Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 16th International Scientific and Practical Conference. 2025. Volume V, 323-327

Application of Deep Learning in Artificial Intelligence Systems for Cyberattack Identification and Prevention

Gergana Varbanova
Department of Information Technologies
Nikola Vaptsarov Naval Academy
Varna, Bulgaria
g.varbanova@naval-acad.bg

***Abstract* — Cyberspace has been established as the fifth domain of warfare, alongside land, sea, air, and space. With the increasing complexity and frequency of cyberattacks, traditional security mechanisms are becoming increasingly inadequate, necessitating the integration of Deep Learning (DL) into cybersecurity. The capability of automated detection and prevention of cyber threats through the analysis of large volumes of data, anomaly identification, and attack prediction—threats that would otherwise remain undetected by conventional security systems—plays a crucial role in ensuring cyberspace security. Deep neural networks, such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and autoencoders, are instrumental in analyzing network traffic and user behaviour, enabling the early identification of cyberattacks. Deep learning can also be applied to automate incident response mechanisms. The high-speed analytical capabilities and self-adaptive nature of these models allow for dynamic cybersecurity defenses, including the autonomous blocking of malicious traffic, the identification of compromised systems, and the reduction of response time. Generative Adversarial Networks (GANs) further enhance security by simulating potential cyberattacks, thereby enabling the testing and refinement of defense strategies. However, these same networks can also be exploited for adversarial purposes, such as manipulating input data to deceive AI-driven cybersecurity systems, leading to incorrect classifications or attack predictions. Deep learning represents a transformative advancement in cybersecurity, offering intelligent, automated, and proactive solutions to combat evolving cyber threats. The continued development of adaptive AI-driven security systems will play a pivotal role in enhancing cyber resilience and safeguarding critical infrastructure against emerging cyber threats. The article examines the role of deep learning in cybersecurity, integrating technological aspects and the legal framework of the European Union (GDPR, eIDAS, AI Act) while analyzing the dual-use risk associated with generative adversarial networks (GANs). The emphasis is placed on the significance of biometric data and the necessity of human oversight, as excessive reliance on automated algorithms may lead to a false sense of security and discriminatory effects.**

Keywords*— *Artificial Intelligence, Cybersecurity, Deep Learning, Generative Adversarial Networks

I. INTRODUCTION

Cyberspace constitutes a powerful resource that provides swift and convenient access to information; however, it has also emerged as the fifth domain of warfare. The advancement of technology has led to its increasing use as a medium for orchestrating attacks aimed at undermining national security and compromising elements of critical infrastructure. Simultaneously, artificial intelligence (AI) systems are experiencing a surge in innovation, and their integration into security and defense is becoming increasingly crucial for analyzing, preventing, mitigating, and countering cyberattacks that target national security or transnational organizations.

For the purposes of this study, it is essential to define the concept of "artificial intelligence," which, in theoretical terms, is understood as the capability of computers or other machines to perform tasks that require human intelligence and rational behaviour—i.e., for a machine to be recognized as simulating human behaviour.

With the adoption of Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), a legal definition of the term "AI system" was established. It is defined as a machine-

Online ISSN 2256-070X
https://doi.org/10.17770/etr2025vol5.8490

based system designed to operate with varying levels of autonomy and capable of exhibiting adaptability after deployment. Such a system, either explicitly or implicitly, generates outcomes—such as predictions, content, recommendations, or decisions—based on the input data it processes, which may subsequently impact a physical or virtual environment.

The European Union's efforts are directed not so much at harmonizing and developing a common policy for building AI systems in specific key sectors [1] as at establishing regulations and imposing limitations on the application of AI where such systems affect fundamental human rights and may lead to their restriction.

Meanwhile, China and the United States have developed clear policies regarding the funding and integration of AI without imposing an overly restrictive regulatory framework. This, in turn, has concentrated high-tech companies and their resources within their respective territories.

II. Materials and Methods

This study employs the comparative legal method to examine the existing European Union legal framework governing artificial intelligence and its role in cybersecurity. The aim is to identify common principles, key differences, and potential gaps in the current legislation.

Additionally, a comparative analysis is conducted on European legal instruments such as the eIDAS and AI Act, focusing on provisions related to personal data protection, biometric technologies, and AI applications in cybersecurity.

The study explores the role of deep learning in cybersecurity, examining models such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Autoencoders, and Generative Adversarial Networks (GANs). These techniques are analyzed in the context of cyber threat detection, automated incident response, and attack simulation. The discussion further considers the legal and ethical challenges surrounding AI-driven cybersecurity, particularly in relation to data protection, algorithmic bias, and regulatory compliance.

III. Results and discussion

In the context of cybersecurity, the development of AI-based systems raises the question of their application as part of defense mechanisms aimed at mitigating the consequences of cyberattacks, including the reduction of material damage to critical infrastructure that is essential for security.

An increasingly significant aspect of cybersecurity is the so-called Defense-in-Depth (DiD) approach, which is based on the concept of constructing multi-layered defense mechanisms [2] through the implementation of Deep Learning (DL) techniques.

The use of deep learning is of critical importance for defense, including in the maritime sector [3], which has also been targeted by cyberattacks in recent years. The vulnerability of various sectors of public life and security generates substantial financial losses, which increase exponentially on an annual basis.

From financial fraud facilitated by social engineering techniques to the disruption of critical infrastructure—such as the energy grid, the maritime sector [4], financial institutions, governments, and supranational organizations—all remain potential victims of cyberattacks.

Deep learning represents an advanced methodology in the field of machine learning, significantly enhancing the automation and precision of data analysis. A key component of deep learning is innovative artificial intelligence (AI) systems, including deep neural networks (DNNs) and recurrent neural networks (RNNs) [5],

Without delving into the technical aspects of the technology, which are beyond the scope of this study, its application in cybersecurity is of paramount importance for identifying sophisticated attacks that are difficult to detect using traditional security systems. These conventional systems often require constant human oversight.

DNN systems enable continuous monitoring and analysis of network traffic, allowing for the detection of anomalous user behaviour, which may indicate potentially malicious activity.

Another type of artificial neural network is the generative adversarial network (GANs), whose primary application is attack simulation. This process significantly supports cybersecurity experts in developing effective defense strategies. By simulating realistic cyberattack scenarios, GANs contribute to the enhancement of early detection methods and countermeasures against malicious actions, as well as the identification of anomalous behaviour in cyber incidents that do not result from human intervention.

AI systems are a valuable resource in processing unstructured and large-scale data, as well as in their automated analysis. This includes the examination of text, images, and behavioural patterns, which, through network analysis, can help identify attacks such as social engineering (phishing), distributed denial-of-service (DDoS) attacks, and other cyber threats.

Prioritizing legislative activity, including criminal law norms, should not come at the expense of technological solutions for cybersecurity protection. Effective analysis, counteraction, and mitigation of cyberattacks require a combination of legal regulation [6] with the implementation of appropriate security systems as well as ensuring the quality of data used in deep learning (DL).

The quality and volume of the data used are critical factors, as inadequately selected input data may lead to erroneous or unpredictable results. To ensure ethical, lawful, and technologically sustainable cybersecurity, Article 10 of the Artificial Intelligence Act establishes imperative requirements concerning training, validation,

and testing data for high-risk AI systems, including those in the field of cybersecurity. This regulation aims not only to enhance the transparency of neural networks, whose interpretability and reliability pose challenges in critical areas such as energy, medicine, and maritime affairs, but also to address specific risks in the cybersecurity domain. To ensure precision and reliability, it is essential to combine deep learning with verification methods, ethical oversight, and an interdisciplinary approach to the validation and analysis of output data.

The process of verification and subsequent control of output data prevents potential false positives, inaccuracies, and distortions in models, which may compromise security mechanisms against cyberattacks or trigger defense protocols due to incorrectly identified behavior as a threat. Additionally, data must be aligned with the specific application of the AI system and reflect the geographical, behavioral, and functional characteristics of the environment in which the high-risk DL system will be deployed. This is critical for preventing incorrect or biased conclusions, which could lead to adverse consequences for security and society.

In the context of cybersecurity, despite significant advancements in the use of artificial intelligence and deep learning, the human factor remains crucial for the effective management of these technologies. While AI accelerates data processing and automates incident response [7], its judgment does not always account for context, necessitating expert oversight. Over-reliance on algorithms creates a risk of a false sense of security, particularly in cases of attacks aimed at manipulating the model. Effective cybersecurity requires a balanced approach in which AI assists but does not replace human expertise.

If deep learning systems rely on incomplete or inaccurate data, the algorithm may block legitimate traffic or fail to identify and prevent an actual cyberattack. Ensuring the security of training data is of particular importance, as such data can themselves become the target of malicious interference aimed at altering and destabilizing the operation of the respective deep learning system. Finally, one of the most serious risks is the excessive trust cybersecurity experts place in algorithms, which may lead to the oversight of unconventional cyberattacks that AI is not trained to detect but could be identified as anomalies by a human analyst.

In the context of cybersecurity, predictive systems based on artificial intelligence (AI) and deep learning play a crucial role in the prevention and management of cyber threats. Their primary objective is the timely identification of threats and the prediction of their potential consequences before any damage is inflicted on an information system or critical infrastructure. These AI-driven systems overcome the limitations of traditional security mechanisms, which primarily rely on predefined rules and databases of known attacks. However, this approach often proves ineffective against emerging threats, including zero-day attacks.

In the field of cybersecurity, predictive models are of paramount importance, as their algorithms possess the capability for dynamic adaptation to the evolving cyber environment. For instance, if a predictive AI system detects anomalous behavior—such as a sudden increase in requests from a particular IP address or attempts to access secured systems—it can automatically block malicious traffic, activate additional authentication measures, or notify a cybersecurity expert for further analysis. The interaction between predictive AI systems and human expertise is essential, as it enables the reduction of response time and minimizes the risk of compromising critical systems.

The integration of predictive AI systems based on deep learning into critical infrastructure represents a significant advancement in cybersecurity. These technologies not only enhance early attack detection mechanisms but also transform the security paradigm, making it predominantly proactive rather than merely reactive.

Undoubtedly, artificial intelligence (AI) systems based on deep learning play a crucial role in cybersecurity [9]. However, these same technologies can be utilized not only as a defensive mechanism but also as a means of executing complex cyberattacks. This dual nature of AI algorithms presents significant challenges to cyberspace security and necessitates the development of advanced countermeasures.

Traditional cyberattacks rely on pre-programmed scripts and techniques such as phishing, social engineering, and Distributed Denial-of-Service (DDoS) attacks. The integration of AI-driven systems into this domain has led to the emergence of more autonomous, faster, and harder-to-detect cyberattacks. AI algorithms not only enable the identification of new vulnerabilities in real time but also facilitate their exploitation before they are detected and mitigated.

Particularly concerning is the use of generative neural networks, which can create highly convincing fake messages, deepfake images, and video content. These technologies can be leveraged to destabilize social groups, manipulate public opinion, or compromise the reputation of public figures. As a result, there is an increasing need for more sophisticated methods to detect AI-generated content and the development of effective counterstrategies against AI-powered cyberattacks.

One of the most effective cyberattacks facilitated by artificial intelligence systems is advanced phishing. In this method, the algorithm analyzes the social profiles of the targeted individual, examining their behavior and generating personalized messages that appear entirely authentic.

The expansion of deepfake technologies has led to their increasing application in the creation of fraudulent audio and video content, which can be used to manipulate targeted groups or specific individuals. Deepfake content is increasingly employed as a tool of social engineering, particularly in the context of financial fraud, resulting in significant financial losses on a global scale.

Another significant aspect in the development of artificial intelligence (AI) and deep learning systems is the use of biometric data. Numerous companies employ fingerprints and facial recognition to manage access to key corporate resources, yet the collection and processing of such data is a process governed by statutory regulations [8]. Biometric data are highly sensitive and fall under a special category of data subject to stringent legal protection throughout all Member States of the European Union. The fundamental principles and requirements for collecting, storing, and processing personal data are derived from the provisions of the General Data Protection Regulation (GDPR), but they are further elaborated in Regulation (EU) No 910/2014 (eIDAS) and the Artificial Intelligence Act.

Pursuant to paragraph 14 of the Artificial Intelligence Act, biometric data may enable the authentication, identification, or categorization of natural persons, as well as the recognition of individuals' emotions. In the context of establishing an eID wallet, the use of biometric identification is permissible. The Act itself restricts profiling based on biometric data and the categorization of individuals according to emotional analysis conducted by AI systems and the results generated by such systems [11]. Thus, when biometric categorization serves only an auxiliary function, paragraph 16 of the Act allows social networks to employ filters that categorize facial or bodily features, enabling users to add or modify images or video content, given that such a filter cannot be employed without the primary service offered by social networks—namely, the online sharing of content. On one hand, this legislative provision appears logical, yet on the other, it presents a substantial risk associated with the use of deepfake content.

From a cybersecurity standpoint, deepfake technologies give rise to social engineering tactics, identity fraud, and falsified communications, targeting both individual citizens and organizations. Generating realistic yet fabricated video or audio materials can mislead users into adopting erroneous beliefs, undermining trust in the cyber domain and potentially compromising public order or specific individuals. The Act acknowledges that, in certain instances, there exists a risk of technical inaccuracies in AI systems intended for remote biometric identification, which may lead to biased outcomes and discrimination against specific categories of natural persons. The Regulation (i.e., the Act) permits a particular category of authorized processing of biometric data and categorization of individuals, namely when such authorization is linked to detecting, preventing, or investigating criminal offenses, provided that appropriate safeguards for the rights and freedoms of third parties are observed and in compliance with Union law—see Article 50 of the Artificial Intelligence Act. High-risk AI systems employed in cybersecurity and the protection of critical infrastructure warrant particular attention. Precisely due to their heightened importance, the Act introduces the so-called "human oversight" [12], enshrined in Article 14, which is of paramount significance for critical systems in order to avert automated decisions with unforeseeable outcomes and to ensure that human control is exercised throughout the entire period in which AI systems are in use.

Furthermore, AI-driven systems designed with malicious algorithms can scan networks and adapt to security mechanisms by automatically encrypting or modifying their code to evade detection. Such an attack is particularly dangerous, as it can operate without human intervention and spread exponentially.

Artificial intelligence systems can facilitate the detection of zero-day vulnerabilities. By analyzing large code bases, an algorithm can identify previously unknown weaknesses, which may subsequently be exploited before developers can address them, potentially enabling a cyberattack. This creates the possibility for the automated generation of zero-day exploits, significantly increasing the security risk for various systems.

As part of modern social engineering, fake social media profiles controlled by artificial intelligence systems can interact with targeted individuals in a manner that appears entirely authentic. As a result, the targeted victim is often deceived into disclosing personal or sensitive information or even initiating substantial financial transactions in favor of malicious actors.

In recent years, there has been a growing use of deepfake videos and audio messages, which can be employed to manipulate public opinion or interfere in electoral processes [10]. This type of attack is particularly dangerous in the context of regional destabilization and ongoing military conflicts, as it facilitates the spread of highly convincing disinformation, making detection and counteraction significantly more challenging.

IV. Conclusion

Ensuring a secure cyberspace is directly linked to the implementation of real-time monitoring through artificial intelligence systems based on deep learning, which can identify anomalies in network traffic and user behavior. This dynamic adaptability of algorithms significantly reduces response time and minimizes the impact of cyberattacks, allowing for timely intervention before attacks cause serious security breaches.

Despite the high efficiency of automated technologies, cybersecurity cannot rely solely on technological solutions. The human factor remains a crucial element in the protection process, as cybersecurity experts play a key role in monitoring, validating artificial intelligence system outputs, and making strategic decisions to counteract sophisticated threats.

Effective cybersecurity requires a balanced approach that integrates automation, human expertise, and a regulatory framework to ensure resilience against the increasing artificial intelligence driven threats.

REFERENCES

[1] S. Mircheska, "Perspectives for the Development of Artificial Intelligence: Security Dimension of Artificial Intelligence in the

Military Sector," *Security and Defense*, vol. 2, pp. 43-55, Dec. 2024. Available: https://doi.org/10.70265/ETNV6785

[2] V. Babanov, "Internals of Defense-In-Depth Strategy in Cybersecurity," *Security and Defense*, vol. 2, pp. 37-42, Dec. 2024. Available: https://doi.org/10.70265/PNEZ3158

[3] S. Lutzkanova, "The Regionalization of European Security and Defense Policies in the Maritime Domain: The Black Sea Case," *Pedagogika-Pedagogy*, vol. 93, pp. 208-216, 2021. Available: https://doi.org/10.53656/ped21-7s.18def

[4] B. Nikolov and D. Nikolov, "A scenario-based approach to cybersecurity training for seafarers," in *Proceedings of the International Association of Maritime Universities Conference 2024*, Massachusetts Maritime Academy, United States, 2024, pp. 385–390.

[5] R. Nishat, N. Arif, and A. Bajwa, "A comprehensive review of machine learning and deep learning applications in cybersecurity: an interdisciplinary approach," *Academic Journal on Science, Technology, Engineering & Mathematics Education*, 2024. Available: https://doi.org/10.69593/ajsteme.v4i04.118

[6] L. Lyubenov, "Domestic Legal Measures for Increasing Security in the Black Sea Region," *Security and Defense*, vol. 2, pp. 9-18, Dec. 2024. Available: https://doi.org/10.70265/VIRM1297

[7] V. Petrova, "A decision hierarchical model of cybersecurity risk assessment," in *Mathematics and Education in Mathematics*, Proceedings of the Fiftieth Spring Conference of the Union of Bulgarian Mathematicians, Burgas, Bulgaria, Sept. 2021, pp. 191-195.

[8] A. Kirkov, "Судебно-экспертная деятельность экспертов в Болгарии. Организация и проблемы," in *Цифровизация деятельности судов: текущие и перспективные задачи*, 2022, pp. 5-9.

[9] M. Oinonen and W. Morsi, "An Automated Wavelet Generation Tool for Cyberattack Detection in Substation Automation Systems," in *2024 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2024, pp. 512-517. Available: https://doi.org/10.1109/CCECE59415.2024.10667244

[10] R. Frick and M. Steinebach, "A Photo and a Few Spoken Words Is All It Needs?! On the Challenges of Targeted Deepfake Attacks and Their Detection," in *Proceedings of the 3rd ACM Workshop on the Security Implications of Deepfakes and Cheapfakes*, 2024. Available: https://doi.org/10.1145/3660354.3660357

[11] I. Nesterova, "Questioning the EU proposal for an Artificial Intelligence Act: The need for prohibitions and a stricter approach to biometric surveillance," *Information Polity*, 2022. Available: https://doi.org/10.3233/ip-211524

[12] L. eEnqvist, "'Human oversight' in the EU Artificial Intelligence Act: What, when and by whom?," *Law, Innovation and Technology*, vol. 15, pp. 508-535, 2023. Available: https://doi.org/10.1080/17579961.2023.2245683