ISSN 2959-6149

Potential Applications and Safety of Large Language Models in Healthcare

Siyin Chen

Abstract:

This study explores the potential applications and associated safety concerns of large language models in healthcare. It particularly highlights the multifaceted applications of large language models (e.g., ChatGPT, MedLM) in healthcare, including data analysis, diagnostics, information retrieval, usage of medical devices, and assistance in tasks. Concurrently, it underscores the risks present alongside these applications, especially concerning data privacy. The study emphasizes the necessity of data privacy protection throughout the entire cycle. By reviewing policies from various countries, it outlines the critical role of refined policies and a clear governmental stance in advancing the application of large language models in healthcare and ensuring their safety.

Keywords: large language model, healthcare, privacy security, healthcare big data

1. Introduction

The emergence of large language models has introduced new opportunities and challenges for the digital transformation of the healthcare sector. As an outstanding representative of deep learning, models such as GPT-4.0 have garnered widespread attention for their successful application in natural language processing. However, their potential extends far beyond, showing valuable applications in clinical practice, diagnosis, drug dispensing, information processing, remote patient monitoring, data sharing, and test analysis. This study aims to delve into the various applications of large language models in healthcare and address potential pitfalls concerning data privacy, national policies, and societal attitudes.

2. Introduction to Large Language Models

Large language models represent a branch of natural language processing technology grounded in deep learning, boasting exceptional capabilities in language comprehension and generation. The historical development of these models can be traced back to the early phases of artificial intelligence, initially focusing on simulating human conversation. With the progression of AI and deep learning, large language models have gradually evolved to process increasingly complex textual information and tasks. Their architectural framework and pre-training mechanisms enable them to understand and generate complex language structures, providing a powerful tool for the medical field. Furthermore, tests such as the United States Medical Licensing Examination (USMLE) have demonstrated that models like ChatGPT can, through training, proficiently acquire the skills and knowledge required for clinical practice. [8,10]

Large language models exhibit core features and principles: 1) Model Structure: Typically utilizing the transformer architecture, these models feature a self-attention mechanism that enhances their ability to capture the contextual nuances of language during input processing. 2) Pre-training Mechanism: They are generally pre-trained on extensive datasets to learn the universal patterns of language, focusing on a deep understanding of language without being optimized for specific tasks initially. 3) Parameter Volume: The term “large” in large language models refers to their vast number of parameters, with models like GPT-3.5 housing billions of parameters, enabling the representation of complex and abstract language structures. 4) Fine-tuning: Post-pre-training, models can be fine-tuned to adapt to specific domains or tasks, allowing optimization with limited labelled data for improved performance in particular areas.

In December 2023, Google introduced a medical large language model named MedLM to its cloud users, building upon the foundations of Med-PaLM and Med-PaLM 2 text models. Med-PaLM, previewed at the end of 2022 and published in “Nature” in July 2023, became the first AI system to surpass the passing threshold for USMLE-style questions, generating accurate, extended responses to health queries. March 2023 saw the launch of Med-PaLM 2, the first to provide answers at a human expert level for USMLE-type questions. To assess the accuracy of model responses, Google established the MultiMedQA benchmark, incorporating seven Q&A datasets covering professional medical examinations, medical research, and consumer inquiries. Med-PaLM emerged as the first model to achieve a passing score on the MedQA dataset for USMLE questions, with an accuracy of 67.7%, while Med-PaLM 2 advanced to 86.5%. Beyond USMLE-type questions, the models' extended responses were evaluated by clinicians and non-clinicians from various backgrounds and countries, with criteria including scientific accuracy, precision, medical consensus, reasoning, bias, and potential for harm. In three consumer healthcare question datasets, both Med-PaLM and Med-PaLM 2 performed well, and in a paired study, Med-PaLM 2's answers were often superior to those of human doctors in most standards. [3,4]

3. Potential Applications in different directions

3.1 Data Analysis

In the realm of data analysis, large language models offer formidable support for disease surveillance, treatment, and patient care. By deeply analyzing clinical data, these models can detect trends in patient conditions, providing early warnings and even assisting physicians in devising personalized treatment plans. Remote patient monitoring is also enhanced, with models analyzing patient data in real-time to identify anomalies, thereby offering timely decision support to healthcare professionals. Google has developed a multimodal model based on PaLM-E for Med-PaLM, enabling synchronous communication of imaging information with physicians, such as chest and breast X-rays, to improve patient care. Currently, it supports dermatology, retinology, radiology (3D and 2D), pathology, health records, and genomics, although it has not yet been clinically deployed [4]. Additionally, The Royal Free NHS Foundation Trust and Google DeepMind's AI division have created a real-time alert system for Acute Kidney Injury (AKI) by monitoring patient blood tests and electronic medical records, sending immediate AKI alerts to clinicians upon detecting deterioration signs. However, this achievement has faced controversies over information security.[17]

3.2 Assistant

Large language models can serve as assistants in electronic medical record entry and preservation, enabling physicians to manage and record patient information more efficiently. Moreover, these models can act as remote virtual medical assistants and medical translation tools, enhancing efficiency and reducing the likelihood of human error. Beyond potential clinical contributions, large language models can also enhance the efficiency and accessibility of medical education. Given the high knowledge demands on medical professionals and the lengthy time required to train medical talent, large language models can assist in medical professional education through speedy and accurate information extraction. As the technology matures, medical personnel may rely on large language models for co-decision-making in medical solutions, potentially alleviating the talent shortage caused by the lengthy and challenging process of medical education and licensure.

3.3 Intelligent Q&A

Large language models hold extensive potential in intelligent Q&A, assisting in clinical decision-making through deep learning capabilities to analyze medical records and literature, providing instant clinical recommendations. These models can offer information on possible diagnoses based on symptoms, medical history, and lab results, aiding in more accurate diagnoses, and reducing errors due to insufficient medical knowledge or misinterpretation. Furthermore, models like the open-source Doctor Dignity on GitHub, based on Meta's Llama with 27 billion parameters and fine-tuned on medical dialogue datasets, have passed the US medical licensing exams. They aim to create personal doctors for patients, accessible on any local device without needing an API (Application Programming Interface) installation [2]. Another example, DoctorGPT, supported by Algomed and OpenAI, which requires API installation and offers personalized advice based on symptoms and history, facilitating efficient and convenient medical consultations for patients. [1]

3.4 Information Retrieval

Large language models also show potential in information retrieval, such as medical knowledge, drug administration, and medication information queries. They can expedite the extraction of information from medical literature and reviews for researchers and doctors, supporting medical research, clinical practice, and education. By synchronizing with hospital medication inventories and learning about medications, these models can assist in creating personalized treatment plans and reducing medication errors and adverse reactions. For patients, they can explain medication effects, usage, and possible side effects in natural language, combining intelligent diagnostic capabilities to suggest basic medication advice based on past clinical treatment plans. [7]

3.5 Medical Devices

The application of large language models in the medical device sector offers innovative prospects. By analyzing extensive data on medical device designs, patient needs, and physician usage habits, models can provide designers with innovative concepts and optimization suggestions. Predicting the failure probability and maintenance needs of devices aids in planning maintenance schedules, and ensuring the reliability of medical equipment. Coupled with simulation technology, models can conduct mock tests on devices to evaluate their performance and safety, bolstering the development and improvement of medical devices. [13]

The broad potential applications of large language models in healthcare demonstrate their multifaceted advantages, enhancing the quality and efficiency of medical services. However, while maximizing their benefits, it is crucial to address potential risks related to data privacy and ethical issues, ensuring sustainable and safe development of large language models in medical applications.

4. Potential Risks

4.1 Data Privacy

With the broad application of large language models in the medical field, protecting patient data privacy emerges as an urgent issue to be addressed. Medical health big data is primarily divided into four categories [5,13]:

1) Clinical Big Data: This includes basic content generated during a patient's medical visit, such as detailed personal information, electronic medical records, medical imaging data, and medication records, which might contain sensitive information like names, ages, addresses, and phone numbers. Improper handling of these data could lead to privacy breaches, raising legal and ethical concerns.

2) Health Big Data: Comprising data from wearable devices, mobile app monitoring, and internet behavior data, these form an individual's electronic health record for monitoring health status. However, this sensitive health information could potentially lead to privacy leaks, especially when interconnected with the internet and medical institutions.

3) Biological Big Data: Encompassing vast datasets from genomics, transcriptomics, proteomics, metabolomics, etc., genetic testing data combined with pathological data may reveal an individual's identity, leading to genetic discrimination. Such information leaks could doubly harm patients' privacy and mental health.

4) Operational Big Data: Medical institutions' operational data includes cost accounting, procurement of drugs, consumables, equipment, and drug development data, which may involve private information about a person's health and financial status. Misuse or leakage of these data could damage patients' trust in medical institutions and may be exploited for commercial purposes, leading to disputes over privacy rights.

Clinical and operational big data, which need to be accessible as training and application data for large language models, are closely linked to patient information. Thus, data privacy protection is indispensable in the development trajectory of large language models. Privacy protection across the entire lifecycle, from data collection, storage, and sharing to analysis, requires specific measures.

For example, during data collection, traditional medical data privacy protection mainly employs anonymization techniques to obscure the link between data and individuals, but these are easily decrypted. More targeted anonymization and differential privacy techniques, such as the DAIMDL algorithm, are needed to weaken data's sensitive attributes. In the data storage phase, encryption technologies suited to health big data and cloud platforms are necessary, along with third-party audits of data integrity. In the data-sharing phase, medical data-sharing platforms like the National Health Information Network (NHIN) in the U.S. have been successfully established. This phase primarily requires identity verification, including user and data authentication for access control, where authentication anonymization can also reduce exposure risks. Lastly, in the data analysis phase, ensuring the confidentiality of data transmission and computation during machine learning training and preventing the leakage of training data are paramount. [5]

4.2 National Policies

Regulations on healthcare data and Al applications vary from country to country, which creates challenges and imposes certain constraints on the use of large language models in healthcare. For example:

1. The UK has legislation that clarifies the sharing of human information: information shared by individuals with professionals should not be used or shared unless the individual understands and agrees to it. National Health Service legislation covers the confidentiality of certain approved research, but patients can opt out of such use to protect their privacy [15, 16].

2. In Germany, there are more specific regulations for the healthcare sector: doctors and other healthcare professionals are generally prohibited from sharing patient data with third parties without the patient's consent; breach of the duty of professional confidentiality is not only sanctioned by the Federal German Medical Association, but also constitutes a criminal offence; and doctors and healthcare professionals can share health data with "contributors", such as service providers acting for the benefit of and on behalf of the healthcare professional, e.g. IT service providers or providers of practice management software. Doctors and healthcare professionals may share health data with "contributors", e.g. service providers acting for the benefit

of and on behalf of healthcare professionals, such as IT service providers or providers of practice management software. [16]

3. On these bases, France explicitly prohibits patient consent for the transmission of medical data: patients cannot consent to the transmission of their healthcare data to third parties. The main exception is in the case of secondary data processing for research, in which case the organisation concerned must first obtain authorisation from the CNIL (Commission nationale de l'informatique et des libertés). [16]

4. Japan: A special law on medical big data is enacted in addition to the law on the protection of personal information, which defines the rights of the subject, the responsibility of the subject, and the mode of dissemination of medical and health data. [9]

5. In Russia, in addition to the laws and regulations related to information sharing, there is a clear indication of compliance related to digital innovation: the Law on Experimentation in the Field of Digital Innovation specifies that it can be carried out in the field of healthcare [9].

Currently in China, the Personal Information Protection Law defines medical and health information as sensitive personal information to be protected, adopts strict protection measures and sufficient purpose necessity, and strengthens processing restrictions and rules, but there are still more legal barriers and blind zones in the application of medical and health data.2023 On 11 April 2023, the State Office of Internet Information Technology issued the "Measures for the Administration of Generative Artificial Intelligence Services (Draft for Opinion) " clarifying that where personal information is involved, the provider assumes the legal responsibility of a personal information processor and fulfils the obligation to protect personal information. If the training data contains personal information, the consent of the subject of personal information should be obtained. At the same time, the provider assumes the obligation to protect the user's input information and usage records in the course of providing services. It shall not illegally retain input information from which the identity of the user can be inferred, shall not conduct profiling based on the user's input information and usage, and shall not provide the user's input information to others.

Based on the relevant laws and policies of various countries, there is also some inspiration for China, such as: amending the Basic Medical Care and Health Promotion Law to increase the legal norms on the processing and utilisation of medical information; formulating a pilot law on artificial intelligence and digital innovation, creating an approval path for the processing of medical data subjects or projects; clarifying the conditions and rules for the processing of anonymised data, enriching the legal norms on the utilisation of de-identified data etc [14].

At the same time, for the patient-oriented application APP or platform (e.g., online intelligent consultation), government regulators need to formulate laws and regulations for Internet healthcare and medical data APP sharing in accordance with relevant policies and regulations, and establish a perfect regulatory mechanism to strictly require and review the compliance of APP privacy policy [12].

4.3 Social attitudes

The public attitude towards large language modelling in healthcare involves multiple dimensions such as ethics and credibility. Based on the analysis of the three-party evolutionary game, patient participation is a crucial part of healthcare data sharing, and once patients refuse to participate, the other two parties will lose the motivation to promote healthcare data sharing. For healthcare providers, the amount of government penalties and subsidies is an important factor that affects their strategic choices, and a reasonable amount of penalties and subsidies will motivate them to protect the privacy of medical data more actively [6].

5. Summary

Large Language Modelling is a major trend for the future and has great potential in a number of fields. In healthcare, which is of great significance to human society and the human economy, the wide range of potential applications of Large Language Models in healthcare can greatly improve the efficiency and accuracy of both clinical and non-clinical applications in the future. Several companies have already launched and continuously improved their healthcare-specific Large Language Models, which are currently focused on intelligent Q&A diagnostics for patients, such as Google's MedLM and open-source Doctor Dignity, etc. In terms of professional testing, there are already a number of Large Language Models in use. In terms of professionalism testing, some of them have already generated answers that are better than those of human medical experts. However, in addition to the further improvement of the accuracy of the answers of the large language model as well as the development of the medical field for the potential data, ethics and other potential pitfalls have not yet been resolved. It is necessary to improve national laws and regulations, to issue clear regulations on privacy protection and data use for medical data used for large language model training, and to ensure how to anonymise and privatise the data, at the same time, if the use of large language models is implemented, the attitude of the government departments has a certain leading role in the attitude of the society, which needs to be actively promoted by the public sector.

Reference

1. DoctorGPT. (n.d.) DoctorGPT. https://doctorgpt.co.in/

2. GitHub. (2021) Doctor Dignity. https://github.com/llSourcell/Doctor-Dignity

3. Google Cloud. (2023) Use MedLM Models. https://cloud.google.com/vertex-ai/docs/generative-ai/medlm/overview?hl=en

4. Google Research. (2023) Med-PaLM: A Large Language Model from Google Research, designed for the medical domain. https://sites.research.google/med-palm/?hl=zh-cn

5. Guo, Z., Luo, Y., Cai, Z., et al. (2021) Overview of Privacy Protection Technology of Big Data in Healthcare. Journal of Frontiers of Computer Science and Technology, 15: 389-402.

6. Han, P., Gu, L., Zhang, J. (2021) Research on Willingness to Share Medical Data from Perspective of Privacy Protection——Based on Tripartite Evolutionary Game Analysis. Journal of Modern Information, 41:148-158.

7. HU, J. (2018) Development Framework and Trend Analysis of Medical Health AI. Chinese Journal of Health Informatics and Management, 15: 485-491.

8. Li,J., Dada, A., Kleesiek J., et al. (2023) ChatGPT in Healthcare: A Taxonomy and Systematic Review. medRxiv.

9. Li. X., Ning, S., Akhmetshin. (2023) Legal techniques of the utilization of medical big Data in Russia and enlightenment. China Health Law, 31: 54-59.

10. McKinsey&Company. (2017) Using Artificial Intelligence to prevent healthcare errors from occurring. In: Second Global Ministerial Summit On Patient Safety.

11. Xiao, A., Lu, Y., Wu, J., et al. (2021) Privacy Security Mechanism of Biomedical Big Data. Computer Applications and Software, 38: 318-322.

12. Zhao, Y., Yan, Z., Shen, Q., et al. (2022) Evaluating Privacy Policy for Mobile Health APPs with Machine Learning. Data Analysis and Knowledge Discovery, 6: 112-126.

13. Kong, X. (2023) Innovation opportunities and challenges of ChatGPT in the medical industry. Zhangjiang Technology Review, 2: 68-71.

14. Xiong, M., Chi, X. (2023) On the security of generative large language model applications——ChatGPT as example. Shandong Social Sciences, 5: 79-90.

15. GOV.UK (2023)Approval standard: confidential patient information. https://www.gov.uk/government/publications/accessing-ukhsa-protected-data/approval-standards-and-guidelines-confidential-patient-information

16. Lamb, S., Tschammler, D., Gottlieb, F., et al. (2023) Health Data in The EU And UK – Regulatory Trends and Developments. https://www.mwe.com/insights/health-data-in-the-eu-and-uk-regulatory-trends-and-developments/

17. Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and technology*, 7(4), 351–367. https://doi.org/10.1007/s12553-017-0179-1