Trust and Transparency in Generative and Non-Generative models for healthcare.

YANNAM SRAVAN KUMAR
Technological Uniiversity Dublin

hcaim human centred artificial intelligence masters

Artificial Intelligence (AI) is transforming healthcare, with generative and non-generative models playing key roles. Generative AI can create synthetic medical data for research, diagnosis, and treatment but raises concerns about accuracy, bias, and privacy. In contrast, non-generative models analyze existing data, aiding diagnosis and risk prediction while facing challenges in transparency and trust.
This study compares the human-centred aspects of generative and non-generative AI in healthcare, focusing on ethical, transparency, and regulatory challenges. It examines trust among healthcare professionals and patients, assesses risks like bias and misinformation, and explores regulatory compliance. Additionally, it evaluates their impact on patient outcomes to determine if the benefits of generative AI justify its risks. Findings will support responsible AI development, ensuring ethical integrity and regulatory adherence in healthcare.

INTRODUCTION

AI is transforming healthcare, offering new solutions for diagnosis, treatment, and patient management. Generative AI, capable of creating synthetic medical data, enhances research and treatment personalization but raises concerns regarding reliability, ethics, and regulation. Non-generative AI, which analyzes real-world data, aids in diagnosis and monitoring but faces challenges with trust, transparency, and bias.

Problem Statement

Both generative and non-generative AI models face human-centred challenges in trust, ethics, and regulatory compliance. Generative models present risks like bias and privacy concerns, while non-generative models struggle with interpretability and transparency. There is a need to evaluate and compare these challenges for responsible AI integration in healthcare.

Research Question

How do generative and non-generative AI models compare in terms of trust, ethical considerations, and regulatory challenges in healthcare?

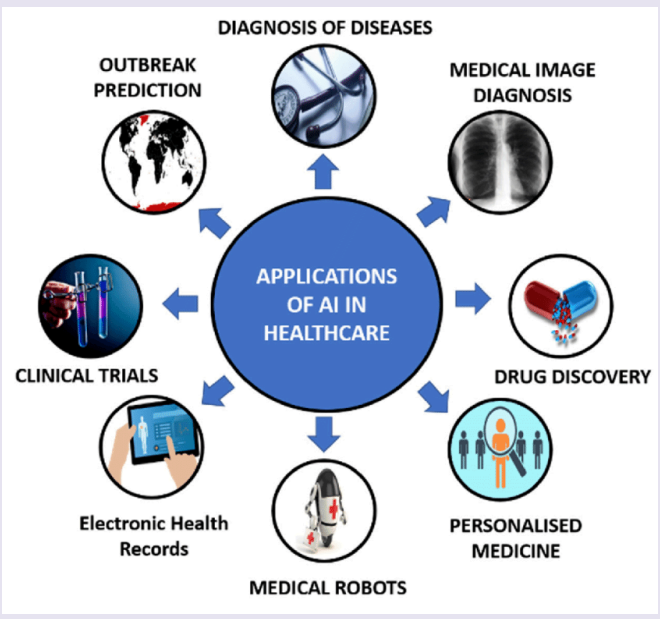

FIG 1 . IIMAGE SHOWING THE APPLICATIONS OF AI IN HEALTHCARE

PRILIMINARY CONSIDERATIONS

This study will compare generative and non-generative AI models in healthcare, focusing on trust, transparency, and regulatory challenges. The research will be conducted over six months, using public medical datasets (e.g., MIMIC-III, NIH Chest X-rays) and synthetic data generated by GANs and VAEs.

AI Models: Generative models (GANs, VAEs) will create synthetic medical data, while CNN-based diagnostic models will analyze real patient data. Explainability techniques (SHAP, Grad-CAM) will be employed to assess model interpretability.

Participants: 30 healthcare professionals (doctors, radiologists, AI researchers) will evaluate AI outputs for accuracy and ethical concerns.

Analysis: The study will use quantitative analysis (accuracy, bias, transparency metrics) and qualitative analysis (expert feedback) to assess the effectiveness and ethical challenges of both AI models. Statistical tools in Python will be used for analysis, ensuring robust evaluation of the models' real-world applicability.

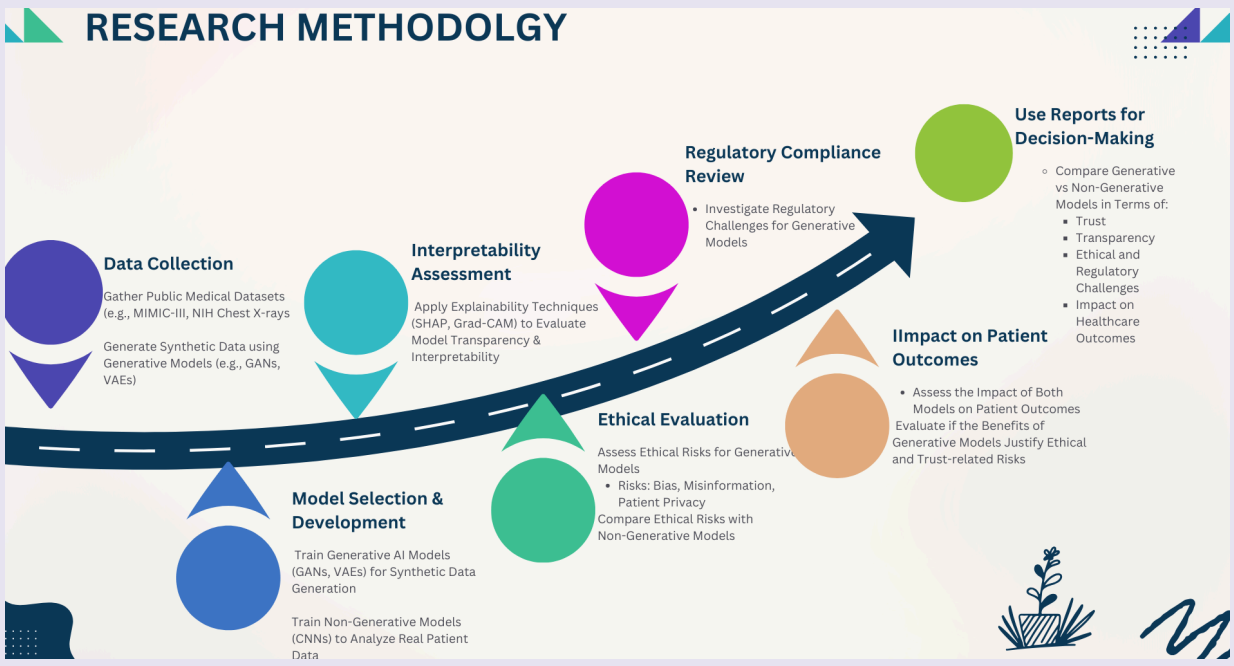

FIG 2 PROPOSED METHODOLY FOR THE THESIS

LITERATURE REVIEW AND METHODOLOGY

Artificial Intelligence (AI) is revolutionizing healthcare, with generative and non-generative models offering distinct benefits and challenges. Generative AI creates synthetic medical images and anonymized patient data, improving research and model training but raising concerns around data authenticity, bias, and trust (Ghassemi et al., 2021). It also lacks verifiability, making transparency difficult (Tschandl et al., 2020). In contrast, non-generative models, like deep learning diagnostic tools, analyze real patient data for disease prediction and monitoring. Though more widely accepted, they face trust issues due to black-box decision-making and bias (Rajpurkar et al., 2018), with XAI methods offering partial improvements (Samek et al., 2019).

Both AI types present regulatory and ethical concerns. Generative AI faces risks related to data integrity and misinformation, while non-generative models raise issues of bias and fairness (Ching et al., 2018). Current regulatory frameworks are insufficient to address these challenges (He et al., 2022).

This study compares how generative and non-generative AI models fare in trust, ethics, and regulatory compliance, aiming to inform responsible AI adoption in healthcare

This study compares generative and non-generative AI models in healthcare, focusing on trust, transparency, and regulatory challenges. Using public datasets (e.g., MIMIC-III, NIH Chest X-rays) and synthetic data from GANs and VAEs, generative models will create synthetic medical data, while CNN-based models will analyze real patient data. Thirty healthcare professionals will evaluate AI outputs for accuracy and ethical concerns. The study will combine quantitative (accuracy, bias, transparency) and qualitative (expert feedback) analysis, using Python tools for robust evaluation of both models' effectiveness and ethical challenges in real-world healthcare settings.

CONCLUSION

This study aims to compares the trust, transparency, and ethical challenges of generative and non-generative AI models in healthcare. Generative AI offers innovative solutions but raises concerns around bias, misinformation, and trust, while non-generative models, though more accepted, face issues with interpretability and bias. The research aims to assess the clinical applicability, ethical risks, and regulatory challenges of both models. Findings will contribute to developing more transparent, ethical AI systems, supporting healthcare professionals and policymakers in making informed decisions. Ultimately, the study will help shape future research and regulatory policies to ensure responsible AI integration in healthcare..

REFERENCES

- Ching, T., Himmelstein, D. S., & Beaulieu-Jones, B. K. (2018). Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface, 15(141), 20170387.
- Ghassemi, M., Naumann, T., & Schulam, P. (2021). A review of AI in healthcare: History, challenges, and opportunities. Artificial Intelligence in Medicine, 110, 28-45.
- He, J., Xu, W., & Zhang, J. (2022). Regulatory challenges and ethical issues of AI in healthcare. Journal of AI Research, 45(1), 25-40.
- Rajpurkar, P., Hannun, A. Y., & Yim, J. (2018). Cardiologist-level arrhythmia detection with convolutional neural networks. Nature Medicine, 24(4), 511-517.
- Samek, W., Wiegand, T., & Müller, K. R. (2019). Explainable artificial intelligence: Understanding, visualizing, and interpreting deep learning models. ITU Journal on Communications and Networks, 11(5), 1-24.

Acknowledgements. The HCAIM (the Human-Centred AI Master's Programme) Project is Co-Financed by the Connecting Europe Facility of the European Union Under Grant №CEF-TC-2020-1 Digital Skills 2020-EU-IA-0068. This Poster Was Created As Part of the Blended IP Event Organized Under Erasmus + Programme of the European Union.