MACHINE LEARNING OR GENERATIVE AI? NEW PERSPECTIVES IN CYBERSECURITY

Mário Marques da Silva [1,2,3] [0000−0002−5498−3220], Fernando Correia[1,2,3] [0000-0002-9530-1986], Héctor Dave Orrillo Ascama [1,2,3][0000−0003−1555−6821], Laercio Cruvinel Júnior[1,2,3][0000−0002−7672−3243]

[1] Department of Engineering and Computer Sciences, Universidade Autónoma de Lisboa, 1169-023 Lisboa, Portugal
[2] Autónoma TechLab, 1169-023 Lisboa, Portugal
[3] Ci2—Centro de Investigação em Cidades Inteligentes, 2300-313 Tomar, Portugal
mmsilva@autonoma.pt, fjcorreia@autonoma.pt, horrillo@autonoma.pt, lcruvinel@autonoma.pt

EXTENDED ABSTRACT

1 Introduction

The proliferation of digital systems in the 4th Industrial Revolution has introduced unprecedented cybersecurity challenges. This paper explores the implementation of Generative AI (GEN AI) in the context of cybersecurity, highlighting their applications in data analysis procedures and automatic control systems. By integrating Machine Learning (ML) for real-time detection and GEN AI for simulating advanced attack scenarios, we can detect cyberattacks at an early stage and minimize the impact on the systems functioning.

With the advent of interconnected devices and the Internet of Things (IoT), cyberattacks have become more sophisticated, necessitating advanced defence mechanisms. Artificial Intelligence (AI) has emerged as a cornerstone in this domain, offering tools for automation, real-time threat detection, and strategic simulation (Samuel, 1959). This paper examines the use of supervised learning methodologies, which rely on labelled datasets (Hinton et al., 2006), and GEN AI, which leverages unstructured data to simulate and preempt cyber threats (OpenAI, 2024). We conclude by discussing the implications for the future of cybersecurity and the anticipated dominance of AI-driven solutions by 2030.

According to Capodieci et al. (Capodieci et al., 2024), there are different types of GEN AI, such as Machine Learning (ML) and Deep Learning (DL), which are a subfield of AI. While ML depends on the human factor to classify and contextualize data, DL is a system with greater autonomy in processing unstructured data. Moreover, Large

Language Models (LLM) represent another type of AI technology, capable of understanding and generating written language (Lee, 2023).

In cybersecurity, given the volume of data that needs to be analysed to identify possible threats, the human factor can be a limiting factor. However, ML has been a fundamental technique in cybersecurity, excelling in detecting known attack patterns and automating routine tasks such as log analysis (Abadi et al., 2016), and to do that, human interaction is required.

DL, an advanced subset of ML, enhances these capabilities through multi-layer neural networks (Goodfellow et al., 2016). DL is part of unsupervised learning, which is effective in anomaly detection and clustering, identifying previously unseen threats without the need for labelled data.

GEN AI, powered by large language and multimedia models, extends these capabilities by creating synthetic data, simulating attack scenarios, and even mimicking social engineering tactics (OpenAI, 2024).

However, its potential misuse, such as generating deepfakes or creating malware variants, highlights the dual-edged nature of this technology (Brynjolfsson, 2014). According to the 2024 report by Capgemini (Capgemini, 2024), in the universe of 1000 organizations analysed, around 45% have encountered deep-fake attacks in the last 2 years, and in 43% there have been financial losses due to the exploitation of deepfake technology.

GEN AI operates along two main vectors: one focused on attacks and the other on testing and strengthening defences. In the attack vector, as mentioned earlier, GEN AI is used to create new forms of attacks on systems by exploiting possible unknown vulnerabilities. On the other hand, the same vector can be proactively applied to generate simulated attack scenarios, which enables the training and continuous improvement of detection systems. This duality not only expands the understanding of emerging threats but also strengthens cyber defences by anticipating and mitigating potential future attacks. In this way, it is possible to leverage the capabilities of SL in combination with GEN AI. While ML can learn to detect known attacks, GEN AI uses this acquired knowledge to infer patterns and generate simulated attack scenarios. This combined approach contributes to minimizing the impact of AI-driven cyberattacks (Siam et al., 2025).

2 Methodology

To explore the integration of ML and GEN AI in cybersecurity, we analyse:

- Applications of ML, including classification, regression, and anomaly detection. While Supervised Learning (SL) clearly identifies an attack type, Unsupervised Learning (UL) can be used to detect an event that is outside of the normal activity.
- GEN AI can be applied in cybersecurity to simulate phishing attacks, malware behaviours, and potential intrusion scenarios, enhancing the ability to predict and mitigate cyber threats.
- A hybrid defence strategy combining the strengths of both paradigms.

SL can be used as a classier, where labels are assigned to data to identify the categories to which that data belongs. For example, the classification of animal images by identifying which animal each image belongs to. It can also be used in regression when it involves the continuous prediction of numerical values based on the input data. Finally, object detection which focuses on identifying and locating multiple objects in images or video. Examples of SL can be found at car park license plate identification processes for automatic payment systems, weather forecasting, credit card fraud detection, among others. On the other hand, UL can be employed to identify anomalous patterns in real-time events. This normally requires the analysis of an operator, to identify or not as a threat.

ML makes it possible to collect data and produce data resulting from previous experiences, improving information quality. In other words, with increased knowledge, performance criteria can be optimized. SL makes it easier to solve various computational problems involving classification and regression procedures. Generically, ML allows for better control over data classes in AI training.

Currently, TBytes of data are produced, so classifying them using human supervision is challenging. Training an ML model requires a lot of computing time and supervision time. ML cannot perform all the complex tasks in the field of Machine Learning, which can be a disadvantage in the case of applying this model. The employment of ML and GEN AI must be weighed up and applied according to the volume of data to be processed and the results to be obtained.

Thus, a hybrid security system can be established, integrating not only classification, regression, and anomaly detection processes but also proactive mechanisms. These mechanisms enable the simulation of attack scenarios generated by GEN AI, which can be used to retrain and enhance the ML detection system, improving its ability to identify and respond to emerging threats.

Our study involved simulated environments using real-world datasets and generative models trained to mimic cyberattack scenarios. Metrics for evaluation included detection accuracy, response time, and the ability to pre-emptively counter novel threats.

3 Conclusions

The integration of ML and GEN AI offers a robust defense strategy against modern cyber threats. Institutions must invest in both technologies to stay ahead in the cybersecurity arms race. By 2030, it is anticipated that AI will play a critical role in resolving up to 95% of cybersecurity incidents. As the digital landscape continues to evolve, the necessity for adaptive, AI-driven solutions becomes increasingly apparent.

Organizations are increasing their use of AI in various cybersecurity areas. Given the growth in areas related to data analysis, application development or cloud operations, the use of a hybrid defense strategy is essential for the continuity of operations.

Finally, the paradigm shifts in the use of AI, through the increased use of GEN AI combined with ML, is a reality. However, the human component must always be

present in the equation to supervise systems' operation and guarantee the presence of ethics in the loop.

KEYWORDS: Cybersecurity and Machine Learning; Generative AI; Hybrid Defense Strategy; Anomaly Detection; Synthetic Data.

References

1. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng. Tensorflow: A system for large-scale machine learning, 2016. URL https://arxiv.org/abs/1605.08695.
2. E. Brynjolfsson. The second machine age : work, progress, and prosperity in a time of brilliant technologies. W.W. Norton & Company, 2014.
3. Capgemini. New defenses, new threats: What ai and gen ai bring to cybersecurity. Technical report, Capgemini Research Institute, 2024.
4. N. Capodieci, C. Sanchez-Adames, J. Harris, and U. Tatar. The impact of generative ai and llms on the cybersecurity profession. In 2024 Systems and Information Engineering Design Symposium (SIEDS), pages 448453, 2024. https://doi.org/10.1109/SIEDS61124.2024.10534674.
5. Goodfellow, Y. Bengio, and A. Courville. Deep Learning. The MIT Press, 2016. ISBN 9780262035613.
6. G. E. Hinton, S. Osindero, and Y.-W. Teh. A fast learning algorithm for deep belief nets. Neural Computation, 18(7):15271554, 2006. https://doi.org/10.1162/neco.2006.18.7.1527.
7. Lee. What are large language models and why are they important? [online] Available: https://blogs.nvidia.com/blog/2023/01/26/whatare-large-language-models-used-for/, Jan. 2023. URL https://blogs.nvidia.com/blog/2023/01/26/what-are-large-language-modelsused-for/.
8. OpenAI. Gpt-4 technical report, 2024. URL https://arxiv.org/abs/2303.08774.
9. L. Samuel. Some studies in machine learning using the game of checkers. IBM Journal of Research and Development, 3(3):210229, 1959. https://doi.org/10.1147/rd.33.0210.
10. A. Siam, M. M. Hassan, and T. Bhuiyan. Artificial intelligence for cybersecurity: A state of the art. In 2025 IEEE 4th International Conference on AI in Cybersecurity (ICAIC), pages 1-7, 2025. https://doi.org/10.1109/ICAIC63015.2025.10848980.