

• 临床教学 •

生成式人工智能大型语言模型在医学教育
实践的探讨*陈湘¹ 邓然² 吴川清¹

[摘要] 随着训练数据的持续增长以及计算能力的不断提升,大型语言模型(large language models,LLMs)在规模和功能上迅速扩展。生成式人工智能(generative artificial intelligence,GAI)领域的大型语言模型,例如 ChatGPT 等,已成为全球发展最为迅猛的应用之一。在医学教育领域,这些生成式人工智能大型语言模型(GAI-LLMs)展现出了巨大的应用潜力。因此,本文介绍了医疗领域中的 GAI-LLMs,并提出了针对医学教育者的有效策略及所面临的挑战,旨在将 GAI-LLMs 有机地融入学习与教学实践,以显著提升学生的数字技能水平,践行习近平总书记关于数字中国的重要指示精神,推动高等教育的数字化建设,培养新一代数字公民。

[关键词] 生成式人工智能;大型语言模型;医学教育

DOI:10.13201/j.issn.1009-5918.2024.06.007

[中图分类号] G434 **[文献标志码]** A

**Exploration of generative artificial intelligence large language models
in medical education practice**CHEN Xiang¹ DENG Ran² WU Chuanqing¹

(¹Department of Stomach Enterochirurgia, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan,430022, China; ²School of Business, Macau University of Science and Technology)

Corresponding author: WU Chuanqing, E-mail: wucq2014@hust.edu.cn

Abstract With the continuous growth of training data and the ongoing improvement in computational capabilities, large language models(LLMs) have rapidly expanded in scale and functionality. In the field of generative artificial intelligence(GAI),large language models such as ChatGPT are among the fastest-growing applications. In the domain of medical education, these generative artificial intelligence-large language models(GAI-LLMs) demonstrate significant potential for application. Hence, this paper introduces GAI-LLMs in the medical domain, presenting effective strategies and challenges faced by medical educators. The aim is to seamlessly integrate GAI-LLMs into learning and teaching practices, significantly enhancing students' digital skills. This endeavor aligns with the important directives regarding the digital transformation of China by President Xi Jinping, thereby advancing the digitization of higher education and fostering a new generation of digital citizens.

Key words generative artificial intelligence; large language models; medical education

党的二十大明确指出,“推进教育数字化,建设全民终身学习的学习型社会、学习型大国”^[1]。党的二十大报告首次将“推进教育数字化”写入报告,赋予了教育在全面建设社会主义现代化国家中新的使命任务,明确了教育数字化未来发展的行动纲领,具有重大意义。习近平总书记在中共中央政治局第五次集体学习时强调,“教育数字化是我国开

辟教育发展新赛道和塑造教育发展新优势的重要突破口”^[2]。中华人民共和国教育部指出,“教育变革首先是从数字资源建设和教育模式探索开始,进而会引发根本性的范式跃迁,形成教育新形态”。近年来,随着大型语言模型(large language models,LLMs)在规模和功能上的迅速扩展,生成式人工智能大型语言模型(generative artificial intelli-

*基金项目:2023 年度国家自然科学基金(No:82372927);2020 年度国家自然科学基金(No:82073324);2024 年度华中科技大学研究生教改项目(培育)(No:2024PY13)

¹华中科技大学同济医学院附属协和医院胃肠外科(武汉,430022)

²澳门科技大学商学院

通信作者:吴川清,E-mail:wucq2014@hust.edu.cn

gence-large language models, GAI-LLMs)已成为全球发展最为迅猛的应用之一。积极开发、探索、升级 GAI-LLMs 辅助的高等教育教学模式,有助于提升学生的数字技能水平,推动高等教育的数字化建设。

GAI-LLMs 特指采用深度学习技术,利用循环神经网络(recurrent neural network, RNN)或变换器等架构构建的模型^[3]。这些模型经训练可解析与生成人类语言,并涵盖自然语言处理的多项任务,包括机器翻译、文本摘要、对话生成与问题回答。其基本原理在于通过大规模文本数据的训练,从中学得语言结构、语法规则以及语义信息。GAI-LLMs 特点在于其能够生成连贯、语法正确的文本,且能对特定任务产生有意义的响应,最具有代表性的模型为 OpenAI 的 GPT 系列(Generative Pre-trained Transformer)。这些模型经过大规模语料库的预训练,然后通过进一步微调以适应特定任务或应用场景^[4]。在医学教育领域中, GAI-LLMs 展现出了巨大的应用潜力^[5]。因此,本文提出了针对医学教育者的有效策略及所面临的挑战,旨在将 GAI-LLMs 有机地融入学习与教学实践,践行习近平总书记关于数字中国的重要指示精神,以培养新一代数字公民。

1 医疗领域中的 GAI-LLMs

自 2022 年 ChatGPT 发布以来,人们对 GAI-LLMs 的潜在应用表现出了日益浓厚的兴趣,并不断深入探索其广泛的应用前景。大量的案例和试点研究已经明确展示了其在多个领域的卓越能力。针对医疗领域, GAI-LLMs 可能具有改变当前医学实践的潜力,从基础的医疗信息提取、医疗实体标准化,到常用的文本分类、信息检索、问答和对话,再到医疗领域特有的医师考试、病历生成、医疗结果预测、药物研发和医学影像分析等各方面对 GAI-LLMs 的应用进行了探索,并初步取得了较为满意的结果。目前,国内外已有的应用于医疗领域的、具有代表性的生成式大型语言模型有: ChatGPT、Med-PaLM、Galactica、GatorTronGPT、PubMedGPT、PMC-LLaMA、MedGPT、山海大模型、百度灵医、添翼医疗大模型、Deepwise MetAI、ClouD GPT 等^[6]。

2 GAI-LLMs 在医学教育领域的应用策略

GAI-LLMs 能够为医学生和医学教育者提供全新的工具和方法,可能会对医学学习和教育的发展及实践产生深远的影响,其出现也迫使医学生和教育工作者迅速调整他们的学习和教学模式。

2.1 理论教学

医学教育者可借助 GAI-LMMs 生成各种交互式学习活动,例如案例研究,这有助于吸引学生的兴趣并巩固他们的知识。通过模型生成的案例,学

生能够在实践中运用理论知识,加深对医学概念的理解,并促进学生与教育者之间的沟通与讨论。另外,由于 GAI-LMMs 具备广泛的知识库和语言生成能力,其可以根据特定的需求和指导,生成连贯、清晰的文本内容。对于医学教育领域的知识和理解, GAI-LMMs 能基于已有的数据和模式,协助医学教育者生成合适的作业题目、考试题目、参考答案等。GAI-LMMs 也能根据学科要求和标准,帮助教育者构思和组织相关的学习内容,以满足教学和评估的需求。

同时,借助 GAI-LMMs,学生能更高效地处理和消化知识,并在学习过程中迅速解决疑惑。学生可利用 GAI-LMMs 即时反馈、快速准确回答问题的特性,在课堂学习中迅速获取信息,提高对关键问题的理解。此外,学生还可通过 GAI-LMMs 生成练习题、模拟考试,或者获取针对复杂概念的额外解释,这有助于增强他们的考试测试能力,并在复杂医学概念的理解上提供课本以外的额外支持。这种互动式学习方式有助于促进学生的自主学习和深度思考,提升他们的学习效果和应对复杂医学知识的能力^[7]。

2.2 临床教学

GAI-LMMs 可基于医学模拟的学习形式,作为真实患者的替代品应用于临床实践以及临床思维能力培训。这一应用为学生提供了模拟真实情境的机会,使其能够在没有真实患者风险的情况下进行大量练习。GAI-LMMs 模拟环境的特点允许学生探索、尝试和犯错,而无需担心可能产生的负面后果。此外,这种实践让学生在逼真的模拟环境中练习使用他们所学的临床专业知识,从而培养并巩固他们的临床决策技能、临床思维能力、专业态度、道德素养、沟通技巧等在临床实践中至关重要的专业素养。GAI-LMMs 模拟教学为学生提供了一个安全且丰富的环境,为其积极发展在临床实践中所需的技能与能力提供了可靠的平台^[8]。

2.3 教学评估

与 GAI-LMMs 的互动能够生成基于特定医学知识背景的言论,这些言论可以被分析,有助于训练学生“评估与评价”等更高级的认知功能,对所学知识有更深刻的理解,除此之外,其还能应用于教学评估,考察学生的批判性临床思维。这一过程中, GAI-LMMs 通过提示生成特定信息,同时学生被迫改变角色,不是作为“学生”而是作为“编辑”,利用自己的专业知识分析、评估模型生成的文本,或尽可能地进行修正,并得出适当的推论^[9]。这一教学评估方法完全不同于传统的评估方案,为医学教育者们开辟了一条新的思路。

2.4 医学研究

GAI-LMMs 具备广泛的知识覆盖和语言处理

能力,对于撰写研究论文具有不可小觑的作用。其出色的能力可以应用于论文写作的多个方面,包括但不限于:梳理并组织观点、提供专业术语和相关信息、确保论文的逻辑结构和连贯性、辅助特定领域信息的获取与解释等,并且能够为撰写论文的各个阶段提供支持,包括研究、起草、修改和润色。

3 GAI-LLMs 在医学教育实践中面临的挑战

尽管 GAI-LLMs 在医学教育领域有着巨大的应用潜力,但其在医学教育实践中仍然面临着一定的问题与挑战。

3.1 错误与偏见

GAI-LLMs 的输出受到其训练数据的影响,如果训练数据中存在偏见或错误,模型可能会产生不准确或有偏见的信息,且有可能伪造、提供虚假信息,例如:当 ChatGPT 响应提问者对其生成的信息提供出处时,ChatGPT 列出了相应的参考文献,但通过 PubMed 和 Google Scholar 验证,结果显示所有被引用的参考文献都不存在,均为 ChatGPT 编造^[10]。另外,由于 GAI-LLMs 为不透明的“黑盒”人工智能,不仅可能存在错误或偏见,并且有可能完全不被发现。在探究 GAI-LLMs 在 ACEM (Australasian College for Emergency Medicine) 初级考试中的表现时,三种模型 (OpenAI 的 GPT 系列、谷歌的 Bard 和微软的 Bing) 均自信地给出了看似正确的错误答案^[7]。而在医学领域,这些错误的信息可能会对学生的临床决策产生负面影响,因此在使用 GAI-LLMs 时需要对其模型输出的内容进行仔细审查和验证。

3.2 学术不诚信及教育公平性

国外有学者利用 ChatGPT 生成了 3 篇关于“field of competence carcinogenesis (研究领域:致癌)”的学术论文,并利用检测程序 Groover 和 Writer 识别是否为 AI 撰写,两个程序都不能检测到为 AI 编写,且 PlagAware 软件也未检测出抄袭迹象。3 篇文章均原封不动地转发给评估老师,其中 2 篇评估结果为“合格”,1 篇“不合格”,不合格的理由为“使用了某些未命名的术语,以及不匹配的文献引用”^[9]。GAI-LLMs 可代替医学生完成论文、报告等多种形式的课后作业,且可在考试期间迅速回答考试问题,易形成不公平的教学评估结果,使作业设计和在线评估都变得更加困难,因而给医学教育者带来了新的挑战^[11]。目前已有多个大学命令禁止使用 GAI-LLMs 进行写作与作业撰写。另外,有机会使用与没有机会使用这些工具的学生和机构之间也形成了一定的不公平性,具有加剧教育不平等的风险。

3.3 隐私及数据安全问题

使用 GAI-LLMs 进行医学案例模拟时,可能涉及真实患者的信息。即使模型数据是匿名的,也

需要确保患者隐私得到充分保护,避免教学活动侵犯患者隐私权。另外,当转换现有的学习内容或创建新的材料时,教育工作者和研究人员应当谨慎,避免将可识别的数据上传至 GAI-LLMs 系统。未经研究参与者的明确书面知情同意,不应尝试公开这些数据。即使有明确同意,建议教育工作者和研究人员首先将数据发布到数据库中^[12]。

3.4 过度依赖

批判性思维和对知识的整合能力是学生需要培养和磨炼的技能,这是任何人工智能工具都无法取代的。但学生可能过度依赖 GAI-LLMs 生成答案、获取信息,而忽略传统的医学课程学习或实际临床实践,这可能会削弱批判性思维和解决问题的能力,还可能导致学生在实践中缺乏自主性和判断力,影响其在临床实践中的应用能力^[13]。

3.5 其他局限性

GAI-LLMs 的局限性还包括无法完全模拟人类的情感、人际交流能力,以及模型可能对某些医学特定领域了解不足等。此外,模型的更新速度可能滞后于医学领域的发展,如 ChatGPT-3.5 获取的信息仅截止到 2021 年 9 月,因此可能有一些更新的内容并未涵盖,导致信息的过时性^[14]。

4 GAI-LMMs 对未来医学教育的影响及思考

尽管 GAI-LLMs 在医学教育领域未来的应用前景仍不确定,但显而易见的是,它们将持续存在并不断发展,整个医学教育行业也将面临着重大的改变和重塑^[15]。为培养新一代更符合 AI 时代背景下的数字公民,本文就 GAI-LLMs 应用于未来医学教育提出以下几点思考。

4.1 医学生的自主学习以及个体的个性化学习

GAI-LMMs 对医学生的自主学习和个性化学习的影响是一个备受关注的话题。这些模型通过其广泛的知识储备和语言生成能力,为医学教育提供了新的可能性。它们可以个性化地响应学生的需求,根据不同的学习风格和节奏提供信息和答案,从而促进了自主学习的发展。在个性化学习方面,GAI-LMMs 能够根据学生的需求和兴趣定制内容,为每个学生量身定制学习体验。这种个性化定制可以提高学习效率,帮助学生更快地掌握知识,同时也增强了学生对学习过程的投入感和参与度。

然而,尽管 GAI-LMMs 在提供个性化学习方面具有潜力,但也存在一些考量。比如,个性化学习是否会导致信息范围的狭窄化,使学生倾向于只关注其已知领域或兴趣范围内的内容?另外,模型本身的算法偏好或特定数据集的偏差可能会影响到学习内容的呈现,这可能引发信息过滤或偏差的问题。因此,在利用 GAI-LMMs 促进医学生个性化学习的过程中,需要平衡个性化和全面性,确保

学生接触到多样化的信息和视角,同时培养其批判性思维和自主学习的能力。此外,监督和指导也是必要的,以确保学生能够正确地评估和利用 GAI-LMMs 所提供的信息。

4.2 临床思维能力和临床实践能力

GAI-LMMs 可以提供丰富的医学知识和案例模拟,对医学生的临床思维能力和实践技能可能产生深远影响。在临床思维方面,GAI-LMMs 可以模拟不同的临床情景,并为学生提供解决问题的思路和方法。通过这些模型,学生可以接触到各种临床案例,从而培养出分析问题、诊断疾病和制定治疗方案的能力。然而,尽管 GAI-LMMs 可以提供大量的理论知识和案例模拟,但它们无法完全替代真实的临床实践经验。临床实践往往涉及到与患者的交流、身体检查、实际治疗和团队合作等方面,这些是模型无法完全模拟的。

因此,在使用 GAI-LMMs 提升医学生的临床思维和实践能力时,关键在于平衡模拟训练和实际临床经验的结合。这意味着在教育中需要强调模型辅助学习与实地实习的相辅相成,以确保学生获得理论知识和实践技能的双重提升。同时,还应培养学生的判断力和决策能力,让他们有能力抉择何时应该依赖模型,何时需要借助实践经验。

4.3 公平公正的教学评估

随着 GAI-LMMs 的兴起,医学教育者们面临着新的问题:如何设计一个不受 GAI-LMMs 影响的评估系统来评估学生对临床专业知识的掌握和熟练程度?由于线上任务不易管控,应尽可能进行现场评估,同时进行充分的监督,杜绝使用 GAI-LMMs 作弊的行为。除了加强现场评估的监督外,还要求教育者改变传统的笔试考察的方法,开发并应用超出 GAI-LMMs 能力范围的评估方案,比如:考察批判性临床思维、标准化患者(standard patients, SP)的诊治、临床技能实操等。另外,在医学研究评估方面,学期论文、科研项目、学术论文等这些书面文本和想法并不一定来源于学生自己的成果,且通过教师和软件均不易分辨是否产自 AI,由此可引起一定的欺骗性问题,因此口头评估(如毕业答辩等)在 AI 时代的背景下起到了越来越重要的作用。

4.4 GAI-LMMs 作为论文作者的争议性

由于具有强大的语言生成功能,GAI-LMMs 引发了人们对未来学术论文和研究成果的质疑,其在医学研究中的贡献近来也引起了广泛的争议。2022年12月,在医学数据库 medRxiv 发布了一篇关于利用 ChatGPT 进行医学教育的预印本,其中 ChatGPT 被列为论文的 12 位作者之一。引发了人们的思考:GAI-LMMs 作为学术论文的作者是否合适?期刊编辑、研究人员和出版商就这一问题

以及 GAI-LMMs 在已发表文献中的地位进行了讨论。*Nature* 新闻团队联系的出版商一致认为,ChatGPT 这样的人工智能并不符合研究作者的标准,因为它们不能对科学论文的内容和完整性负责。但是一些出版商表示,GAI-LMMs 对论文写作的贡献可以在作者名单以外的部分得到认可,比如致谢^[16]。GAI-LMMs 正在广泛传播和应用,为了避免类似争议引起的问题,有关部门应尽快出台学术论文中使用 GAI-LMMs 的相关政策。

4.5 医学生人文素养教育

随着越来越多的机械且琐碎的任务被 AI 接管,人们对 AI 不可替代的人文技能(如同理心、沟通等)的需求和重视程度也越来越高。GAI-LMMs 的出现为医学生的人文素养教育提供了新的可能性,但同时也带来了一些思考与挑战。GAI-LMMs 能够辅助医学生更深入地理解人文素养领域的概念和内容。它们可以提供广泛的知识、历史背景和伦理原则等,为学生提供全面的知识视角,有助于培养医学生的综合素养,使学生能够更好地理解医疗工作中的人性化、沟通、伦理等方面的重要性。然而,值得注意的是,尽管 GAI-LMMs 能够提供大量信息,但医学生的人文素养并非仅靠信息获取就能完全培养。人文素养涉及对道德伦理的理解、对患者需求的敏感性以及人性化医疗的实践。这些方面需要更多的实践、模拟场景和人际互动来培养,而模型无法完全替代这些实际经验。此外,对于 GAI-LMMs 生成的与人文素养相关的内容,医学生需要具备批判性思维和分析能力,以筛选和评估信息的可靠性和适用性。这样的技能也是人文素养的一部分,需要在医学教育中得到强调和培养^[13]。

因此,GAI-LMMs 对医学生人文素养教育的影响是双重的。它们提供了广泛的信息和视角,为医学生的人文素养教育带来了便利和丰富性。然而,使用这些模型的同时,也需要结合实践、模拟、人际互动等方式,以全面培养医学生的人文素养,要求学生不仅掌握相关知识,更应在临床实践中体现出对患者和医疗伦理的理解和尊重。在 GAI-LMMs 盛行的当下,医学教育者也需要更多地注重培养医学生的人文关怀、沟通能力和伦理责任感,以确保未来医务人员能够综合运用技术和人文素养,提供更全面和细致的医疗服务。

4.6 人工智能伦理教学

GAI-LMMs 这样的人工智能只是一种工具,使用责任在于使用者本人。作为社会的一员,我们有责任确保人工智能的使用既有效又符合道德规范。鉴于目前缺乏关于人工智能在医学中使用的指导方针以及上述提及的 GAI-LMMs 应用于医学教育面临的挑战,有学者建议医学教育应始终包含人工

智能伦理方面的教学^[17]。

5 小结

人工智能现如今已渗透入医学学习、临床管理以及日常生活的方方面面。在医疗实践不久的将来,数字工具、应用程序和电子病历将成为医疗保健系统的重要组成部分。医学教育机构以及教育者们需让医学生和毕业生具备在这一 AI 时代茁壮成长的能力。

GAI-LLMs 在医学教育中的应用是一项新兴技术,其引入可能为医学教育实践带来颠覆性的影响。值得注意的是,尽管 GAI-LLMs 由于其广泛的知识库和语言生成能力可作为学生学习和教育工作者教学的助手,但它们不能取代人类的智慧或复制人类思维的复杂性,也给医学教育带来了一定的问题和挑战。因此,教育者们应让学生清楚认识到这一新兴技术的两面性,并传授合理使用 GAI-LLMs 的技能与方法,同时相关部门也应尽快出台相关政策,确保 GAI-LLMs 应用于医学教育的科学性及合理性。医学教育机构和医学教育者应积极开发、探索、升级 GAI-LLMs 辅助的高等教育教学模式,将 GAI-LLMs 有机地融入学习与教学实践,践行习近平总书记关于数字中国的重要指示精神,推动高等教育的数字化建设,培养新一代数字公民。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 薛二勇,李健,黎兴成.推进中国教育数字化的战略与政策[J].中国电化教育,2023,30(1):25-32.
- [2] 吴书光,陈翠荣.中国高等教育数字化良好转型生态系统构建研究[J].临沂大学学报,2023,45(6):166-176.
- [3] 张熙,杨小汕,徐常胜.ChatGPT 及生成式人工智能现状及未来发展方向[J].中国科学基金,2023,37(5):743-750.
- [4] Abd-Alrazaq A,Alsaad R,Alhuwail D,et al. Large Language Models in Medical Education: Opportunities,Challenges,and Future Directions[J].JMIR Med Educ,2023,9:e48291.
- [5] Bair H,Norden J. Large Language Models and Their Implications on Medical Education[J]. Acad Med, 2023,98(8):869-870.
- [6] 颜见智,何雨鑫,骆子焯,等.生成式大语言模型在医疗领域的潜在典型应用与面临的挑战[J].医学信息学杂志,2023,44(9):23-31.
- [7] Smith J,Choi PM,Buntine P. Will code one day run a code? Performance of language models on ACEM primary examinations and implications[J]. Emerg Med Australas,2023,35(5):876-878.
- [8] Tokuç B,Varol G. Medical Education in the Era of Advancing Technology[J]. Balkan Med J, 2023, 40(6):395-399.
- [9] Moritz S,Romeike B,Stosch C,et al. Generative AI (gAI) in medical education: Chat-GPT and co [J]. GMS J Med Educ,2023,40(4):54.
- [10] Vignesh R,Pradeep P,Balakrishnan P. A Tête-à-tête with ChatGPT on the impact of artificial intelligence in medical education[J]. Med J Malaysia, 2023, 78(4):547-549.
- [11] Shaheen A,Azam F,Amir M. ChatGPT and the Future of Medical Education: Opportunities and Challenges [J]. J Coll Physicians Surg Pak, 2023, 33(10):1207.
- [12] Jowsey T,Stokes-Parish J,Singleton R,et al. Medical education empowered by generative artificial intelligence large language models[J]. Trends Mol Med, 2023,29(12):971-993.
- [13] Heng JJY,Teo DB,Tan LF. The impact of Chat Generative Pre-trained Transformer (ChatGPT) on medical education[J]. Postgrad Med J, 2023, 99(1176): 1125-1127.
- [14] Mohammad B,Supti T,Alzubaidi M,et al. The Pros and Cons of Using ChatGPT in Medical Education: A Scoping Review [J]. Stud Health Technol Inform, 2023,305:644-647.
- [15] 瞿星,杨金铭,陈滔,等. ChatGPT 对医学教育模式改变的思考[J]. 四川大学学报(医学版),2023,54(5): 937-940.
- [16] Stokel-Walker C. ChatGPT listed as author on research papers; many scientists disapprove[J]. Nature, 2023,613(7945):620-621.
- [17] Weidener L,Fischer M. Teaching AI Ethics in Medical Education: A Scoping Review of Current Literature and Practices[J]. Perspect Med Educ, 2023, 12(1): 399-410.

(收稿日期:2024-05-12)