Large Language Models in Medical Education and Quality Concerns

Pandey M and Mishra V*

Chief People Officer, Digital Musketeers, India

College of Healthcare Management and Economics, Gulf Medical University, Ajman, UAE

*Corresponding author: Vinaytosh Mishra, Associate Professor, College of Healthcare Management and Economics, Gulf Medical University, Ajman, UAE, Email: vinaytosh@gmail.com

Keywords: Medical Education; Chat GPT; HTA

Abbreviations: HTA: Health Technology Assessment; NLP: Natural Language Processing; HIPPA: Health Insurance Portability and Accountability Act.

Background

The world is witnessing increased digitalization in the recent past [1]. Artificial Intelligence has become helpful in solving many perennial problems in healthcare and large language models are one of those applications [2]. Large language models can be utilized in medical education to improve the learning and training of medical students [3]. These models comprehend complicated medical ideas and moral quandaries using a massive quantity of medical data and natural language processing, through which students receive more in-depth and interesting learning resources [4]. The advent of Generative Pre-trained Transformer applications such as Chat GPT and integration of these with popular search engines such as Bing will make the adoption of these technologies in healthcare fast. Medical information search will go with sea changes in the coming two years. But some serious ethical and quality issues need to be addressed before promoting the use of these technologies in medical education. Thus, there is a need for Health Technology Assessment (HTA) before introducing LLM or GPT into the formal education system [5]. The functioning of a language model is depicted in Figure 1.

Figure 1: Functioning of Large Language Models.
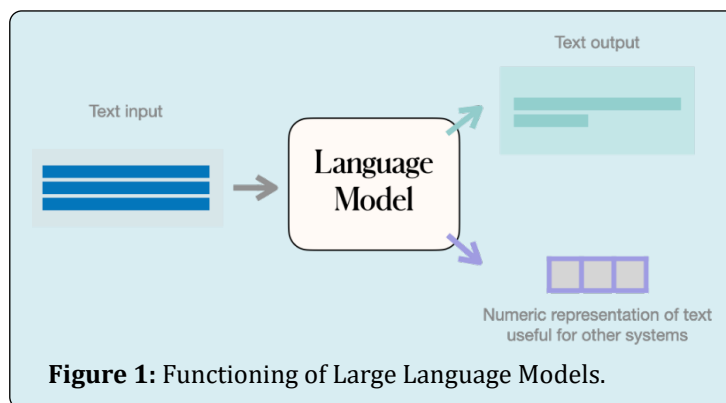
Use of LLM in Medical Education

Large language models are used in medical education in the following ways:

- **Case simulations**: These models can help medical students better comprehend the complexities of patient care by helping them assess and comprehend difficult medical cases and ethical situations.

- **Teaching & Learning:** In the context of medical education and large language models refers to the application of cutting-edge machine learning and artificial intelligence approaches to improve teaching and learning by creating natural language processing (NLP) based image creation. But there is a lot that needs to be done in this area.
- **Medical terminology**: By helping medical students learn it, large language models can make it simpler for them to comprehend complicated ideas and communicate clearly with patients and other healthcare professionals.

Medical students can use large language models to learn about medical ethics, including patient autonomy, informed consent, and confidentiality. By doing this, they can acquire the skills and information necessary to deliver the best possible patient care. It can help them prepare for medical laws such as Health Insurance Portability and Accountability Act (HIPPA). The use of extensive language models in medical education has the potential to completely transform how medical students are educated by giving them access to a wealth of information.

Large language models can assist students to make clinical judgments that are based on ethical considerations and industry best practices. In recent years the utility of large language models has increased in the field of medical education. The probability of a given word or phrase being used in a medical context can be predicted by these models. The models also intend to identify words that have a likelihood to be included in medical dictionaries but are currently excluded. It has been demonstrated that large language models are efficient for several activities, including recognizing uncommon and potentially harmful pharmacological side effects, increasing the accuracy of diagnosis codes, and offering insights into the efficacy of therapies. Additionally, summaries of clinical studies and the identification of patients at risk for particular diseases have been produced using massive language models. The effectiveness and precision of medical education materials may be increased by the use of large language models. Large language models can also be utilized to provide tailored content for audiences, such as patients with certain medical issues or students with learning needs. As already mentioned earlier, large language models, like Chat GPT, can be utilized in medical education as a learning aid for students. They can be used, for instance, to respond to inquiries, offer clarifications, and spark conversation on medical ethics. Another concern with the medical education database was not including information about social determinants of healthcare and LLM-based solutions have the potential to address this limitation. Figure 2 depicts the potential of LLM as a holistic information source for healthcare (Figure 2).
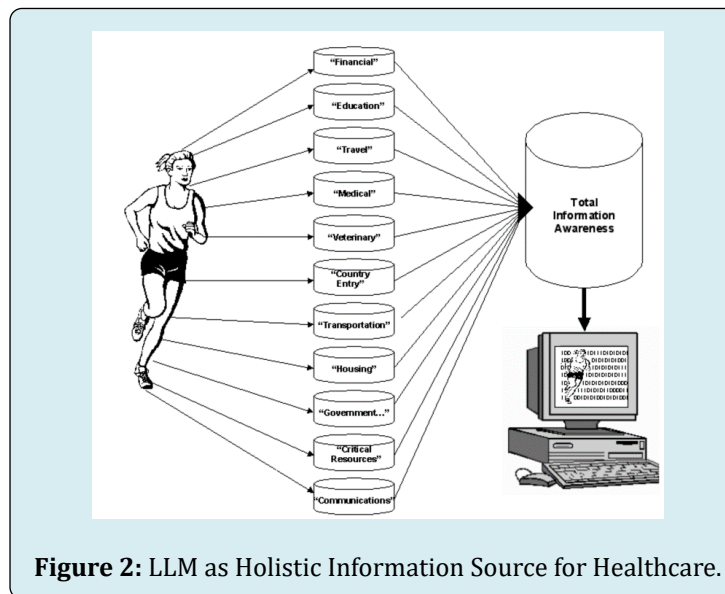

Figure 2: LLM as Holistic Information Source for Healthcare.

Quality and Ethical Concerns

It's crucial to keep in mind that these models cannot replace human expertise and that any predictions made using them should always be subjected to a critical review and cross-checked with data from other trustworthy sources. When using AI in medical education, additional ethical factors including protecting privacy and preventing biases in training data should be considered.

To enhance and streamline healthcare, the U.S. government passed a law 10 years ago that actively pushed the adoption of electronic health records. Beyond the purview of clinical trials, the vast amount of data in

these now-digital records could be utilized to address the following very specific questions: What dosage of this drug is appropriate for patients of this height and weight? What about individuals who have a particular genetic profile?.

Unfortunately, the majority of the information that could provide the answers to these queries is buried in complicated doctor's notes. Current methods make it difficult for computers to comprehend these notes; to extract information, numerous machine-learning models must be trained. To enhance and streamline healthcare, the U.S. government passed a law 10 years ago that actively pushed the adoption of electronic health records. Beyond the purview of clinical trials, the vast amount of data in these now-digital records could be utilized to address the following very specific questions: What dosage of this drug is appropriate for patients of this height and weight? What about individuals who have a particular genetic profile?

Conclusion

A time-consuming and expensive procedure, training each model requires domain specialists to categorize large amounts of data. Models created for one hospital also don't operate well at others. Thus, the interoperability of these devices needs to be addressed before promoting them on a large scale. A system that learns from a little amount of labeled data works well at many hospitals and uses a single model to extract multiple sorts of information would be great. Yet how? Researchers at MIT employed a very large, GPT-3-style model to perform tasks like expanding overloaded jargon and acronyms and extracting prescription regimens to obtain that crucial medical information. Another notable application is the Green Button project of Stanford University.

The use of LLM and GPT in medical education is many but there is grave privacy, security, interoperability, and above all quality concerns. These bottlenecks need to be addressed to realize the full potential of Artificial Intelligence in medical education.

References

1. Mishra V, Sharma MG (2022) Digital transformation evaluation of telehealth using convergence, maturity, and adoption. Health Policy and Technology 11(4): 100684.

2. Mishra V, Samuel C, Sharma SK (2015) Use of machine learning to predict the onset of diabetes. Int J Recent Adv Mech Eng 4(2).

3. Yang X, Chen A, Pour Nejatian N (2022) A large language model for electronic health records. NPJ Digital Medicine 5(1): 194.

4. Gu Y, Tinn R, Cheng H (2021) Domain-specific language model pretraining for biomedical natural language processing. ACM Transactions on Computing for Healthcare 3(1): 1-23.

5. Mishra V, Singh J (2022) Health Technology Assessment of Telemedicine Interventions in Diabetes Management: Evidence from UAE. FIIB Business Review 23197145221130651.

6. Raphael D (2006) Social determinants of health: present status, unanswered questions, and future directions. International Journal of Health Services 36(4): 651-677.

7. Longhurst CA, Harrington RA, Shah NH (2014) A 'green button' for using aggregate patient data at the point of care. Health affairs 33(7): 1229-1235.

Pandey M and Mishra V. Large Language Models in Medical Education and Quality Concerns. J Qual Healthcare Eco 2023, 6(1): 000319.

Copyright© Pandey M and Mishra V.