Ethical Considerations in Implementing Explainable AI for Healthcare Decision Support

Swetha Arra[1] and Dr. Ramesh Kumar[2]

Research Scholar, Department of Computer Science and Engineering[1]

Research Guide, Department of Computer Science and Engineering[2]

NIILM University, Kaithal, India

Abstract: *As artificial intelligence (AI) is making significant advancements in the healthcare sector, transparency and dependability in AI-driven clinical decision-making are becoming increasingly crucial. Explainable AI (XAI), which provides understandable explanations for the predictions and recommendations made by AI systems, is one approach put out to address this issue. This research examines the use of XAI techniques in healthcare and its effects on legal compliance, professional confidence, and patient outcomes. In the context of healthcare research, the possible benefits and drawbacks of certain XAI approaches—such as rule-based models, decision trees, and model-agnostic approaches—are discussed. The integration of XAI into healthcare systems is also discussed in terms of future directions, challenges, and ethical considerations.*

Keywords: Explainable AI, Healthcare Transparency

REFERENCES

[1]. Hamamoto, R. (2021). Application of artificial intelligence for medical research. In Biomolecules (Vol. 11, Issue 1, pp. 1–4). MDPI. https://doi.org/10.3390/biom11010090.

[2]. Anton, N., Doroftei, B., Curteanu, S., Catãlin, L., Ilie, O. D., Târcoveanu, F., & Bogdănici, C. M. (2023). Comprehensive Review on the Use of Artificial Intelligence in Ophthalmology and Future Research Directions. In Diagnostics (Vol. 13, Issue 1). MDPI. https://doi.org/10.3390/diagnostics13010100.

[3]. Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P., & Parasa, S. (2022). On evaluation metrics for medical applications of artificial intelligence. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-09954-8.

[4]. Car, L. T., Dhinagaran, D. A., Kyaw, B. M., Kowatsch, T., Joty, S., Theng, Y. L., & Atun, R. (2020). Conversational agents in health care: Scoping review and conceptual analysis. In Journal of Medical Internet Research (Vol. 22, Issue 8). JMIR Publications Inc. https://doi.org/10.2196/17158.

[5]. le Glaz, A., Haralambous, Y., Kim-Dufor, D. H., Lenca, P., Billot, R., Ryan, T. C., Marsh, J., DeVylder, J., Walter, M., Berrouiguet, S., & Lemey, C. (2021). Machine learning and natural language processing in mental health: Systematic review. In Journal of Medical Internet Research (Vol. 23, Issue 5). JMIR Publications Inc. https://doi.org/10.2196/15708.

[6]. Laptev, V. A., Ershova, I. V., & Feyzrakhmanova, D. R. (2022). Medical Applications of Artificial Intelligence (Legal Aspects and Future Prospects). Laws, 11(1). https://doi.org/10.3390/laws11010003

[7]. Guo, Y., Hao, Z., Zhao, S., Gong, J., & Yang, F. (2020). Artificial intelligence in health care: Bibliometric analysis. Journal of Medical Internet Research, 22(7). https://doi.org/10.2196/18228.

[8]. Alami, H., Lehoux, P., Auclair, Y., de Guise, M., Gagnon, M. P., Shaw, J., Roy, D., Fleet, R., Ahmed, M. A., & Fortin, J. P. (2020). Artificial intelligence and health technology assessment: Anticipating a new level of complexity. In Journal of Medical Internet Research (Vol. 22, Issue 7). JMIR Publications Inc. https://doi.org/10.2196/17707.

[9]. Sharma, M., Savage, C., Nair, M., Larsson, I., Svedberg, P., & Nygren, J. M. (2022). Artificial Intelligence Applications in Health Care Practice: Scoping Review. In Journal of Medical Internet Research (Vol. 24, Issue 10). JMIR Publications Inc. https://doi.org/10.2196/40238.

[10]. Ma, X., Niu, Y., Gu, L., Wang, Y., Zhao, Y., Bailey, J., & Lu, F. (2021). Understanding adversarial attacks on deep learning based medical image analysis systems. Pattern Recognition, 110.

[11]. Ward, A., Sarraju, A., Chung, S., Li, J., Harrington, R., Heidenreich, P., Palaniappan, L., Scheinker, D., & Rodriguez, F. (2020). Machine learning and atherosclerotic cardiovascular disease risk prediction in a multi-ethnic population. Npj Digital Medicine, 3(1). https://doi.org/10.1038/s41746- 020-00331-1.

[12]. Li, W., Song, Y., Chen, K., Ying, J., Zheng, Z., Qiao, S., Yang, M., Zhang, M., & Zhang, Y. (2021). Predictive model and risk analysis for diabetic retinopathy using machine learning: A retrospective cohort study in China. BMJ Open, 11(11). https://doi.org/10.1136/bmjopen-2021-050989.

[13]. Bharati, S., Mondal, M. R. H., & Podder, P. (2023). A Review on Explainable Artificial Intelligence for Healthcare: Why, How, and When? IEEE Transactions on Artificial Intelligence. https://doi.org/10.1109/TAI.2023.3266418.

[14]. Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Medical Informatics and Decision Making, 20(1). https://doi.org/10.1186/s12911- 020-01332-6.

[15]. Hasan Sapci, A., & Aylin Sapci, H. (2020). Artificial intelligence education and tools for medical and health informatics students: Systematic review. In JMIR Medical Education (Vol. 6, Issue 1). JMIR Publications Inc. https://doi.org/10.2196/19285.

[16]. Salman, M., Ahmed, A. W., Khan, A., Raza, B., & Latif, K. (2017). Artificial Intelligence in Bio-Medical Domain An Overview of AI Based Innovations in Medical. In IJACSA) International Journal of Advanced Computer Science and Applications (Vol. 8, Issue 8). www.ijacsa.thesai.org

[17]. Rundo, L., Tangherloni, A., & Militello, C. (2022). Artificial Intelligence Applied to Medical Imaging and Computational Biology. In Applied Sciences (Switzerland) (Vol. 12, Issue 18). MDPI. https://doi.org/10.3390/app12189052

[18]. Choudhury, N., & Begum, S. A. (2016). A Survey on Case-based Reasoning in Medicine. In IJACSA) International Journal of Advanced Computer Science and Applications (Vol. 7, Issue 8). www.ijacsa.thesai.org

[19]. Gerke, S., Minssen, T., & Cohen, G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. In Artificial Intelligence in Healthcare (pp. 295–336).Elsevier. https://doi.org/10.1016/B978-0-12-818438- 7.00012-5

[20]. Bhattad, P. B., & Jain, V. (2020). Artificial Intelligence in Modern Medicine – The Evolving Necessity of the Present and Role in Transforming the Future of Medical Care. Cureus. https://doi.org/10.7759/cureus.8041

[21]. Alam, M. N., Singh, V., Kaur, M. R., Kabir, M. S. (2023). Big Data: An overview with Legal Aspects and Future Prospects. In Journal of Emeging Technologies and Innovative Research (Vol. 10, Issue 5).

[22]. Alam, M. N., Kaur, B., & Kabir, M. S. (1994). Tracing the Historical Progression and Analyzing the Broader Implications of IoT: Opportunities and Challenges with Two Case Studies. networks (eg, 4G, 5G), 7, 8.

[23]. Kabir, M. S., & Alam, M. N. (2023). IoT, Big Data and AI Applications in the Law Enforcement and Legal System: A Review.

[24]. Guo, J., & Li, B. (2018). The Application of Medical Artificial Intelligence Technology in Rural Areas of Developing Countries. In Health Equity (Vol. 2, Issue 1, pp. 174–181). Mary Ann Liebert Inc. https://doi.org/10.1089/heq.2018.0037.

Copyright to IJARSCT
www.ijarsct.co.in

ISSN
2581-9429
IJARSCT

652