Future Medicine AI

A clinician's guide to large language models

Giovanni Briganti*,1,2 (iD)

[1] Chair of AI & Digital Medicine, Faculté de Médecine, Université de Mons, Avenue du Champs de Mars 6, B7000 Mons, Belgium
[2] Département des Sciences Cliniques, Université de Liège, Quartier Hôpital, B4000 Liège, Belgium
*Author for correspondence: giovanni.briganti@hotmail.com

The rapid advancement of artificial intelligence (AI) has led to the emergence of large language models (LLMs) as powerful tools for various applications, including healthcare. These large-scale machine learning models, such as GPT and LLaMA have demonstrated potential for improving patient outcomes and transforming medical practice. However, healthcare professionals without a background in data science may find it challenging to understand and utilize these models effectively. This paper aims to provide an accessible introduction to LLMs for healthcare professionals, discussing their core concepts, relevant applications in healthcare, ethical considerations, challenges, and future directions. With an overview of LLMs, we foster a more collaborative future between healthcare professionals and data scientists, ultimately driving better patient care and medical advancements.

Tweetable abstract: AI's foundation models (e.g., GPT, LLaMA) are transforming healthcare. This paper offers an accessible intro for healthcare pros, covering core concepts, applications, ethics, challenges & future. #AIHealthcare

In recent years, artificial intelligence (AI) has made remarkable advancements in various fields, including healthcare [1]. Among these developments, large language models (LLMs) have emerged as a powerful tool for generating, understanding, and manipulating human-like text. These models hold great promise for revolutionizing medical practice and improving patient outcomes [2]. However, healthcare professionals without a background in data science may find it challenging to understand and utilize these models effectively.

To provide a clearer picture of how AI can aid clinicians, it is beneficial to look at its core capabilities in healthcare: natural language processing (NLP), image processing, and prediction. NLP involves computer algorithms understanding and interacting with human language. For example, in healthcare, NLP can help extract critical information from unstructured clinical notes, improving patient stratification for treatment plans [3]. Image processing, another crucial area of AI, is invaluable in specialties like radiology and pathology. AI algorithms can assist in interpreting imaging data such as x-rays or MRIs, identifying abnormalities that might be challenging for the human eye to detect. For instance, AI can help diagnose conditions such as lung cancer from chest radiographs or detect signs of diabetic retinopathy from retinal images [4,5]. Last, AI shines in prediction tasks, using historical and real-time data to make forecasts. An example of this is predicting disease progression based on a patient's electronic health records, which can assist in timely and personalized intervention [6,7].

The purpose of this paper is to provide healthcare professionals with an accessible introduction to LLMs. The paper aims to explain core concepts, highlight relevant applications in healthcare, and address ethical considerations and challenges. The scope of the paper is to provide a solid foundation for healthcare professionals to explore the potential of these models in their own practice and collaborate with data scientists and AI experts.

The growing adoption of AI in healthcare underscores the need for medical practitioners to be aware of and understand the implications of LLMs. As these models continue to evolve and find their way into various aspects of medical practice, it is crucial for healthcare professionals to be prepared for their potential impact on diagnostics, treatment planning, research, and patient care. By providing an accessible introduction to LLMs, this paper aims to facilitate a more informed and collaborative future for healthcare professionals and data scientists alike.

Future Medicine

LLMs: an overview

LLMs are large-scale machine learning models that serve as a base for various AI applications. They are pre-trained on extensive amounts of data, often from diverse sources, to develop a comprehensive understanding of human language, relationships, and patterns. These models can then be fine-tuned for specific tasks, enabling them to generate high-quality outputs in a wide range of domains, including healthcare [8].

The core concepts of LLMs include:

- Pre-training: this phase involves training the model on large-scale datasets, which allows the model to learn language patterns, grammar, and contextual relationships;
- Fine-tuning: after pre-training, the model is fine-tuned on task-specific data to adapt its knowledge to the specific domain, such as medical diagnostics or treatment planning;
- Transfer learning: this concept refers to the ability of LLMs to leverage knowledge acquired during pre-training and apply it to new, related tasks, thus reducing the need for extensive training data for each task.

Such concepts may be hard to grasp at first for healthcare professionals, so let us make an example for healthcare data.

LLMs are first pre-trained on a large-scale medical dataset, including diverse sources like electronic health records, medical literature, and clinical trial reports. After pre-training, the models are fine-tuned on healthcare-specific applications, such as medical diagnosis or treatment recommendation, by learning from task-specific data. The fine-tuning process involves training the model on a dataset of anonymized patient-reported symptoms and corresponding diagnoses, allowing it to adapt its general medical knowledge to a specific task. The resulting model can then be used by healthcare professionals to generate a ranked list of potential diagnoses, helping to expedite the diagnostic process and improve patient care. Transfer learning allows us to leverage the knowledge acquired during pre-training and fine-tuning to adapt the model to a related but distinct task. For example, we can further fine-tune the model on a new dataset containing effective treatment options. This process enables the model to associate specific diagnoses with their corresponding treatments, building upon its existing knowledge of patient-reported symptoms and medical conditions.

The resulting treatment recommendation model can then be used by healthcare professionals to suggest personalized treatment options for their patients based on their reported symptoms and confirmed diagnosis. The AI-generated list of potentially effective treatments can support more informed treatment decisions and tailored patient care.

Let us consider the development of an LLM-based model to aid the general practitioner in dealing with diabetes patients. As a step-by-step reasoning, we would have:

- Pre-training ('learning medical language'): the model begins by learning from a wide array of medical data - textbooks, research papers, and anonymized patient reports. It gains a fundamental understanding of medical language but lacks a specialized focus;
- Fine-tuning ('specializing in chronic disease management'): the model then receives additional training on information specifically related to chronic diseases like diabetes. This includes data on patient symptoms, blood sugar levels, and the resulting diagnoses. As a result, the AI tool becomes proficient in understanding the context of chronic disease management and can help physicians interpret patient-reported data and lab results, identifying when a patient's diabetes might be under poor control;
- Transfer learning ('tailoring treatment recommendations'): finally, the model is provided with data associating specific patient states (such as elevated blood sugar levels despite current medication) with corresponding adjustments in treatment (like changing the type or dose of insulin). This way, the AI tool can suggest possible treatment modifications to the physician when a patient's data suggests their current regimen isn't working well.

In this example, the pre-training process involves learning general medical language understanding from a large-scale dataset. Fine-tuning focuses on adapting this knowledge to a specific task, in this case, diabetes. Transfer learning enables the model to build upon its existing knowledge and adapt to a related but distinct task, such as predicting the most effective treatment options for a patient based on their reported symptoms and diagnosed medical condition.

It is crucial to emphasize that AI-generated suggestions should be treated as a supplementary tool, with healthcare professionals using their clinical expertise and judgment to make the final diagnosis and treatment decisions.

Applications in healthcare
Medical diagnostics
LLMs have shown promise in assisting healthcare professionals with diagnosing medical conditions [9]. By analyzing patient data, such as medical history, symptoms and test results, these models can generate differential diagnoses and suggest further tests or interventions [10–12]. This helps reduce diagnostic errors, expedite the process, and enhance the overall quality of care [13].

Treatment planning & optimization
AI-driven models can support the development of personalized treatment plans by considering patient-specific factors and analyzing large volumes of medical literature [14]. They can assist in optimizing treatment strategies, predicting treatment response, and identifying potential adverse effects [13]. This enables healthcare professionals to make more informed decisions and tailor treatments to individual patients, ultimately improving outcomes.

Moor *et al.* [15] have proposed the concept of generalist medical AI (GMAI), that encompasses various applications of AI in healthcare, aiming to enhance patient care and improve clinical workflows. Some potential applications include: bedside decision support, where GMAI models assist clinicians in making informed decisions by providing data summaries, explanations and treatment recommendations; interactive note-taking, where GMAI models help reduce administrative tasks by drafting documents for clinicians to review and approve; chatbots for patients, enabling personalized support and advice outside clinical settings; and text-to-protein generation, where GMAI models design protein sequences and structures based on textual prompts, potentially accelerating drug discovery and development.

Medical research & drug discovery
LLMs can be employed to facilitate medical research by automating the extraction of relevant information from scientific literature, identifying patterns in data, and generating hypotheses [16,17]. They can also accelerate drug discovery by predicting molecular properties, identifying potential drug candidates, and suggesting new chemical structures [18]. This expedites the research process and helps bring new therapies to market more efficiently.

Challenges & limitations
Quality of training data
The performance of LLMs depends on the quality and representativeness of the data they are trained on. In healthcare, high-quality and diverse data is crucial to ensure accurate and generalizable results. However, obtaining and curating such data can be challenging, given issues related to data privacy, consent and varying data collection practices [19].

Interpretability & explainability
LLMs can act as 'black boxes', making it difficult to understand how they arrive at specific conclusions or recommendations. In healthcare, where decision-making has significant consequences, it is essential for professionals to understand the rationale behind AI-generated outputs. Developing more interpretable and explainable models remains an ongoing challenge in the field [20].

Generalizability & real-world performance
While LLMs may demonstrate impressive performance in controlled research settings, their real-world performance can vary. Factors such as data quality, model limitations and differences in patient populations may impact the generalizability of AI-driven insights. Healthcare professionals should be cautious when applying these models in practice and consider validating their performance on local or domain-specific data.

Legal & regulatory concerns
The AI Act proposed for Europe [21], the more comprehensive and specific regulatory attempt at AI, has been met with The AI Act's has been met with skepticism [22], as its focus on healthcare is mostly limited to references to the medical devices regulation (MDR) and *in vitro* diagnostic medical devices regulation (IVDR). The AI Act classifies medical AI under the MDR or IVDR as high-risk if it requires a third-party conformity assessment. Other high-risk AI uses in healthcare include emergency first response services dispatching and risk assessment in life and health insurance. Medical AI could be classified as unacceptable risk if they involve vulnerable groups. However,

if medical AI does not meet the high-risk or unacceptable risk criteria, it falls into the minimal risk category. The MDR currently remains the most relevant approach to regulating AI-based medical devices under the AI Act [22]. Some countries have (temporarily) halted certain LLMs [23].

As AI becomes more integrated into healthcare, legal and regulatory issues will likely arise. These may include questions around liability for AI-driven recommendations, data protection requirements, and the need for standardized AI evaluation protocols. Healthcare professionals should stay informed about the evolving legal landscape and engage in discussions around the responsible use of AI in their practice.

Conclusion

As research in AI and LLMs continues to progress, we can expect new models with improved performance, interpretability, and applicability in healthcare. These advancements may lead to even more accurate diagnoses, personalized treatments and a deeper understanding of complex medical conditions.

The successful integration of LLMs into healthcare requires close collaboration between healthcare professionals and data scientists [1]. By working together, they can develop models that address the unique challenges and requirements of the medical domain, ultimately leading to better patient outcomes.

LLMs hold great promise for telemedicine and global health applications. They can help bridge the gap in access to healthcare resources by providing remote diagnostic support, treatment planning, and health monitoring services. By leveraging these models, healthcare professionals can expand their reach and provide care to underserved populations, improving health equity worldwide.

In conclusion, the future of LLMs in healthcare is promising, with many opportunities for improving patient care and advancing medical knowledge. As healthcare professionals become more familiar with these models and their potential, they can harness their power to transform healthcare and deliver better outcomes for patients around the globe.

Future perspective

Four future perspectives for the use of LLM in medicine are identified.

Precision medicine and personalized treatment: by identifying nuanced patterns in patient data, these models could assist in developing personalized treatment plans that take into account individual differences in genetics, lifestyle and environment.

Predictive analytics: whether predicting disease progression, patient responses to treatments, or forecasting disease outbreaks, these models' predictive abilities could have far-reaching implications for healthcare delivery and public health.

Real-time decision support: LLMs could potentially be integrated into electronic health records (EHRs) and other clinical decision support systems, providing real-time guidance to healthcare professionals. They could help analyze patient data on the spot, suggest differential diagnoses, recommend suitable treatment options, or even identify potential risk factors, all contributing to improved patient care.

Health education and patient engagement: LLMs could be utilized to develop personalized health education materials for patients, improving health literacy and patient engagement. These models could generate easy-to-understand information tailored to each patient's health status and educational needs, empowering patients to take a more active role in managing their health.

Executive summary

- This paper offers an accessible introduction to large language models (LLMs) for healthcare professionals, particularly for those without a background in data science or artificial intelligence (AI). As AI continues to permeate the healthcare sector, there is a need for clinicians to understand the principles of LLMs and their potential impact on medical practice.
- The paper first introduces the basic concepts and mechanics of LLMs, employing simple language and clear illustrations. The focus is on explaining how these models learn and generate human-like text, with real-world examples from the medical field, such as predicting disease progression and recommending treatments based on patient data.
- In addition to explaining the 'how' of LLMs, the paper explores 'why' these models are important in healthcare. It outlines several relevant applications of LLMs, such as streamlining medical diagnostics, enhancing treatment planning, and facilitating patient care. These use cases illustrate how LLMs could revolutionize various aspects of healthcare, leading to improved patient outcomes.
- The paper concludes by emphasizing the need for cross-disciplinary collaboration between healthcare professionals and data scientists. The goal is to empower clinicians with a foundational understanding of LLMs and inspire them to collaborate effectively with AI experts, fostering a future where AI is skillfully integrated into medical practice for the benefit of patients.
- In this paper, the aim is to demystify the concept of LLMs in healthcare and facilitate informed AI adoption in clinical practice.

References

1. Briganti G, Le Moine O. Artificial intelligence in medicine: today and tomorrow. *Front. Med.* (2020). www.frontiersin.org/articles/10.3389/fmed.2020.00027/full

2. Touvron H, Lavril T, Izacard G, Martinet X, Lachaux MA, Lacroix T *et al.* Llama: open and efficient foundation language models. ArXiv Prepr ArXiv230213971 (2023).

3. Iroju OG, Olaleke JO. International Journal of Information Technology and Computer Science (IJITCS). *Int. J. Inf. Technol. Comput. Sci.* 7(8), 44–50 (2015).

4. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJ. Artificial intelligence in radiology. *Nat. Rev. Cancer* 18(8), 500–510 (2018).

5. Grzybowski A, Brona P, Lim G, Ruamviboonsuk P, Tan GS, Abramoff M *et al.* Artificial intelligence for diabetic retinopathy screening: a review. *Eye* 34(3), 451–460 (2020).

6. Zhang Z, Hong Y. Development of a novel score for the prediction of hospital mortality in patients with severe sepsis: the use of electronic healthcare records with LASSO regression. *Oncotarget* 8(30), 49637–49645 (2017).

7. Cooper JA, Ryan R, Parsons N, Stinton C, Marshall T, Taylor-Phillips S. The use of electronic healthcare records for colorectal cancer screening referral decisions and risk prediction model development. *BMC Gastroenterol.* 20, 1–16 (2020).

8. OpenAI. GPT-4 Technical Report. (2023). http://arxiv.org/abs/2303.08774

9. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. *J. Med. Syst.* 47(1), 33 (2023).

10. Chen S, Wu M, Zhu KQ, Lan K, Zhang Z, Cui L. LLM-empowered chatbots for psychiatrist and patient simulation: application and evaluation. ArXiv Prepr ArXiv230513614 (2023).

11. Huang H, Zheng O, Wang D, Yin J, Wang Z, Ding S *et al.* ChatGPT for Shaping the future of dentistry: the potential of multi-modal large language model. ArXiv Prepr ArXiv230403086 (2023).

12. Kleesiek J, Wu Y, Stiglic G, Egger J, Bian J. An opinion on ChatGPT in health care – written by humans only. Vol. 64, Journal of Nuclear Medicine. *Soc. Nuclear Med.* 701–703 (2023).

13. Chirino A, Wiemken T, Furmanek S, Mattingly W, Chandler T, Cabral G *et al.* High consistency between recommendations by a pulmonary specialist and ChatGPT for the management of a patient with non-resolving pneumonia. *Nort. Healthc. Med. J.* (2023).

14. Briganti G. Intelligence artificielle: une introduction pour les cliniciens. *Rev. Mal. Respir.* (2023). www.sciencedirect.com/science/article/pii/S0761842523000906

15. Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ *et al.* Foundation models for generalist medical artificial intelligence. *Nature* 616(7956), 259–265 (2023).

16. Park YJ, Kaplan D, Ren Z, Hsu CW, Li C, Xu H *et al.* Can ChatGPT be used to generate scientific hypotheses?. ArXiv Prepr ArXiv230412208 (2023).

17. Rahman MM, Terano HJ, Rahman MN, Salamzadeh A, Rahaman MS. ChatGPT and academic research: a review and recommendations based on practical examples. *J. Educ. Manag. Dev. Stud.* 3(1), 1–12 (2023).

18. Méndez-Lucio O, Nicolaou C, Earnshaw B. MolE: a molecular foundation model for drug discovery. (arXiv) (2022). http://arxiv.org/abs/2211.02657

19. Mashoufi M, Ayatollahi H, Khorasani-Zavareh D, Boni TTA. Data quality in health care: main concepts and assessment methodologies. *Methods Inf. Med.*, (2023).

20. Dwivedi R, Dave D, Naik H, Singhal S, Omer R, Patel P *et al.* Explainable AI (XAI): core ideas, techniques, and solutions. *ACM Comput Surv.* 55(9), 1–33 (2023).

21. European Parliament. Proposal for a regulation of the European Parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. (2021). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

22. Palmieri S, Goffin T. A blanket that leaves the feet cold: exploring the AI act safety framework for medical AI. *Eur. J. Health Law* 1(aop), 1–22 (2023).

23. Intelligenza artificiale: il Garante blocca ChatGPT. Raccolta illecita di dati personali. Assenza di sistemi per la verifica dell'età dei minori (2023). www.garanteprivacy.it:443/home/docweb/-/docweb-display/docweb/9870847