International Journal of Society Reviews (INJOSER) e-ISSN: 3030-802X
Vol. 2 No. 7 2025, pages: 374~380

ETHICS AND SECURITY IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING: CURRENT PERSPECTIVES IN COMPUTING

Loso Judijanto
IPOSS Jakarta, Indonesia
losojudijantobumn@gmail.com

Alim Hardiansyah
Universitas Sultan Ageng Tirtayasa
alim.hardiansyah@untirta.ac.id

Opan Arifudin
STIT Rakeyan Santang
opan.arifudin@yahoo.com

Abstract

The application of artificial intelligence (AI) and machine learning (ML) continues to show significant development and make a huge impact in various fields, from industry to daily life. However, these technological advancements also come with ethical and security challenges that cannot be ignored. This article explores crucial issues related to data privacy, bias in AI systems, and potential security threats that can arise from the use of AI and ML. From an advanced computing perspective, a comprehensive and collaborative approach between developers, users, and regulators is needed to ensure that the benefits of AI and ML can be optimised without compromising ethical and security standards. This research underscores the importance of developing robust security policies and technologies, as well as implementing methods to reduce bias and protect data privacy to support fair, responsible and safe use of AI.

Keywords: Ethics, Security, Artificial Intelligence, Machine Learning, Computing.

Introduction

In the ever-evolving digital era, Artificial Intelligence (AI) and Machine Learning (ML) have become key technologies that drive innovation in various sectors. From industrial automation, healthcare, to personalising user experience on digital platforms, AI and ML offer great potential to improve the efficiency and effectiveness of processes and services (Caner & Bhatti, 2020) .

The increasing use of artificial intelligence (AI) and machine learning (ML) in various sectors has a pivotal role in driving efficiency, innovation, and improving service quality. In the healthcare sector, AI and ML are used to diagnose diseases more accurately and quickly, design more personalised treatments, and optimise hospital management. In industry, these technologies improve productivity through production process automation, predictive maintenance, and supply chain optimisation (Bad et al., 2020) . Meanwhile, in the financial sector, AI and ML help analyse risks, detect fraud, and provide more targeted financial services. In addition, the potential of AI and ML in smart city development, customised education, and environmental preservation makes it an

important pillar in sustainable development and innovation towards a smarter and more effective future. However, along with its widespread application, a number of issues related to ethics and security also arise, which are equally important to be considered (Saghiri et al., 2022) .

One of the main issues is algorithm bias that can reinforce social injustice. Algorithms used in AI and ML are often trained using historical data that may contain biases, so the resulting decisions can be unfair to certain groups. For example, gender-biased job recruitment algorithms or facial recognition algorithms that are more accurate for individuals with a certain skin colour (Montasari, 2024)

In addition, transparency in the algorithmic decision-making process is a major concern. Many AI models, especially highly complex ones such as deep learning, operate as black-boxes that are difficult for humans to understand. The inability to explain how a decision is made fuels concerns about the responsibility and reliability of these systems (Boopathi et al., 2023) .

Security is another crucial aspect that has been impacted by advances in AI and ML. Attacks on AI models, such as adversarial attacks that can trick image or speech recognition systems, show that the reliability of results from AI systems can be easily manipulated. This lack of robustness of AI models opens the door for potential exploits that can be fatal to critical systems such as autonomous vehicles or cybersecurity applications (Siriwardhana et al., 2021) .

In addition, the misuse of AI for malicious purposes such as the creation of deepfakes, cyber-scams, and the spread of disinformation also adds a layer of complexity in securing this technology. Efforts to develop safe and ethical AI require consideration not only from the technical but also from the legal realm and various attempts have been made by the scientific community, industry, and government to address these challenges (Bertino et al., 2021) . There is a strong push to develop ethical AI and ML through the application of the principles of fairness, accountability, and transparency (FAT). In addition, initiatives to develop security standards and compliance with regulations such as the General Data Protection Regulation (GDPR) are also seen as important steps towards the responsible application of technology (Syed & ES, 2024) .

However, the dynamics and complexity of this issue require a sustained and collaborative approach. Further studies and adaptive policies are needed to ensure that AI and ML can contribute positively without compromising ethical and security values.

Research Methods

The study in this research uses the literature method. The literature research method is a systematic approach in collecting, evaluating, and synthesising information from various sources relevant to the research topic. This research involves a comprehensive literature search, including scientific journals, books, reports, and other publications, to identify related findings and compile an in-depth understanding of the subject under study (Sahar, 2008) ; (Arikunto;, 2000) . This process includes not only

summarising and analysing the material, but also critiquing and comparing different perspectives and previous findings to present a more holistic and informative narrative. Thus, literature research methods help researchers understand recent developments in a particular field of study, identify research gaps, and establish a strong theoretical basis for further research (Fadli, 2021) .

Results and Discussion

Ethical Issues in AI and ML Development

The development of artificial intelligence (AI) and machine learning (ML) brings with it various ethical issues that developers, researchers, and policymakers need to be aware of. Firstly, bias in data and algorithms is one of the main challenges. AI models may reflect, or even reinforce, human prejudices hidden in the training data. For example, if the historical data used to train an AI system is unbalanced or has racial, gender, or social biases, then the results produced by the system can also be biased. This can lead to discrimination in decision-making, such as in job recruitment, loan appraisals, and legal decisions (Bharadiya ., 2023)

Secondly, the issue of data privacy and security is also an important concern. AI and ML typically rely on large amounts of data to learn and operate effectively. This demands the collection, storage, and analysis of personal data that can compromise individual privacy if not managed carefully. Unauthorised use of data, data leakage, and exploitation of data for unethical purposes can violate privacy rights and pose serious security risks (Koshiyama et al., 2022) .

Third, transparency and accountability in AI systems is another challenge. AI algorithms generally function as elusive "black boxes", making it unclear how decisions are made. This makes it difficult for users and affected parties to assess or question AI decisions. Therefore, it is important to develop transparent AI, where the decision-making process can be explained in detail, and there are adequate mechanisms for accountability in case of errors or misuse (Huang et al., 2022) .

Fourth, there are concerns about the social and economic impacts of AI and ML. Automation driven by these technologies may replace human jobs, causing major changes in the labour structure and unemployment. In the long run, this could affect economic and social inequality if the transition is not managed well. Therefore, labour inclusion and retraining strategies need to be developed to mitigate these negative effects (Mohamed et al., 2023) .

Finally, there is the issue of moral and legal responsibility in automated decisions made by AI systems. Who is responsible when an AI makes a wrong or harmful decision? This demands the development of clear legal and ethical frameworks to ensure that the use of AI is in line with human values and applicable laws (Jin & Ye, 2022) . The adoption of AI should be done with the interests of society in mind, and with strict enforcement of regulations to prevent potentially harmful misuse of the technology.

Security challenges arising from the use of AI and ML

The use of artificial intelligence (AI) and machine learning (ML) presents multiple and complex security challenges. Firstly, adversarial attacks against AI models are an increasingly real threat. These attacks involve manipulating input data to trick AI models into making incorrect predictions. For example, a slightly modified image can cause a face or object recognition system to fail to recognise it correctly. This poses a great risk especially in critical applications, such as cybersecurity, autonomous vehicles, and surveillance (Al-Worafi ., 2023)

Secondly, the collection and storage of large amounts of data by AI and ML also pose a risk of privacy breaches and data breaches. Many AI systems require extensive personal data to operate efficiently, such as health data, financial information, and online behaviour. If these data are not properly protected, they are vulnerable to misuse by hackers or irresponsible parties. Data leaks can have serious consequences, ranging from identity theft to financial and reputational losses (Ahmed et al., 2021) .

Thirdly, the issue of data integrity is also a big challenge in the context of AI and ML. AI models rely heavily on the quality of data used for training and operation. Corrupted, incomplete, or incorrect data can result in biased and inaccurate results. In addition, malicious actors may try to insert fake data or tamper with datasets to disrupt model performance, which can have a serious impact on the reliability of AI systems in making decisions (Santosh & Gaur, 2022) .

Fourth, the security of AI models used in edge computing is also an important concern. AI models deployed on devices such as smartphones, smart cameras, and IoT sensors often have limited storage and computing space, making it more difficult to implement strong security measures. In addition, because these devices are often connected to wider networks, they can be entry points for cyberattacks that target connected networks or systems (Cobianchi et al., 2022) .

Finally, the proliferation of deepfakes and other AI manipulation technologies adds another layer of complexity to the security challenge. Deepfakes can be used to create highly convincing but fake video or audio content, which can be used for disinformation, blackmail or defamation. These technologies demand the development of more sophisticated detection methods and strict regulations to police the use and dissemination of such content. Overall, addressing the security challenges arising from the use of AI and ML requires a holistic approach, involving technology, policy, and public awareness and education.

Conclusion

The use of artificial intelligence (AI) and machine learning (ML) is having a huge impact on many aspects of life, but also poses significant ethical and security challenges. From a current perspective in computing, one of the main concerns is how these technologies may affect privacy and personal data. With AI and ML increasingly relying on

big data to train their models, the risk of privacy breaches and possible misuse of data by irresponsible parties is an issue that cannot be ignored. Developers and users of these technologies need to ensure that data is collected and managed appropriately, and protected from unauthorised access.

In addition, the issue of bias in AI and ML also requires serious attention. Models trained with unrepresentative or biased data can result in unfair and discriminatory decisions. This can be especially detrimental when applied in critical sectors such as recruitment, bank credit, and law enforcement. Therefore, it is imperative to develop methods that can detect and reduce bias in AI models as well as apply strict ethical standards in the process of developing and applying these technologies.

The security of AI systems is also a critical aspect that requires strong security policies and technologies. Cyber-attacks targeting AI models, data manipulation, and the use of technologies such as deepfakes for malicious purposes, require a comprehensive approach to protect these systems from threats. Stakeholders in both technology and regulation must work together to ensure that AI and ML innovations can be used safely and responsibly, supporting human progress without compromising ethics and security.

References

Ahmed, S., Hossain, M., Kaiser, M., Noor, M., & ... (2021). Artificial intelligence and machine learning for ensuring security in smart cities. *... -Driven Mining, Learning ...*, *Query date: 2025-02-07 08:58:43.* https://doi.org/10.1007/978-3-030-72139-8_2

Al-Worafi, Y. (2023). Artificial intelligence and machine learning for drug safety. *Technology for Drug Safety: Current Status and Future ...*, *Query date: 2025-02-07 08:58:43.* https://doi.org/10.1007/978-3-031-34268-4_7

Arikunto;, S. (2000). *Research Management* (Jakarta). Rineka Cipta. //172.0.0.24%2Felibrary%2Findex.php%3Fp%3Dshow_detail%26id%3D2341%26keywords%3D

Bertino, E., Kantarcioglu, M., Akcora, C., & ... (2021). AI for Security and Security for AI. *... Application Security ...*, *Query date: 2025-02-07 08:58:43.* https://doi.org/10.1145/3422337.3450357

Bharadiya, J. (2023). AI-driven security: How machine learning will shape the future of cybersecurity and web 3.0. *American Journal of Neural Networks and ...*, *Query date: 2025-02-07 08:58:43.* https://www.researchgate.net/profile/Jasmin-Bharadiya-4/publication/371562853_AI-Driven_Security_How_Machine_Learning_Will_Shape_the_Future_of_Cybersecurity_and_Web_30/links/6489d4e2712bd82962231476/AI-Driven-Security-How-Machine-Learning-Will-Shape-the-Future-of-Cybersecurity-and-Web-30.pdf

Boopathi, S., Pandey, B., & Pandey, D. (2023). Advances in artificial intelligence for image processing: Techniques, applications, and optimisation. *Handbook of Research on Thrust...*, *Query date: 2025-02-07 08:58:43.* https://www.igi-global.com/chapter/advances-in-artificial-intelligence-for-image-processing/328027

Bad, B., Ekmekci, P., & Arda, B. (2020). A critical perspective on guidelines for responsible and trustworthy artificial intelligence. *Medicine, Health Care and Philosophy*, *Query date: 2025-02-07 08:58:43.* https://doi.org/10.1007/s11019-020-09948-1

Caner, S., & Bhatti, F. (2020). A conceptual framework on defining business strategy for artificial intelligence. *Contemporary Management Research*, *Query date: 2025-02-07 08:58:43.* https://cmr-journal.org/article/view/19970

Cobianchi, L., Verde, J., Loftus, T., & ... (2022). Artificial intelligence and surgery: Ethical dilemmas and open issues. *Journal of the ...*, *Query date: 2025-02-07 08:58:43.* https://journals.lww.com/journalacs/fulltext/2022/08000/Artificial_Intelligence_and_Surgery__Ethical.17.aspx?context=FeaturedArticles&collectionId=1

Fadli, M. R. (2021). Understanding the design of qualitative research methods. *HUMANIKA*,*21* (1), 33-54. https://doi.org/10.21831/hum.v21i1.38075

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2022). An overview of artificial intelligence ethics. *... on Artificial Intelligence*, *Query date: 2025-02-07 08:58:43.* https://ieeexplore.ieee.org/abstract/document/9844014/

Jin, K., & Ye, J. (2022). Artificial intelligence and deep learning in ophthalmology: Current status and future perspectives. *Advances in Ophthalmology Practice and Research*, *Query date: 2025-02-07 08:58:43.* https://www.sciencedirect.com/science/article/pii/S2667376222000555

Koshiyama, A., Kazim, E., & Treleaven, P. (2022). Algorithm auditing: Managing the legal, ethical, and technological risks of artificial intelligence, machine learning, and associated algorithms. *Computer*, *Query date: 2025-02-07 08:58:43.* https://ieeexplore.ieee.org/abstract/document/9755237/

Mohamed, N., Almazrouei, S., Oubelaid, A., & ... (2023). Artificial intelligence (AI) and machine learning (ML)-based information security in electric vehicles: A review. *2023 5th Global ...*, *Query date: 2025-02-07 08:58:43.* https://ieeexplore.ieee.org/abstract/document/10175817/

Montasari, R. (2024). Addressing Ethical, Legal, Technical, and Operational Challenges in Counterterrorism with Machine Learning: Recommendations and Strategies. *... and the International Security in the Fourth Industrial ...*, *Query date: 2025-02-07 08:58:43.* https://doi.org/10.1007/978-3-031-50454-9_10

Saghiri, A., Vahidipour, S., Jabbarpour, M., & ... (2022). A survey of artificial intelligence challenges: Analysing the definitions, relationships, and evolutions. *Applied Sciences*, *Query date: 2025-02-07 08:58:43.* https://www.mdpi.com/2076-3417/12/8/4054

Sahar, J. (2008). A critique of qualitative research. *Indonesian Nursing Journal*,*12* (3), 197-203. https://doi.org/10.7454/jki.v12i3.222

Santosh, K., & Gaur, L. (2022). *Artificial intelligence and machine learning in public healthcare: Opportunities and societal impact*. books.google.com. https://books.google.com/books?hl=en&lr=&id=bWRXEAAAQBAJ&oi=fnd&pg=PR7&dq=ethics+security+in+artificial+intelligence+machine+learning&ots=2IGt_MIATv&sig=r9KXVWmGC-6EehCZ3fD9DOUHRZ4

Siriwardhana, Y., Porambage, P., & ... (2021). AI and 6G security: Opportunities and challenges. *2021 Joint European ...*, *Query date: 2025-02-07 08:58:43.* https://ieeexplore.ieee.org/abstract/document/9482503/

Syed, F., & ES, F. (2024). AI in Securing Pharma Manufacturing Systems Under GxP Compliance. *... in Cybersecurity and Artificial Intelligence*, *Query date: 2025-02-07 08:58:43.* https://www.researchgate.net/profile/Amelia-Ethan/publication/385720826_AI_in_Securing_Pharma_Manufacturing_Systems_

Under_GxP_Compliance/links/6732bfb268de5e5a30739a95/AI-in-Securing-Pharma-Manufacturing-Systems-Under-GxP-Compliance.pdf