International Journal on Applications of Graph Theory in Wireless Ad hoc Networks and Sensor Networks (GRAPH-HOC) Vol.17, No.1, March 2025

THE IMPORTANCE OF MACHINE LEARNING TO INDIVIDUALS' RIGHTS TO PROTECT THEIR DATA

Basma Mohamed[1], Khaled Eid Abdel Moneim Abdel Fattah[2]

[1]Department of Information Systems, Giza Higher Institute for Managerial Sciences, Tomah, Egypt
[2]Department of Private Law, Giza Higher Institute for Managerial Sciences, Tomah, Egypt

ABSTRACT

The widespread use of machine learning to address a wide range of practical issues has resulted in the requirement to gather and analyze massive amounts of data, some of which are regarded as sensitive and personal, raising major data protection issues. The existing EU legislation on data protection and AI typically demand privacy-enhancing technologies (PETs) to reach a general trustworthiness and to secure personal data. In the increasingly digitally connected world of today, data privacy has become crucial. The growth of data in many different fields has made protecting sensitive information more crucial than ever. To improve data privacy in the digital era, this abstract examines state-of-the-art machine learning approaches. A workable solution to data privacy concerns is provided by the machine learning subset of artificial intelligence. This paper explores the potential of machine learning methods to improve confidentiality safeguards across a variety of applications. Machine learning models can provide proactive defense against cyber-attacks by analyzing massive amounts of data and potentially identifying vulnerabilities and breaches in real-time.

KEYWORDS

Data Protection; Data Security; Data Quality

1. INTRODUCTION

In the digital age, data has emerged as a vital resource that drives innovation, guides decisions, and facilitates the smooth operation of modern civilization. However, the abundance of data has led to hitherto unheard-of issues with privacy and confidentiality. The safeguarding of data against unauthorized access, breaches, and exploitation has become increasingly imperative due to the massive generation and sharing of sensitive information by individuals and companies [1]. This increasing concern has led to the development and application of state-of-the-art technologies, especially machine learning, to improve data confidentiality. This introduction lays the framework for a detailed examination of machine learning methods for preserving data secret by providing a broad overview of the evolving data privacy landscape and highlighting the crucial role of machine learning [2]. Furthermore, machine learning can enhance data anonymization techniques, facilitating the exchange of beneficial insights while protecting personal privacy [3].Margam [4] delved into the ethical considerations and challenges surrounding AI implementation in healthcare, including data privacy, algorithmic bias, and regulatory frameworks.

By eliminating learned knowledge concerning hazardous material, Graves et al. thorough empirical study [5] proved that these methods are, in fact, safe, effective, and maintain the

DOI: 10.5121/jgraphoc.2025.17101

performance of trained models. Presented by Foster et al. [6], selective synaptic dampening (SSD) is a revolutionary two-step, post hoc, retrain-free method for machine unlearning that is quick, effective, and does not require long-term training data storage. Unlearning was investigated for the regression problem by Tarun et al. [7], especially in deep learning models. Extensive research has been conducted on unlearning in classification and simple linear regression. A subset of the model parameters with the greatest semantic overlap was found on an individual sample level by Mehta et al. [8] using a variation of a new conditional independence coefficient called L-CODEC. A machine unlearning strategy for WSL by adjusting the model parameters was presented by Tang et al. [9]. The discussion and debate about artificial intelligence (AI) and its implications for human rights, including gaps and problems as well as implications for human rights principles, were the main topics of Rodrigues [10]. Sarker [11] sought to provide a technical point of reference for decision-makers in a variety of real-world scenarios and application sectors, as well as for experts in academia and business. Further information is available in the literature [12–15].

1.1. Significance of Data Protection

Businesses now need to manage a variety of data kinds in order to remain relevant. Organizations continue to dread data breaches as a result of low-security practices notwithstanding this growth over time[16]. Actually, 9,198,580,293 in a decade were cited in a recent report from a public tracking that analyzes the breach-level index. These violations have serious repercussions, including:

- Damaged brand reputation and lost sales
- Error in vital business information.
- Privacy invasion and identity theft.
- Hidden costs
- Consequences for the law; bankruptcy

2. OUTLINE OF AI AND ML IN DATA SECURITY

2.1. AI Applications for Data Protection

Businesses have recently learned that AI is a very useful tool for data security[17]. It is capable of predictive analytics, meaning it can examine historical data. By utilizing this strength, algorithms are able to anticipate security breaches before they occur. Organizations may remain ahead of cyber dangers by using this proactive approach.

Artificial Intelligence can also identify the kind of datasets and identify a missing data string with the use of neural languages. This can aid in preventing data loss as a result of human error. AI is also helpful in automating data cleansing and preparation with the aid of data model generation. This guarantees quick operation.

2.2. Application of ML to Data Security

The goal of machine learning is to gradually improve accuracy by mimicking human learning through the use of data and algorithms[18]. ML has applications in a wide range of business domains, including:

Fraud detection
Genuine or fraudulent image recognition
Processing natural language

Machine learning has numerous applications in the detection of suspicious activities related to data security[15]. ML is able to ingest and analyze data in order to identify trends, threats, and attack methods for cybercrime. This makes it easier for security teams to stay informed about such threats and take the appropriate precautions to fend them off before they get worse.

3. Applications of AI and ML for Data Protection:

Let's now examine a few more scenarios where artificial intelligence and machine learning are crucial to safeguarding our data[20]. The following are these cases:

- **Recognizing Dangers**:

  Machine learning and artificial intelligence are always guarding our data by keeping an eye out for cybercriminals attempting to gain access to our online systems[21].

- **Analysis of Behavior**:

  Artificial intelligence is capable of analyzing user behavior to discover baseline patterns and eliminate deviations that could point to unusual activity. An extra layer of security was added to the behavioral analysis, which focuses on alerting threats based on action rather than predefined signatures[22].

- **Finding Abnormalities**

  Machine learning and artificial intelligence are excellent at recognizing odd objects. Let's say, though, that something about our online technology seems off. Either AI or ML will then sound an alert to safeguard our internet technology in that scenario[23].

- **Monitoring in Real Time:**

  The monitoring of user behavior, system operations, and network traffic is made possible by artificial intelligence. By identifying these questionable activity, possible security risks can be mitigated, and a prompt reaction made possible[24].

- **The use of encryption:**

  Together with encryption, AI and ML communicate with each other using a secret code that they both know. To ensure that no one can comprehend or figure out this method, they make sure it is extremely strong[25].

- **Modular Security Protocols:**

  Machine learning algorithms can improve threat detection by continuing to gather knowledge from new data[26]. This development is particularly crucial in combating cyber threats that may change their method over time.

4. RESTRICTIONS AND CONSIDERATIONS

As indicated previously, the employment of AI and ML has come under investigation for numerous reasons. This comprises:

-Growing Ethical Concerns:

There are ethical concerns with using AI and ML to secure data, particularly with relation to privacy. Privacy rights may inadvertently be violated by automated models. This calls for giving ethical standards and laws considerable thought[27].

-Attacks by Adversaries Vulnerability:

AI/ML systems can be tricked by malicious actors through data manipulation. This action puts data security at serious risk since hackers could use security flaws to get around protections[28].

-Insufficient Understanding of the Black-Box Models:

Many AI professional models work as “black boxes .” This complicates the interpretation of their decision-making procedures. Transparency and trust are largely dependent on understanding how these models get to their findings. Particularly in delicate business domains like data protection, this is valid [29].

-The Algorithms' Biases:

Preconceptions from data training may be inherited by AI models, which could produce biased outcomes. Algorithmic decisions or biases in the data may affect some groups more than others. This may exacerbate injustice and encourage inequality[30].

-Abuse of Automation:

Excessive basing business choices on AI without comprehensive human evaluation of the offered facts might lead to unforeseen outcomes. Contextual comprehension depends on human judgment even with models' alleged accuracy. This is especially important in complicated situations where moral considerations are paramount[31].

-Risks Associated with Data Security and Privacy:

Large-scale dataset processing and archiving for AI training has emerged as a critical component of contemporary IT operations. However, it has posed extra security and privacy vulnerabilities[32]. For this reason, protecting this data from unwanted access is essential to averting possible breaches.

-Expensive Setup:

Even though it appears inexpensive, using AI/ML to protect data is expensive. This is because continuing maintenance, hiring qualified staff, and making investments in technology are all necessary[33]. This additional expense might put a major financial pressure on smaller companies.

-Absence of Prominent Structure:

One of the questions hanging around the use of AI models is the need for standardized frameworks for evaluation. This restriction may result in inconsistent compliance and security protocols[34].

5. DATA QUALITY LIMITATIONS

The unfortunate reality about AI models is that garbage in, rubbish out. They rely greatly on the quality of supplied data[35]. So, the accuracy of the model is directly impacted by any errors, incompleteness, or out-of-date information. These could result in incorrect findings. Avoiding these problems necessitates a thorough strategy that integrates technological advancements, moral concerns, and legal compliance to guarantee ethical and successful AI/ML application in data protection.

6. PROSPECTIVE PATTERNS AND ADVANCEMENTS

Although generative AI models have recently gained popularity in the IT sector, new developments have a big impact on data security[36]. Over time, the industry may experience a plethora of technological innovations, such as:

6.1. Explainable AI (XAI)

The business community longs for AI models that tackle the "black-box" issue by providing clear justifications for their choices. Therefore, the goal of ongoing research by AI developers is to make complex models easier to understand. Reaching this goal will promote understanding and trust in important applications like data security[37].

6.2. Preserving Privacy in Machine Learning

Models may be trained on decentralized data thanks to privacy-preserving methods like federated learning and homomorphic encryption. This guarantees comprehensive safeguards against the disclosure of private data. As encryption protocols and federated learning algorithms continue to progress, privacy preservation will become increasingly important. This will enable businesses to safely leverage insights from dispersed databases[38].

6.3. Security Orchestration, Automation, and Response (SOAR) in an Automated Form

Another expanding area of AI/ML use in business is the SOAR platforms. By automating incident response, it improves overall security posture and speeds up reaction times. Adaptive learning methods allow SOAR systems to continuously improve reaction strategies in response to changing threats. As a result, incident response will become more fluid and efficient[39].

6.4. Intelligence on Threats Powered by AI:

With 2,200 daily threats, averaging 39 per second, cyber threats remain on the horizon. Managing cybercrimes and related illegal actions online may undergo a significant evolution with the use of integrated artificial intelligence threat intelligence for real-time analysis.

Businesses anticipate that more sophisticated NLP and machine learning algorithms will improve their capacity to sort through enormous volumes of threat data.

As a result, threat identification would be more precise and timelier[40].

Other AI/ML trends are as follows:

Data normalization and de-identification
Moral leadership
Zero-trust framework for security
These technical developments will have a big impact on how quickly and widely organizations adopt AI/ML within their company.

7. DEEPER COMPARISON OF AI-BASED DATA PROTECTION TECHNIQUES

AI is transforming data security by providing more dynamic and advanced solutions than conventional techniques. Here is a more thorough comparison of various important AIbased data security methods:

7.1. Data Masking Driven by AI

Conventional data masking uses static techniques and preset rules to hide sensitive information; it frequently has to be updated by hand and has trouble keeping up with changing data patterns.

Data Masking Powered by AI:
Contextual Masking: By comprehending the relationships and context included in data, AI algorithms enable more accurate and practical masking while maintaining the usefulness of the data for testing and analytics.

AI's adaptive learning ensures that masking techniques continue to work well in dynamic data settings by constantly learning about new data kinds and patterns.
Scalability and Performance: AI is appropriate for real-time applications and large-scale data settings due to its ability to manage massive data quantities effectively.
Comparison: In terms of accuracy, flexibility, scalability, and automation, AI-powered data masking has a number of benefits over conventional techniques. For basic masking, classical techniques could be easier to execute, but AI offers more dynamic and reliable protection for complicated and changing data environments.

7.2. AI-Powered Threat Identification and Avoidance

Conventional security systems are susceptible to developing malware and zero-day attacks because they frequently rely on signature-based detection and preset rules.

Security Systems Driven by AI:

Behavioral Analysis: AI keeps an eye on user behavior patterns and spots irregularities, including odd login behavior or data access patterns, that can point to a security breach.

Intrusion Detection Systems (IDS): AI-driven IDS can minimize the harm caused by cyberattacks by instantly detecting and reacting to network intrusions.

Automated reaction: AI may reduce reaction time and human error by automating responses to specific threats, such as banning suspect IP addresses or quarantining phishing emails. Comparison: When compared to conventional systems, AI-driven security solutions have improved threat detection and prevention capabilities. AI is an essential weapon in the fight against changing cyberthreats because of its capacity to learn and adapt to new dangers, evaluate intricate patterns, and automate responses.

7.3. AI-Powered Authentication and Access Control

Conventional access control frequently depends on roles and static rules, which can be rigid and ineffective in dynamic settings.

Access Control Enhanced by AI

Risk-Based Authentication: AI is able to determine the degree of risk associated with each login attempt by taking into account a number of variables, including location, device, and user behavior, and then modify the authentication requirements appropriately.

Adaptive Access Control: AI may dynamically modify access rights in response to real-time risk assessments and continually monitor user activities, guaranteeing that only authorized users have access to critical information.

Comparatively speaking, AI-enhanced access control provides more flexible and detailed control over data access than conventional techniques. AI improves security and lowers the chance of unwanted access by analyzing contextual data and adjusting to shifting risk levels.

7.4. Data Loss Prevention (DLP) using AI

Conventional DLP: Frequently uses content matching and preset rules, which can be difficult to administer and result in false positives.
AI-Assisted DLP:
Content Understanding: AI is able to comprehend the meaning and context of data, which improves the accuracy of sensitive information detection and lowers false positives.

Anomaly detection: AI can identify odd user activity or data exfiltration patterns that can point to data leakage, hence preventing data loss.

Comparatively speaking, AI-powered DLP is more accurate and efficient than conventional DLP solutions. AI improves data protection and lowers the risk of data breaches by comprehending data context and identifying abnormalities.

Challenges and Considerations:

Data Availability and Quality: AI models need a lot of high-quality data to be trained, yet this might be difficult to get and categorize.

computer Resources: AI-based solutions, particularly when implemented on a big scale, may demand a substantial amount of computer resources.

Explainability and Transparency: It might be difficult to comprehend how complicated algorithms generate decisions in AI models, yet doing so is essential for accountability and trust.

8. CONCLUSION

In conclusion, the topic of how AI affects the safety of personal data is complicated and constantly changing. We've seen how advanced threat detection, real-time monitoring, and effective incident response have been made possible by AI and machine learning technologies, revolutionizing data security. However, there are drawbacks to integrating AI with data security, including bias and poor data quality, adversarial assaults, and opaque decision-making procedures.

To ensure the appropriate and ethical use of AI in data security and beyond, businesses must follow best practices such as completing comprehensive risk assessments, establishing robust data governance standards, and adopting continuous monitoring and model validation processes. Furthermore, privacy concerns in AI pose questions about discrimination, ethical use, and human control. The General Data Protection Regulation (GDPR) has a substantial impact on how personal data is used in AI systems.

Future privacy and artificial intelligence demand serious policy analysis and a change in the way that privacy regulation is thought of. To balance the advantages of AI with privacy concerns, it is imperative to give top priority to data governance, transparency, and individual privacy rights. We must adapt our rules and procedures in tandem with AI's continued advancement to guarantee its ethical and responsible application while safeguarding personal information.

REFERENCES

[1] Sharma, S., Alam, A. M., & Chen, K. (2021, September). Image disguising for protecting data and model confidentiality in outsourced deep learning. In *2021 IEEE 14th International Conference on Cloud Computing (CLOUD)* (pp. 71-77). IEEE.

[2] Fan, L. (2018). Image pixelization with differential privacy. In *Data and Applications Security and Privacy XXXII: 32nd Annual IFIP WG 11.3 Conference, DBSec 2018, Bergamo, Italy, July 16–18, 2018, Proceedings 32* (pp. 148-162). Springer International Publishing.

[3] Gallier, J. (2011). *Geometric methods and applications: for computer science and engineering* (Vol. 38). Springer Science & Business Media.

[4] Margam, R. The Intelligent Prescription: Unveiling The Ai-Driven Future Of Healthcare. *Journal ID*, *9339*, 1263.

[5] Graves, L., Nagisetty, V., & Ganesh, V. (2021, May). Amnesiac machine learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 35, No. 13, pp. 11516-11524).

[6] Foster, J., Schoepf, S., & Brintrup, A. (2024, March). Fast machine unlearning without retraining through selective synaptic dampening. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 11, pp. 12043-12051).

[7] Tarun, A. K., Chundawat, V. S., Mandal, M., & Kankanhalli, M. (2023, July). Deep regression unlearning. In *International Conference on Machine Learning* (pp. 33921-33939). PMLR.

[8] Mehta, R., Pal, S., Singh, V., & Ravi, S. N. (2022). Deep unlearning via randomized conditionally independent hessians. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1042210431).

[9] Tang, Y., Gao, Y., Luo, Y. G., Yang, J. C., Xu, M., & Zhang, M. L. (2024). Unlearning From Weakly Supervised Learning. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*.

[10] Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, *4*, 100005.

[11] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, *2*(3), 160.

[12] Batiha, M. I., Amin, M., Mohamed, B., & Jebril, H. I. (2024). Connected metric dimension of the class of ladder graphs. *Mathematical Models in Engineering,10* (2), 65–74.

[13] Almotairi, S., Alharbi, O., Alzaid, Z., Almutairi, B., & Mohamed, B. (2024). The Secure Metric Dimension of the Globe Graph and the Flag Graph. *Advances in Operations Research*, *2024*(1), 3084976.

[14] Batiha, I. M., Mohamed, B., & Jebril, I. H. (2024). Secure metric dimension of new classes of graphs. *Mathematical Models in Engineering*,*10*(3), 1-7 , https://doi.org/10.21595/mme.2024.24168 .

[15] Almotairi, S., Alharbi, O., Alzaid, Z., Hausawi, Y. M., Almutairi, J., & Mohamed, B. (2024). Computing the connected dominant metric dimension of different graphs. *Advances and Applications in Discrete Mathematics*, *41*(6), 505-520.

[16] Ulven, J. B., & Wangen, G. (2021). A systematic review of cybersecurity risks in higher education. *Future Internet*, *13*(2), 39.

[17] Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard business review*, *96*(1), 108-116.

[18] Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN computer science*, *2*(6), 420.

[19] Sagar, R., Jhaveri, R., & Borrego, C. (2020). Applications in security and evasions in machine learning: a survey. *Electronics*, *9*(1), 97.

[20] Tschider, C. A. (2018). Regulating the internet of things: discrimination, privacy, and cybersecurity in the artificial intelligence age. *Denv. L. Rev.*, *96*, 87.

[21] Zeadally, S., Adi, E., Baig, Z., & Khan, I. A. (2020). Harnessing artificial intelligence capabilities to improve cybersecurity. *Ieee Access*, *8*, 23817-23837.

[22] Liang, Y., Samtani, S., Guo, B., & Yu, Z. (2020). Behavioral biometrics for continuous authentication in the internet-of-things era: An artificial intelligence perspective. *IEEE Internet of Things Journal*, *7*(9), 9128-9143.

[23] Nair, M. M., Deshmukh, A., & Tyagi, A. K. (2024). Artificial intelligence for cyber security: Current trends and future challenges. *Automated Secure Computing for Next-Generation Systems*, 83-114.

[24] Mandal, V., Mussah, A. R., Jin, P., & Adu-Gyamfi, Y. (2020). Artificial intelligence-enabled traffic monitoring system. *Sustainability*, *12*(21), 9177.

[25] Knott, B., Venkataraman, S., Hannun, A., Sengupta, S., Ibrahim, M., & van der Maaten, L. (2021). Crypten: Secure multi-party computation meets machine learning. *Advances in Neural Information Processing Systems*, *34*, 4961-4973.

[26] Shah, V. (2021). Machine learning algorithms for cybersecurity: Detecting and preventing threats. *Revista Espanola de Documentacion Cientifica*, *15*(4), 42-66.

[27] Stahl, B. C., & Wright, D. (2018). Ethics and privacy in AI and big data: Implementing responsible research and innovation. *IEEE Security & Privacy*, *16*(3), 26-33.

[28] Blowers, M., & Williams, J. (2020, May). Artificial intelligence presents new challenges in cybersecurity. In *Disruptive Technologies in Information Sciences IV* (Vol. 11419, pp. 75-81). SPIE.

[29] Pierce, R., Sterckx, S., & Van Biesen, W. (2022). A riddle, wrapped in a mystery, inside an enigma: How semantic black boxes and opaque artificial intelligence confuse medical decision-making. *Bioethics*, *36*(2), 113-120.

[30] Kartal, E. (2022). A Comprehensive Study on Bias in Artificial Intelligence Systems: Biased or Unbiased AI, That's the Question!. *International Journal of Intelligent Information Technologies (IJIIT)*, *18*(1), 1-23.

[31] Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International journal of information management*, *57*, 101994.

[32] Samtani, S., Kantarcioglu, M., & Chen, H. (2021). A multi-disciplinary perspective for conducting artificial intelligence-enabled privacy analytics: Connecting data, algorithms, and systems. *ACM Transactions on Management Information Systems (TMIS)*, *12*(1), 1-18.

[33] Attaran, M., & Deb, P. (2018). Machine learning: the new 'big thing' for competitive advantage. *International Journal of Knowledge Engineering and Data Mining*, *5*(4), 277-305.

[34] Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., ... & Zhou, B. (2023). Trustworthy AI: From principles to practices. *ACM Computing Surveys*, *55*(9), 1-46.

[35] Fang, B., Yu, J., Chen, Z., Osman, A. I., Farghali, M., Ihara, I., ... & Yap, P. S. (2023). Artificial intelligence for waste management in smart cities: a review. *Environmental Chemistry Letters*, *21*(4), 1959-1989.

[36] Abumalloh, R. A., Nilashi, M., Ooi, K. B., Tan, G. W. H., & Chan, H. K. (2024). Impact of generative artificial intelligence models on the performance of citizen data scientists in retail firms. *Computers in Industry*, *161*, 104128.

[37] Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., ... & Hussain, A. (2024). Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation*, *16*(1), 45-74.

[38] Ma, J., Naas, S. A., Sigg, S., & Lyu, X. (2022). Privacy-preserving federated learning based on multi-key homomorphic encryption. *International Journal of Intelligent Systems*, *37*(9), 5880-5901.

[39] ReddyAyyadapu, A. K. (2023). Optimizing Incident Response in Cloud Security with Ai And Big Data Integration. *Chelonian Research Foundation*, *18*(2), 2212-2225.

[40] Arif, H., Kumar, A., Fahad, M., & Hussain, H. K. (2024). Future Horizons: AI-Enhanced Threat Detection in Cloud Environments: Unveiling Opportunities for Research. *International Journal of Multidisciplinary Sciences and Arts*, *3*(1), 242-251.