) note that machine learning is frequently employed for detection purposes, but its utilization in other phases of the cybersecurity model is uncommon. Researchers have endeavored to address this issue by conducting literature reviews covering a broad range of ML applications, many of which discuss the ramifications and constraints of machine learning (Ansari et al., 2022; Das & Morris, 2017; Ford & Siraj, 2014; Mohamed, 2023; Muser & Garriott, 2021).

Additionally, some scholars have decided to target their research to a wider demographic (Apruzzese et al., 2023; Ansari et al., 2022). Nevertheless, there is a continued persistence to

focus on the advantages of employing ML algorithms to fortify conventional defensive measures.

This study builds on previous work by focusing on the domains where ML could provide solutions to the most urgent challenges in cybersecurity from a holistic perspective. However, it should not be regarded as an all-encompassing review of the current applications of ML in cybersecurity. It is an attempt to identify areas within diverse cybersecurity domains where ML could have a transformational impact. This determination was predicated on two factors: the academic community's acceptance of ML as a more effective substitute to conventional approaches and whether the technology was being adopted in practice. The study finds that while ML represents the future for cybersecurity, it faces obstacles in the form of technological constraints that must be surmounted prior to its revolutionizing the industry.

Before delving into an examination of MLs impact, it is imperative to identify the precise elements within each domain where the transformation is already underway. A concise overview of the fundamentals of ML will then follow.

Policies, Standards, and Guidelines

A review of the academic literature and the current state of ML developments in the security policy domain indicate that ML could have a significant impact on the implementation, evaluation, and dissemination of corporate policies, standards, and guidelines. Security policies establish the norms, expectations, and general posture an organization assumes while ensuring data confidentiality, integrity, and availability. Policies inform staff and others in the workplace about the guidelines and restrictions that govern the utilization of the organization's data and information assets (Whitman & Mattord, 2013, p. 170). Setting clear expectations includes the implementation of additional standards and guidelines as necessary to ensure that employees

understand how to comply with company policies. Standards furnish unambiguous instructions that specify the precise actions that are required to adhere to a policy. Guidelines are advisory rather than mandatory directives that personnel may consult to verify their adherence.

Policy Implementation and Evaluation

Ensuring information and physical security policies are consistently applied and enforced is essential for their effective implementation and legal defensibility (Whitman & Mattord, 2013, p. 197). In addition, a company must communicate how it intends to monitor and terminate conduct that it deems inappropriate. For example, an organization that chooses to enforce a policy on website access restrictions, must notify staff members that controls will be employed to monitor network activities; and as a result, access to all objectionable websites will be denied. Furthermore, to avoid being held liable if challenged for enforcing employee compliance, it must provide supporting documentation demonstrating that the content was disseminated to and accessed by the intended recipients of the policy (Whitman & Mattord, 2013, p. 197).

To assess the efficacy and quality of security policies, organizations must monitor internal communication channels, such as networks and all end points using automated tools to determine whether employees are complying. Additionally, organizations can evaluate current policies against established laws, standards, and best practices. Frameworks such as ISO 27001, NIST, or CIS can be used to benchmark policies and detect any deficiencies or deviations.

Governance, Risk Management, and Compliance

The contemporary landscape places significant emphasis on IT governance, risk management, and regulatory compliance as prerequisites for organizational success. This is primarily due to the proliferation of federal and state laws, the growing intricacy of commercial operations, and the perpetual introduction of new technologies. As a result, practitioners have consolidated these

domains into a single framework known as GRC (Pratt, 2023). However, although risk management and compliance are considered pillars of IT governance, MLs impact can be better understood when they are viewed as distinct domains.

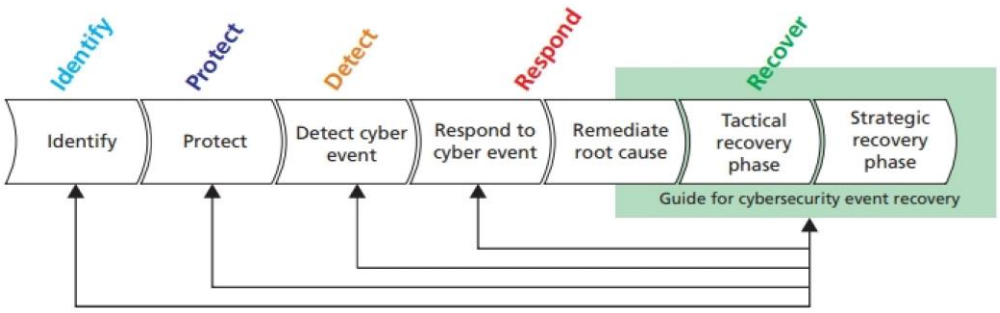
IT Governance

By leveraging crucial management functions in the areas of strategic planning and decision-making, the incorporation of ML technology has the capacity to fundamentally transform IT governance. Effective governance ensures the alignment of an organization's information technology with its business objectives, the provision of benefits to stakeholders, the minimization of risks, compliance with regulations, the establishment of streamlined and effective internal controls, and the efficient and effective management of IT resources (Otero, 2019, p. 134). Governance is the implementation of the board of directors' jurisdiction, oversight, and guidance for the organization. Strategic planning comprises a methodical execution of governance obligations with the objective of assessing and directing the objectives, activities, and motivations of an organization by means of carrying out pivotal decisions and initiatives. Like security policies, IT governance can be implemented and evaluated by comparing current governance practices to established maturity metrics, standards, and best practices using frameworks such as COBIT (Control Objectives for Information and Related Technologies), ITIL (Information Technology Infrastructure Library), ISO/IEC 38500, and the NIST Cybersecurity Framework.

NIST Cybersecurity Framework

To help understand the impact of ML on cybersecurity, it is useful to view how it is being applied from the perspective of the Cybersecurity Framework. The NIST framework categorizes processes into five concurrent and ongoing functions and offers a comprehensive strategic

overview of an organization's approach to managing cybersecurity risk to determine where to invest time and resources. At its core is a collection of intended cybersecurity tasks and outcomes for each function that are classified into categories which encompass the entire set of possible cybersecurity objectives of an organization. Governance and risk management procedures are assessed in relation to the framework's tiers. These tiers serve as an indicator of the degree to which the cybersecurity risk management practices of the organization align with the characteristics of the framework. The five core functions are listed below:



Source: NIST Special Publication 800·61, Rev. 2: The Computer Security Incident Handling Guide.

Identify by conducting an audit to determine which systems are vital to operations and identify potential risks that could have a detrimental impact on those systems.

Protect is to establish and execute appropriate safeguards to ensure the delivery of critical infrastructure services.

Detect involves implementing relevant activities for detecting the presence of a cybersecurity incident.

Respond by taking appropriate action in response to an identified cybersecurity occurrence.

Recover by taking the necessary steps to restore operations to their condition prior to the catastrophe.

Risk Management Process

Risk management is the process of identifying, assessing, managing, and reducing risk (Gibson & Igonor, 2020). NIST Special Publication 800-37 Rev. 2 defines risk as the degree of peril to which an entity is exposed due to a potential incident or occurrence. ML is likely to have the largest impact in the areas of risk assessment and control strategies for risk mitigation, including identifying a risk response, selecting controls, implementing and testing controls, and evaluating controls.

Risk Assessment

The initial phase of the risk management process is risk assessment, which involves identifying assets and their value; identifying threats and vulnerabilities to the asset; prioritizing the threats and vulnerabilities; determining the likelihood of a threat exploiting a vulnerability; and determining the impact of a risk. Additionally, there are two primary risk assessment methodologies: qualitative risk assessment evaluates, and ranks identified risks according to their severity and the probability of their consequences, whereas quantitative risk assessment (QRA) computes risk using collected data. One of the first steps of a QRA is the qualitative categorization of the frequency and severity of potential incidents. This task is normally performed by a team and can be very time and resource-consuming for large organizations. Compiling a threat inventory is a prerequisite for vulnerability assessment. Since vulnerabilities can remain hidden, the ideal timing for their identification is before an adversary threatens an asset. System logs, previous incidents, reports, response teams, and security testing are all sources that can aid in the identification of vulnerabilities.

Control Strategies for Risk Mitigation

The primary objective of control strategies for risk mitigation is to diminish the likelihood of a risk event taking place and, if it does transpire, to restrict the severity of the resulting damage. Other risk-handling strategies include risk acceptance, sharing, avoidance, and risk transfer. Nevertheless, in terms of the influence of machine learning on cybersecurity, risk reduction and prevention has thus far emerged as the principal beneficiary. The specific areas with the most potential for advancement include malware detection and analysis, intrusion detection/prevention, various physical security measures, firewall technology, and spam detection. These software applications are specifically designed to monitor systems and network traffic with the intent of detecting detrimental or unauthorized activity and access. This is achieved by managing extensive volumes of unprocessed, chronologically arranged data that are continuously generated, all while adhering to predetermined processing time constraints. (Nguyen et al., 2020).

Conventional retail and commercial **malware detection** software uses signature-based methods to detect and eliminate malware. This approach detects malicious software by utilizing signatures or codes that have been obtained from past malware attacks and are stored in a database referred to as a "blacklist." By cross-referencing the signature obtained during the assault with those in the blacklists, the system is able to identify the attack. It is critical to promptly compare the signatures to prevent the attack. While this approach has generally demonstrated an ability to secure systems, it has been shown to be ineffective when confronted with novel threats that have never been seen before or have yet to be added to the blacklist (Ansari et al., 2022).

Regulatory Compliance

Machine learning is making significant strides in the domain of regulatory compliance. Many organizations, including educational institutions, financial establishments, and medical facilities, that collect, utilize, and disclose the personal information of individuals have a responsibility to protect the privacy of those involved. Developing streamlined processes to ensure compliance with internal policies and privacy regulations, including but not limited to the Health Insurance Portability and Accountability Act (HIPAA) and the Gramm-Leach-Bliley Act (GLBA), has become an exceptionally intricate challenge for modern businesses (Garg et al., 2011).

Incident Response Management

A company's efforts to prepare for, respond to, and recover from an incident constitute incident response (IR). Response teams diligently examine and monitor systems as an essential element of the security testing process, with the objective of identifying and responding to signs of unauthorized access. As expeditious and effective incident containment and resolution is of the utmost importance to organizations, IR planning and coordination must be meticulous. Irrespective of the magnitude of the impact, an organization commences the incident response plan upon discovering an occurrence that influences it.

Machine Learning Basics

Machines require three components for the learning process to occur: a massive amount of data for training, validating, and testing; an algorithm or logical program that converts data sets into a learning model; and a training strategy for fine-tuning the accuracy and performance of the algorithm and learning model.

Using data to train an algorithm or program to arrive at a decision without explicit programming is a difficult and time-consuming process. The procedure begins with selecting the appropriate algorithm and acquiring a vast amount of data that the algorithm can learn from. The data is used to "train" the algorithm. There are four different approaches for training algorithms that comprise the separate ways machines learn: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. These broad categories determine whether labeled data sets were used to train the algorithm. Labels serve as an example of the desired prediction or outcome, referred to as the target or output.

Supervised Learning

The utilization of labeled data sets is mandatory when training supervised learning algorithms. Bharadiya (2023) describes the process of implementing supervised learning to accurately classify emails as spam. Data, representing the "features" of a spam email must be inputted into the algorithmic program to instruct it on the nature of spam. To enable the algorithm to differentiate spam, it is necessary to designate a subset of the data sets as spam. These labelled data sets will function as examples for the algorithm, which will use the data to develop a mathematical model that depicts the relationship between the expected output and the labelled inputs. When an algorithm is presented with an image of an email that it has never seen before, it applies the model to determine whether the email is spam. Further refinement of the model's accuracy and performance can be achieved through the inclusion of additional clues, specifically data sets that have been labeled as spam.

Unsupervised Learning

Labeled data sets are not an essential component of the Machine Learning process for all algorithms. Unsupervised learning and reinforcement learning utilize unlabeled data for training. The fundamental concept underlying unsupervised learning is for the model to autonomously process unstructured data to derive meaning from it. Unsupervised learning systems acquire knowledge through the process of clustering data or objects according to their similarities, or by discerning noteworthy connections, concealed patterns, or trends, which is known as association rule learning. Clustering, like other unsupervised learning algorithms, has the capability to identify shared characteristics among unlabeled and uncategorized data points in enormous datasets without any explicit guidance. A method for efficiently arranging large datasets into more comprehensible formats is known as clustering. The method of association identifies connections between variables. Although unsupervised learning is a more cost-effective and expedient method for completing training tasks, its applicability is limited to clustering, and it yields less precise outcomes.

Semi Supervised Learning

To train algorithms using semi-supervised learning (SSL), a small quantity of labeled data is combined with a large quantity of unlabeled data. Semi-supervised learning has proven to be a viable substitute in situations where the availability of labels is limited and unlabeled data is plentiful.

Reinforcement Learning

This ML model is constructed by reinforcement learning utilizing an entirely automated feedback system that enables the model to engage with its environment. Frequently, rewards and punishments constitute the feedback that the agent algorithm employs to optimize its behavior.

Once the algorithm is programmed to maximize overall and long-term rewards with the intention of generating an optimal solution, the agent acquires knowledge through trial and error as it endeavors to identify the most rewarding activities. The Markov decision process forms the fundamental basis upon which reinforcement learning systems are built. During this procedure, an agent in a particular state within an environment must select the optimal action from the numerous actions that are feasible given its current state. Specific behaviors offer incentives to enhance motivation. Upon transitioning to the subsequent state, it will acquire the capability to engage in novel and gratifying activities. The cumulative reward is determined as the agent accumulates the rewards it has received over time in exchange for the actions it elects to carry out. Although reinforcement learning has generated considerable interest within the domain of artificial intelligence, its practical implementation and widespread adoption remain constrained.

Deep Leaning

Deep learning is a kind of machine learning approach that uses multi-layered neural networks to gradually extract higher-level features from unprocessed input. It utilizes the same learning mechanisms previously mentioned. However, there is a distinction predicated on the structure of the algorithm, the diminished necessity for human intervention, and the augmented data prerequisites. Machine learning algorithms execute tasks by utilizing simple linear regression or decision trees that rely on statistical patterns and inference. In contrast, deep learning algorithms employ a layered structure of algorithms composed of "nodes" known as artificial neural networks to analyze data and draw conclusions in a logical manner akin to human reasoning. Subsequently, deep learning algorithms operate with a substantially reduced need for human intervention. The development of "deep" learning algorithms, the proliferation of internet data

accessibility, and low-cost computation have all contributed to the latest advances in machine learning.

Machine Learning Technologies

Predictive and behavioral analytics, particularly User and Entity Behavior Analytics (UEBA), are the main catalysts for advancements in cybersecurity. These techniques are employed to detect patterns and enhance security measures. User and entity behavior analytics (UEBA) is a cybersecurity system that employs machine learning algorithms to detect anomalies in the conduct of individuals within a corporate network, encompassing users, routers, servers, and endpoints (Mishra & Shekhawat, 2021). Predictive cybersecurity analytics examines patterns derived from past cyber incidents and ongoing network activities to anticipate and prevent future attacks. According to Mishra and Shekhawat (2021), predictive systems possess the capacity to perform a thorough assessment of potential risks and detect inadequacies in existing control protocols. These two technologies can be found in numerous Cybersecurity domains because of their ability to empower machine learning algorithms to forecast future opportunities, risks, and trends by analyzing historical data.

Discussion

ML-driven cybersecurity solutions offer advanced functionalities, including adaptive learning, pattern recognition, and proactive threat detection (specifically, malware). These systems can identify trends, anomalies, and potential hazards through the analysis of vast quantities of data that conventional methods may fail to notice. The finding of the present investigation demonstrates how specific areas within key cybersecurity domains will likely be impacted by advancements in ML, providing opportunities for innovative research and the development of practical solutions to address the current rise in cyberattacks.

Policies, Standards, and Guidelines

Machine learning operations predominantly influence physical and information security policies, standards, and guidelines via policy enforcement. Enforcement of issue-specific policies, like Acceptable Use, Access Control (system-specific policies), and program policies is carried out through the utilization of ML enhanced automated testing, video surveillance systems (VSS), intrusion detection systems (IDS), and electronic physical access control systems (ePACS).

Corporate Policy Compliance Monitoring and Enforcement

The potential of machine learning to significantly enhance the effectiveness of monitoring and regulating employee compliance with physical and information security policy is generating considerable enthusiasm. AI compliance monitoring tools automate the monitoring and detection of compliance violations using intelligent analytics and machine learning algorithms. The ineffectiveness and monotony of conventional compliance monitoring methods have escalated in tandem with the intricacy and magnitude of data in the contemporary business environment. Nevertheless, due to the efficient processing and analysis of enormous volumes of data, ML models have the capability to detect anomalies and patterns that would be challenging to discern through manual means. This significantly enhances the detection of non-compliant behavior. By comparing malicious traffic to a known profile, supervised methods may be used to detect future attacks, provided that well-labeled data on previous assaults is available. Conversely, unsupervised algorithms possess the capability to detect attacks solely on the grounds of their atypical and inconsistent characteristics (Muser & Garriott, 2021). Marshall et al. (2021) provided evidence that infosec behaviors—user activities that are challenging to monitor due to their subjective nature involving attitudes, beliefs, and perceptions—can compromise the efficacy of automated controls. Frequently, for instance, personnel who are dubious or skeptical

increase the risk to the organization when they use the same password to authenticate corporate and external services. The identification of these vulnerabilities is improbable through the examination of device configurations or network traffic. In the Marshall study, infosec behaviors extracted from the literature and evaluated as plausible indicators of noncompliance with security policies were used to train ML systems to predict infosec behaviors. They claim that the model can be used to predict threatening infosec behaviors before they appear in the network.

ML is transforming how companies protect their employees and assets from social engineering attacks initiated by email spam and phishing. Prominent internet service providers such as Yahoo, Gmail, and Outlook have effectively implemented machine learning models to detect and classify unsolicited bulk emails (UBEs), which are email spam and phishing attempts. Gangavarapu et al. (2020) presented an exhaustive description of the procedure utilized to extract behavior-based features and email content in their study. Furthermore, they delineated the procedure for determining the optimal feature set and specific attributes that are appropriate for the identification of Unsolicited Bulk Emails (UBEs). The academics' suggested models demonstrated a classification accuracy of 99% for UBEs.

Communicating Policies, Standards, and Guidelines

Employees are considered the most vulnerable component of a company's security framework, as they frequently inadvertently open malicious attachments and links, divulge passwords, or fail to encrypt sensitive documents. The objectives of transformational awareness programs, proposed by Carpenter (2019), encompass the modification of employee behaviors, thought processes, routines, and beliefs, as well as the promotion of positive organizational social and cultural change. Carpenter contends that the keys to effective transformation awareness training are to emulate the strategies and approaches of experts in other industries, such as marketing,

psychology, communications, and more. Significantly more ML has been implemented in these disciplines than in cybersecurity. This finding suggests that machine learning has the capacity to support the integration of cybersecurity with other fields to fundamentally alter the dissemination and communication of corporate policies.

Governance, Risk Management, and Compliance

By providing board members with real-time data, predictive analytics, strategic planning tools, and enhanced efficiency, machine learning (ML) is fundamentally transforming the governance responsibilities of members. The utilization of this capability enables board members to effectively navigate swiftly changing corporate landscapes, make well-informed decisions, and sustain a competitive advantage. By utilizing the capabilities of machine learning, board members can proficiently tackle the complex obstacles present in the modern business landscape and steer their organizations towards increased success and sustainability (Bhattacharya, 2018). Throughout history, board members have obtained information regarding the activities of the organization through a combination of reports, presentations, and in-person meetings. Despite the ongoing value of these strategies, artificial intelligence (AI) brings forth an innovative element to the process of decision-making by employing data and automation. The author of Bhattacharya's (2018) research investigated the utilization of machine learning (ML) technologies in relation to organizational strategic decision-making. Training ML systems to assist the firm's C-suite comprised the study. The utilization of machine learning technologies was illustrated in Bhattacharya's case studies, which enabled board members to simulate various scenarios and evaluate the effectiveness of distinct solutions.

Risk Assessment

One of the initial steps in **Quantitative Risk Analysis** (QRA) involves categorizing the frequency and severity of potential hazards in a qualitative manner. ML systems assist companies in prioritizing mitigation efforts by assessing and ranking probable threats. In their study, Macedo et al. (2022) introduced a system that employs machine learning classifiers to extract crucial information and knowledge from previous risk assessments. This system then generates qualitative estimations of the repercussions and frequency of unexpected events. The results demonstrate that the technique is a highly valuable instrument for supporting analysts throughout the initial stages of QRA. The potential of ML to utilize predictive analytics will revolutionize the process of identifying and prioritizing threats and vulnerabilities, accurately predict the possibility of a threat exploiting a vulnerability and forecast the impact of a cybersecurity event.

Regulatory Compliance Monitoring and Enforcement

Over the course of the last five years, regulatory compliance monitoring and enforcement have been significantly improved by research into the application of machine learning algorithms, specifically in the enhancement of "policy checking" technology.

Garg et al (2011) undertook initiatives to close the gap by developing, implementing, and evaluating an algorithm that could authenticate the adherence of audit logs to privacy and security regulations. The application utilized a policy logic that was adequately explicit in its ability to simulate real privacy regulations, including HIPAA, while remaining feasible and solvable through rigorous static analysis. This inquiry confirmed that the model exhibited adequate effectiveness to be operationalized in real-world scenarios. At present, Feng et al. (2023) are involved in interdisciplinary research concerning autonomous vehicles, specifically to

train reinforcement learning systems to assess adherence to safety regulations through the utilization of vehicle performance indicators.

Control Strategies for Risk Mitigation

Malware Detection and Analysis

All organizations are susceptible to cyberattacks. Malware is the vehicle or focal point of most attacks. The objective of control strategies must be to restrict malware infections through detection and cleaning. Prior to the implementation of machine learning (ML) technology in the field of cybersecurity, threat detection was a laborious and time-consuming process that could result in failed detection. Given the trajectory of advancements in malware detection research and the concerted efforts of government agencies, it is probable that the malware menace will be completely eradicated within the coming decade. In addition to eradicating the threat by addressing its root cause, ML can mitigate future attacks by bolstering incident response measures such as post-mortem investigations and malware analysis (Saeed et al., 2013). To address a dynamic threat environment, traditional approaches are being fortified through the integration of artificial intelligence and machine learning technologies. AI's capacity for continuous learning results in its intelligence advancing with every detected hazard, thereby fortifying forthcoming security protocols.

Intrusion Detection and Prevention

Machine learning has set itself apart from other technologies through the implementation of its wide-ranging methodologies to protect against cyber threats such as botnets, intrusion detection and prevention, denial of service attacks, and spam detection. Machine learning demonstrates exceptional efficacy in the detection of cyber threats. Machine learning (ML) systems are designed to identify potential security vulnerabilities in network data through the analysis of

unforeseen patterns. An ML may detect multiple unsuccessful login attempts originating from a foreign IP address as indicative of a possible security concern. By employing machine learning techniques in risk management, organizations can efficiently detect and mitigate these perils. The machine learning system has the capability to autonomously activate protective measures, such as blocking a suspicious IP address. Prompt implementation of this response is of the utmost importance to avert data breaches and subsequent intrusion. A considerable proportion of intrusion detection systems heavily depend on machine learning approaches, primarily due to their ability to dynamically adjust to new and unidentified threats.

Threat Intelligence

Alomari et al (2023) have conducted an extensive study on the impact of ML on Cyber Threat Intelligence (CBI). In the future, Artificial Intelligence and Machine Learning will be more deeply integrated into CBI, resulting in a substantial revolution in cybersecurity strategies. The researchers anticipate the transformative effects to involve integrating advanced technologies into traditional cybersecurity methodologies, specifically concerning automated threat detection, predictive analytics, and adaptive threat responses. The implementation of ML and AI in computer telephony integration (CTI) systems would enable the analysis and processing of enormous amounts of data at a rate that was previously unattainable. Accordingly, this measure will significantly bolster the ability to identify and mitigate emerging risks in a timely and precise manner. The future of cybersecurity is expected to be profoundly influenced by the growing utilization of increasingly advanced Threat Intelligence Platforms (TIPs). It is anticipated that in the future years, TIPs will progress into more sophisticated systems that provide deeper levels of integration and more comprehensive insights. Future platforms of this nature will have the capacity to collect and analyze data from an even greater variety of sources,

such as hidden and inaccessible regions of the internet, social media, and the Internet of Things (IoT) ecosystem, which is expanding at an accelerated rate (Alomari et al., 2023). By means of this expansion, a considerably more extensive and diverse viewpoint of the threat landscape will be presented.

Physical Security Controls

Modern Physical systems consist of multiple elements and measures such as site layout and security configuration, visibility of critical areas using lighting and video cameras, access control— from simple locks to keypads and biometric access, perimeter protection, intrusion detection (motion sensors, cameras, and tripwire alarms), hiring/training, and incident response and testing. Enforcement of issue-specific policies, like Acceptable Use, Access Control (system-specific policies), and program policies is carried out through the utilization of automated testing, video surveillance systems (VSS), intrusion detection systems (IDS), and electronic physical access control systems (ePACS). Scholars are concentrating their efforts on machine learning to drive innovation in each of these security systems (Lane & Brodley, 1997; Shankar et al., 2020; Thornton et al., 2015; Tsakanikas & Dagiuklas, 2018; Ucar & Ozhan, 2017; Vanin et al., 2022).

Incident Response Management

By utilizing rule-based logic and/or machine learning, incident response automation can simplify the incident response procedure. It can be employed by teams to automate tasks such as creating a responder conversation channel, adding responders to an incident, or initiating a conference bridge. Incident management is an essential component in the supervision of IT services, their enhancement, and the attainment of organizational objectives. Prihandono et al (2020) confirmed that predictions of future incidents may be facilitated by employing ML techniques. To forecast

IT incidents, they examined the factors that contribute to incidents by employing both early and modern machine learning techniques (Random Forest, SVM, Multilayer Perceptron, RNN, LSTM, and GRU).

Conclusion

The frequency and sophistication of cyberattacks targeting critical infrastructure systems are growing at an alarming rate. To successfully detect and mitigate the potential risks inspired by outside and adversaries within, cybersecurity experts must have a thorough knowledge of recent technologies for the purpose of applying it to address emerging threats.

In recent years, the integration of AI and Machine Learning across all sectors has resulted in significant advancements within several cybersecurity domains. This study investigates the impact of machine learning on cybersecurity through the identification of six cyber domains where machine learning could lead to substantial progress in resolving the complex challenges confronting contemporary cybersecurity: The present inquiry corroborates the conclusions drawn by prior investigators who detected discrepancies in the alignment between cyber research and practical application when compared with other industries.

References

Alomari, E. S., Nuiaa, R. R., Alyasseri, Z. A. A., Mohammed, H. J., Sani, N. S., Esa, M. I., & Musawi, B. A. (2023). Malware detection using deep learning and correlation-based feature selection. Symmetry, 15(1), 123.

Ansari, M. F., Dash, B., Sharma, P., & Yathiraju, N. (2022). The Impact and Limitations of Artificial Intelligence in Cybersecurity: A Literature Review. International Journal of Advanced Research in Computer and Communication Engineering.

Apruzzese, G., Laskov, P., Montes de Oca, E., Mallouli, W., Brdalo Rapa, L., Grammatopoulos, A. V., & Di Franco, F. (2023). The role of machine learning in cybersecurity. Digital Threats: Research and Practice, 4(1), 1-38. (Apruzzese et al.,2023)

Bharadiya, J. (2023). Machine Learning in Cybersecurity: Techniques and Challenges. European Journal of Technology, 7(2), 1-14.

Bhattacharya, P. (2018, November). Artificial Intelligence in the Boardroom: Enabling 'Machines' to 'Learn'to Make Strategic Business Decisions. In 2018 Fifth HCT Information Technology Trends (ITT) (pp. 170-174). IEEE.

Carpenter, P. (2019). Transformational security awareness: What neuroscientists, storytellers, and marketers can teach us about driving secure behaviors. John Wiley & Sons.

Cichonski, P. R., Millar, T., Grance, T., & Scarfone, K. (2012). Computer Security Incident Handling Guide : Recommendations of the National Institute of Standards and Technology. https://doi.org/10.6028/nist.sp.800-61r2

Cybersecurity framework. (n.d.). GSA. https://www.gsa.gov/technology/it-contract-vehicles-and-purchasing-programs/information-technology-category/it-security/cybersecurity-framework

Eian, I. C., Yong, L. K., Li, M. Y. X., Qi, Y. H., & Fatima, Z. (2020). Cyber attacks in the era of covid-19 and possible solution domains.

Feng, S., Sun, H., Yan, X. et al. Dense reinforcement learning for safety validation of autonomous vehicles. Nature **615**, 620–627 (2023). https://doi.org/10.1038/s41586-023-05732-2

Ford, V., & Siraj, A. (2014, October). Applications of machine learning in cyber security. In Proceedings of the 27th international conference on computer applications in industry and engineering (Vol. 118). Kota Kinabalu, Malaysia: IEEE Xplore.

Gangavarapu, T., Jaidhar, C. D., & Chanduka, B. (2020). Applicability of machine learning in spam and phishing email filtering: review and approaches. Artificial Intelligence Review, 53, 5019-5081.

Garg, D., Jia, L., & Datta, A. (2011, October). Policy auditing over incomplete logs: theory, implementation and applications. In Proceedings of the 18th ACM conference on Computer and communications security (pp. 151-162).

Gibson, D., & Igonor A. (2020). Managing Risk In Information Systems (3rd ed.). Jones & Bartlett Learning.

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From chatgpt to threatgpt: Impact of generative ai in cybersecurity and privacy. IEEE Access.

Jada, I., & Mayayise, T. O. (2023). The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review. Data and Information Management, 100063.

Lane, T., & Brodley, C. E. (1997, October). An application of machine learning to anomaly detection. In Proceedings of the 20th national information systems security conference (Vol. 377, pp. 366-380). Baltimore, USA.

M. A. Prihandono, R. Harwahyu and R. F. Sari, "Performance of Machine Learning Algorithms for IT Incident Management," *2020 11th International Conference on Awareness Science and Technology (iCAST)*, Qingdao, China, 2020, pp. 1-6, doi: 10.1109/iCAST51195.2020.9319487.

Macedo, J. B., das Chagas Moura, M. J., Ramos, M., Lins, I. D., & Zio, E. (2022). Machine learning-based models to prioritize scenarios in a Quantitative Risk Analysis: An application to an actual atmospheric distillation unit. Journal of Loss Prevention in the Process Industries, 77, 104797

Marshall, B., Curry, M., Crossler, R. E., & Correia, J. (2021). Machine learning and survey-based predictors of InfoSec non-compliance. ACM Transactions on Management Information Systems (TMIS), 13(2), 1-20.

Mishra, P. K., & Shekhawat, K. (2021, March). Quantitative Risk Analysis with AI based Prediction Model. In 2021 8th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 763-768). IEEE.

Mohamed, N. (2023). Current trends in AI and ML for cybersecurity: A state-of-the-art survey. Cogent Engineering, 10(2), 2272358.

Muser, M., & Garriott, A. (2021). Machine learning and cybersecurity: Hype and reality. Center for Security and Emerging Technology (CSET), Georgetown University. https://cset. georgetown. edu/wp-content/uploads/Machine-Learning-and-Cybersecurity. pdf.

Nguyen, G., Dlugolinsky, S., Tran, V., & García, Á. L. (2020). Deep learning for proactive network monitoring and security protection. ieee Access, 8, 19696-19716.

NIST SP 800-37 Rev. 2 Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy. (2018). In *NIST*. https://https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-37r2.pdf

Otero, A. R. (2018). Information Technology control and audit. In Auerbach Publications eBooks. https://doi.org/10.1201/9780429465000

Pratt, M. (2023, December 28). What is GRC? The rising importance of governance, risk, and compliance. CIO. https://www.cio.com/article/230326/what-is-grc-and-why-do-you-need-it.html

Saeed, I. A., Selamat, A., & Abuagoub, A. M. (2013). A survey on malware and malware detection systems. International Journal of Computer Applications, 67(16).

Sarker, I. H. (2023). Machine learning for intelligent data analysis and automation in cybersecurity: current and future prospects. Annals of Data Science, 10(6), 1473-1498

Shankar, K., Iyer, V., Iyer, K., & Pandhare, A. (2020, February). Intelligent video analytics (iva) and surveillance system using machine learning and neural networks. In 2020 International Conference on Inventive Computation Technologies (ICICT) (pp. 623-627). IEEE.

Thornton, C., Cohen, O., Denzinger, J., & Boyd, J. E. (2015). Automated testing of physical security: Red teaming through machine learning. Computational Intelligence, 31(3), 465-497

Tsakanikas, V., & Dagiuklas, T. (2018). Video surveillance systems-current status and future trends. Computers & Electrical Engineering, 70, 736-753.

2023 Data Breach Investigations report. (n.d.). Verizon Business. https://www.verizon.com/business/resources/reports/dbir/

Ucar, E., & Ozhan, E. (2017). The analysis of firewall policy through machine learning and data mining. Wireless Personal Communications, 96, 2891-2909.

Vanin, P., Newe, T., Dhirani, L. L., O'Connell, E., O'Shea, D., Lee, B., & Rao, M. (2022). A study of network intrusion detection systems using artificial intelligence/machine learning. Applied Sciences, 12(22), 11752.

Whitman, M. E., & Mattord, H. J. (2013). Management of information security. http://ci.nii.ac.jp/ncid/BB0274567X