Applications and Challenges of Generative AI in the Medical Field

Zhenchao Tao
http://orcid.org/0000-0001-8142-9164
The First Affiliated Hospital of USTC West District, China

Lei Li
http://orcid.org/0000-0002-5374-7293
Hefei University of Technology, China

ABSTRACT

The medical field is one of the important vertical application areas for generative artificial intelligence (AI), which has been employed in disease diagnosis, disease prediction, and clinical research, bringing tremendous opportunities for the demand and development of medical data and providing convenience for patients, healthcare professionals, AI practitioners, and regulatory authorities. However, the integration of generative AI in the medical field presents numerous challenges related to risk perception, value shaping, organizational transformation, institutional change, and policy response. Therefore, the rapid and extensive application of generative AI will inevitably raise more ethical issues, and effective regulation will create a favorable environment for its development. While generative AI is expected to enhance healthcare standards, it is crucial to focus on the responsible use and regulation of AI technology to prevent misuse and unethical behavior and ensure the full protection of human dignity.

1. INTRODUCTION

On November 30, 2022, OpenAI released the ChatGPT chatbot model, which significantly improved the human-machine interaction experience by employing a "Transformer architecture large model + human feedback reinforcement learning algorithm" for natural language processing (Ouyang, Wu, Jiang, et al. 2022). This showcased the immense potential of large-scale artificial intelligence (AI) models in applications, demonstrating a more "human-like" intelligent performance. Subsequently, major global technology companies such as Google, Microsoft, Baidu, and Alibaba successively launched their own general AI language models. The remarkable generative capabilities and significant social value of generative AI have attracted widespread attention across various fields (Mondal, Das, & Vrana, 2023).

DOI: 10.4018/407369

It has disrupted previous perceptions and distances from AI technology, showcasing powerful "creative" abilities, ranging from image and music generation to the development of new machine learning algorithms. Generative AI technology is being rapidly integrated into numerous new applications (Bozkurt, Xiao, Lambert, et al. 2023). Experts predict that with the breakthrough of generative AI large models, the availability and industrial informatization level of AI technology will significantly improve, and the industrial implementation timeline of general AI technology is expected to gradually shorten from 20 years to less than 5 years. Industrial AI is poised for rapid development (Zhang, Zhang, Li, et al. 2023).

The integration of AI in healthcare has a long and evolving history. As early as the 1970s, rule-based expert systems such as MYCIN and INTERNIST-I laid the foundation for computer-assisted diagnosis, although their reliance on hand-coded knowledge limited scalability (Schaffner, 1981). The 2000s witnessed the widespread adoption of electronic health records (EHRs), catalyzing the accumulation of large-scale structured and unstructured clinical data (Burton, Anderson, & Kues, 2004). Parallel advances in computer vision enabled automated interpretation of radiological images, while machine learning techniques began to show promise in risk prediction and outcome modeling. The next major milestone came with the rise of deep learning in the 2010s, driven by innovations in neural network architectures and the availability of GPU acceleration (Prince, 2012). Convolutional neural networks (CNNs) achieved radiologist-level performance in medical image classification tasks, and recurrent neural networks (RNNs) and attention mechanisms were deployed in electronic health data modeling (Shiri, Perumal, Mustapha, et al. 2023). However, these models were often task-specific, lacked flexibility, and struggled with generalization. The emergence of the Transformer architecture marked a paradigm shift. Transformers enabled scalable self-attention mechanisms that proved highly effective in processing sequential data, leading to the development of pre-trained language models such as BERT and GPT. These models demonstrated strong generalization capabilities, paving the way for generative AI. Generative AI, particularly large language models (LLMs), represents a synthesis of these technological evolutions. It leverages massive pre-training on diverse data sources and fine-tuning with human feedback to produce coherent, context-aware responses. In the healthcare context, this translates into unprecedented capabilities: from summarizing complex clinical notes and generating patient-specific reports to assisting in drug discovery and treatment planning.

The vertical application of generative AI requires industries to possess vast amounts of high-quality data in order to fully leverage its technological capabilities, and the healthcare sector precisely meets this requirement. Medical data represents one of the largest categories of data in human society (Albahri, Duhaim, Fadhel, et al. 2023). Generative AI is bound to bring tremendous opportunities to the healthcare field, leading healthcare and wellness into the next creative world (Eysenbach, 2023). However, the birth of new technologies often presents an unpredictable double-edged sword that can bring convenience and opportunities to the healthcare field while also giving rise to new problems, such as data privacy, ethical issues, and legal regulations (Cascella, Montomoli, Bellini, et al. 2023). This article provides an overview of the advantages and challenges of generative AI applications in the healthcare field, aiming to provide deeper understanding and guidance for the medical community and researchers.