

DOI: 10.61189/059703zeipzv

· 综述 ·

医学数字人 GPT 的研究现状及展望

韦 球¹, 蒋维芃², 杨超勉¹, 白春学^{2*}

1. 南宁市第一人民医院呼吸与危重症医学科, 南宁 530022

2. 复旦大学附属中山医院呼吸与危重症医学科, 上海 200032



[摘要] 当前,医学数字人 GPT 的研究主要集中于其在医疗健康领域中的应用。这种技术能够通过自动解读医疗影像和电子病历,帮助医生更快更准确地做出诊断,提高诊断精度和效率。同时,它还可以提供个性化的健康教育和患者关怀,从而改善患者的体验,并提高患者的满意度和依从性。此外,GPT 能够自动化处理大量的文本数据,显著降低医疗人员的工作量,降低医疗成本。其预诊断和健康管理功能也有助于预防和早期发现疾病,减少后期治疗的成本。在科研方面,GPT 可以识别医疗数据中的异常现象,帮助科研人员发现新的治疗方法或疾病预测模型。它还能根据已有的医学知识自动生成新的假说和实验方案,为科研人员提供实用的建议。此外,GPT 还可以通过推理和逻辑思维,帮助解决医学难题,促进科研的进展。展望未来,医学数字人 GPT 有着广阔的发展前景。随着技术的不断进步和医疗需求的日益增长,GPT 在医疗健康领域的应用将更加广泛和深入。它不仅可以提高医疗服务的质量和效率,还可以推动医学科研的创新和发展。同时,随着人们对隐私和数据安全的关注度不断提高,如何确保敏感医疗数据的安全存储和处理,避免数据泄露的风险,维护患者隐私和数据合规性,也将是医学数字人 GPT 未来发展的重要考虑因素。

[关键词] 医学数字人;GPT;自然语言处理;医疗;保健

[中图分类号] R-1/TP399 **[文献标志码]** A

Current status and prospects of digital human GPT in medicine

WEI Qiu¹, JIANG Weipeng², YANG Chaomian¹, BAI Chunxue^{2*}

1. Department of Pulmonary and Critical Care Medicine, The First People's Hospital of Nanning, Nanning 530022, Guangxi, China

2. Department of Pulmonary and Critical Care Medicine, Zhongshan Hospital, Fudan University, Shanghai 200032, China

[Abstract] Research on digital human GPT in medicine mainly focuses on its applications in healthcare. This technology can help doctors make diagnoses faster and more accurately by automatically interpreting medical images and electronic medical records, thereby improving diagnostic accuracy and efficiency. At the same time, it can provide personalized health education and patient care, which can improve the patient experience and increase patient satisfaction and compliance. In addition, GPT can automate the processing of large amounts of textual data, significantly reducing the workload of medical staff and reducing medical costs. Its prediagnosis and health management functions can also help detect and prevent diseases early, reducing the cost of later treatment. When applied in scientific research, GPT can identify anomalies in medical data and help researchers discover new treatments or disease prediction models. It can also automatically generate new hypotheses and protocols based on existing medical knowledge, providing practical recommendations for researchers. In addition, GPT can help solve medical problems and promote the progress of scientific research through reasoning and logical thinking. Looking forward to the future, digital human GPT in medicine has broad development prospects. With the continuous advancement of technology and the increasing demand for medical care, the application of GPT in the medical and health fields will be deeper and more extensive. It can not only improve the quality and efficiency of medical services but also promote innovation in and the development of medical research. At the same time, with the increasing demand for

[收稿日期] 2024-02-28

[接受日期] 2024-03-25

[基金项目] 上海市健康科普人才能力提升专项 (JKKPYC-2023-A20), 上海市科学技术委员会上海工程技术研究中心建设计划 (20DZ2254400), 上海市科技创新行动计划 (21DZ2200600), 上海市市级科技重大专项 (ZD2021CY001), 上海市临床重点专科建设项目 (shslczdk02201), 南宁市公立医院高质量发展项目 (南宁市呼吸系统疾病防治体系建设项目)。Supported by Project of Promoting Ability of Medical Science Popularization for Young Talents in Shanghai (JKKPYC-2023-A20), Establishment of Shanghai Engineering Technology Research Center of Science and Technology Commission of Shanghai Municipality (20DZ2254400), Shanghai Action Plan for Science, Technology and Innovation (21DZ2200600), The Shanghai Municipal Science and Technology Major Project (ZD2021CY001), Shanghai Municipal Key Clinical Specialty (shslczdk02201), High-quality Development Projects of Public Hospitals in Nanning (Construction of Respiratory Disease Prevention and Control in Nanning) .

[作者简介] 韦 球, 博士, 副主任医师. E-mail: weiqiu2017@163.com

* 通信作者 (Corresponding author). Tel: 021-64041990, E-mail: bai.chunxue@zs-hospital.sh.cn

privacy and data security, we must understand how we can ensure the safe storage and processing of sensitive medical data, avoid the risk of data leakage, and maintain patient privacy and data compliance to ensure the future development of digital human GPT in medicine.

[Key Words] digital human in medicine; GPT; natural language processing; medical; healthcare

随着科技的不断发展,生成式人工智能技术如GPT在医疗领域的应用逐渐受到关注。GPT是一种基于大规模文本数据进行训练的生成式人工智能技术,具有生成逻辑性和连贯性文本的能力。在医学领域,GPT被应用于处理和分析大量的医疗文本数据,如病历、科研论文和临床试验报告等。通过对这些数据的深度学习和理解,医学数字人GPT能够模拟医生的专业知识和经验,提供医学建议、辅助诊断和治疗方案等。医学数字人GPT的研发和应用为医学研究、诊断和治疗带来了革命性的变革。它不仅可以提高医生的诊疗效率和准确性,还可以为患者提供更个性化、精准的医疗服务。同时,医学数字人GPT也在医学教育、药物研发和健康管理等领域展现出广阔的应用前景。然而,医学数字人GPT的研发和应用仍面临一些挑战和问题,如数据的质量和可用性、隐私保护和伦理问题等。为了解决这些问题,需要不断探索新的技术和方法,推动医学数字人GPT技术的安全、可靠和广泛应用。

1 医学数字人 GPT 研究现状

目前,医学数字人GPT的研究正处于蓬勃发展的阶段。随着人工智能技术的不断进步,GPT在医学领域的应用已经受到广泛关注,并取得了一系列重要成果。一些前沿的研究机构已经成功构建了基于GPT的医学数字人模型。这些模型通过对海量的医学文献、临床数据和医学知识进行深度学习和训练,展现了在疾病诊断、治疗方案推荐、药物研发等方面的强大能力。它们能够模拟医生的专业知识和经验,提供准确的医学建议,辅助医生进行临床决策,从而提高诊疗效率和准确性。此外,医学数字人GPT在医学教育、健康管理和医学研究等领域也展现出了广阔的应用前景。它们可以作为智能教学助手,为医学实习生和规培生提供实践操作的机会,帮助他们更好地理解 and 掌握医学知识。同时,医学数字人GPT还可以根据用户的基因、生理和病理指标等信息,为用户提供个性化的健康建议和预防措施,帮助人们更好地管理自己的健康。然而,医学数字人GPT的研究仍面临一些挑战和问题。

1.1 国际医学数字人 GPT 研究现状 随着人工智能技术的不断进步和全球范围内对医疗健康领域的持续关注,GPT在医学领域的应用逐渐受到国际社会的广泛认可,并取得了一系列重要成果。在国际上,许多知名的医疗机构、研究团队和高校都积极投入到医学数字人GPT的研究中^[1-3]。他们应用大规模医学文本数据、临床数据和医学知识库进行模型训练和优化,致力于构建适应不同语境和医学需求的GPT模型。这些模型在疾病诊断、治疗方案推荐、药物研发等方面展现出了强大的能力,为医生提供了准确的辅助决策支持,提高了诊疗效率和准确性^[4-6]。此外,国际合作与交流也在推动医学数字人GPT的研究进程。各国的研究机构通过共享数据、合作研究和经验交流等方式,共同探索GPT在医学领域的应用潜力和创新方向。这种跨国合作不仅促进了技术的快速发展,还为解决全球性的医疗健康问题提供了更多可能性。目前,国际社会需要加强合作与协调,共同制定相关标准和规范,推动医学数字人GPT的安全、可靠和广泛应用。

1.2 中国医学数字人 GPT 研究现状 随着人工智能技术的不断发展和中国医疗健康需求的增加,医学数字人GPT在国内的应用和研究逐渐受到关注。国内已经有一些医疗机构、高校和企业开始投入到医学数字人GPT的研究中。他们应用深度学习、自然语言处理(natural language processing, NLP)等技术,构建基于GPT的医学数字人模型,通过对海量的医学文献、临床数据和医学知识进行学习和训练,提高模型在医学领域的准确性和应用效果^[7-8]。在医学领域,GPT被应用于处理和分析大量的医疗文本数据,包括病历、科研论文等。通过对这些数据的深度学习和理解,医学数字人GPT能够模拟医生的专业知识和经验,提供医学建议、辅助诊断和治疗方案等,从而提高医生的诊疗效率和准确性,为患者提供更个性化、精准的医疗服务。此外,国内的研究机构还积极探索医学数字人GPT在医学教育、健康管理和医学研究等领域的应用。例如,应用GPT构建智能教学系统,为医学学生提供实践操作的机会和个性化的学习体验,用于教学查房;应用GPT分析用户的基因、生理和病理生理指标等信息,为其提供个性化的健康建议和预防措施^[9-10]。

2 医学数字人 GPT 应用场景

2.1 赋能教育与培训 医学数字人与 GPT 的结合,将极大提升教育与培训的质量和效率,让学习更加个性化、高效。这不仅是技术的进步,更是医学教育的 1 次重要革新^[10-11]。在教育领域,GPT 还可以作为医学生的学习助手,为他们互动式解答问题、模拟考试场景等。在智能辅助诊断与治疗方面,医学数字人 GPT 能够通过学习和理解大量的病例数据和诊治经验,协助医生更精确地诊断和制定更有效的治疗方案。GPT 可以根据患者的症状描述、检查结果和其他相关信息,为医生提供初步的诊断意见和辅助决策支持。这有助于提高医生的诊断准确率和治疗效果,从而更好地满足患者的需求。

2.2 辅助医学研究 通过深度学习和 NLP 技术,GPT 能够迅速分析大量的医学文献和临床数据,为医学研究人员提供准确、全面的信息支持。在医学研究中,数据分析和文献综述是非常重要的环节^[12-13]。然而,由于医学文献数量庞大、内容复杂,研究人员往往需要花费大量时间和精力进行筛选和整理。而医学数字人 GPT 则能够通过 NLP 技术,自动识别和提取文献中的关键信息,帮助研究人员快速了解研究领域的前沿动态和最新进展。此外,医学数字人 GPT 还可以辅助医学研究人员进行实验设计和分析数据。通过深度学习,GPT 能够预测实验结果和趋势,为研究人员提供有价值的参考意见。同时,GPT 还能够对实验数据进行自动化处理和分析,提高数据处理的准确性和效率^[14-15]。

2.3 辅助病情分析与预诊断 在初级医疗环节,基于 GPT 的模型可以分析病人的症状描述,并给出可能的诊断,以辅助医生进行更准确的判断。这对于在疫情等特殊情况下进行远程医疗服务,或者在医疗资源紧张的地区提供医疗服务,具有重要的价值。

如我们研究的 PNapp 5A 和 OSapp 5A(获得中国医学装备协会第九届临床创新与发明大赛重庆赛区二等奖),这种程序结合了深度学习和 NLP 技术,能够理解并解析复杂的医学文本和数据。在病情分析方面,医学数字人 GPT 可以迅速处理患者的病历、症状描述和检查结果等大量信息。通过深度分析这些数据,GPT 能够提取出关键特征,帮助医生更准确地判断患者的病情。此外,GPT 还能根据已有的医学知识和经验,为医生提供最佳的诊断思路 and 方向,从而辅助医生做出更明智的决策。在预诊断方面,医学数字人 GPT 可以基于患者的初步症

状描述或检查结果,结合医学知识库中的大量数据,给出可能的疾病类型和风险程度^[16-18]。这种预诊断的能力在疫情等特殊情况下尤为重要,如 nCapp 5A 可以帮助医生快速筛选出高风险患者,优先安排检查和治疗。同时,GPT 还可以根据患者的病情发展趋势,预测可能的并发症和恶化风险,从而提醒医生采取更积极的治疗措施。

2.4 辅助医学影像诊断 医学数字人 GPT 在医学影像诊断方面也展现出了应用潜力。由于其结合了深度学习和医学影像处理的能力,使得机器能够辅助医生进行更快速、更准确的影像诊断。通过训练和优化,医学数字人 GPT 能够识别和分析医学影像中的关键特征,如病变位置、形态、大小等,从而帮助医生判断疾病的类型和严重程度^[16,19]。例如,我们研发的 BAIMGPT 即可赋能诊断肺结节中的早期肺癌。此外,GPT 还能根据已有的医学知识和经验,为医生提供可能的诊断思路 and 方向,进一步辅助医生做出准确的诊断决策。在实际应用中,医学数字人 GPT 可以处理各种类型的医学影像,如 X 光片、CT 扫描、MRI 等。通过自动化处理和分析这些影像数据,GPT 能够减轻医生的工作负担,提高工作效率,同时减少人为因素导致的误诊和漏诊。

2.5 赋能远程医疗服务 医学数字人 GPT 可为远程医疗服务赋能。借助这种先进的 NLP 技术,不仅能够实现医生和患者之间的实时交流,还能提供更为精准、高效的医疗服务。首先,医学数字人 GPT 能够模拟医生的专业知识和经验,对患者进行初步的问诊和咨询。患者可以通过文字、语音或视频与数字人 GPT 交互,描述自己的症状和病情,GPT 则能够分析这些信息,给出可能的诊断和建议。其次,医学数字人 GPT 可以辅助医生进行远程诊断和治疗方案的制定^[20]。医生可以通过 GPT 获取患者的相关病历资料、检查结果和医学影像等信息,进行全面的分析和评估。GPT 还能根据医生的指示和要求,生成个性化的治疗方案和注意事项,为患者提供针对性的远程医疗服务^[9,21]。此外,医学数字人 GPT 还能用于远程医疗教育和科普宣传。GPT 可以生成易于理解和接受的医学知识和健康建议,帮助患者和公众提高健康意识和自我管理 ability。简言之,医学数字人 GPT 可极大地扩展医疗资源的覆盖范围,提高医疗服务的可及性和质量,有望为更多人带来更好的医疗体验和健康保障。

2.6 辅助诊断与决策支持 医学数字人 GPT 可以结合深度学习和 NLP 的能力,使得机器能够理解并解析复杂的医学文本和数据,从而为医生提供精准

的诊断和决策支持^[4,7]。在辅助诊断方面,医学数字人 GPT 可以迅速处理患者的病历、症状描述、检查结果等大量信息,并通过深度分析这些数据,提取出关键特征,帮助医生更准确地判断患者的病情。此外,GPT 还能根据已有的医学知识和经验,为医生提供可能的诊断思路 and 方向,从而辅助医生做出更明智的诊断决策。在决策支持方面,医学数字人 GPT 可以基于患者的病情和诊断结果,结合医学知识库中的大量数据,为医生提供个性化的治疗方案和注意事项。这种决策支持的能力可以帮助医生避免遗漏重要信息或考虑不周的情况,从而提高治疗的效果和安全性。医学数字人 GPT 在辅助诊断与决策支持方面的应用,将极大地提高医疗服务的效率和准确性,为患者带来更好的就医体验。然而,需要注意的是,GPT 的辅助作用并不能完全替代医生的专业判断和经验,医生仍然需要对诊断结果进行综合分析和确认,以确保最终决策的准确性和可靠性。

2.7 辅助医疗健康管理工作 医学数字人 GPT 在辅助医疗健康管理工作方面具有重要作用。(1)个性化健康计划。基于患者的健康数据、生活习惯和目标,医学数字人 GPT 能够为患者制定个性化的健康管理计划,包括饮食建议、运动处方、生活习惯调整等,帮助患者改善健康状况,降低患病风险^[22-23]。(2)健康监测与预警。GPT 可以持续监测患者的健康状况,并根据数据分析结果,及时发现潜在的健康问题。当检测到异常情况时,GPT 能够向患者发出预警信息,提醒患者及时就医或调整健康管理方式^[24-25]。(3)健康教育。GPT 能够根据患者的需求和兴趣,提供个性化的健康教育内容。帮助患者更好地理解健康知识,提高健康意识,更好地进行自我管理。(4)持续沟通与支持。GPT 可以作为患者的健康伴侣,全时空提供健康咨询和支持。这种持续的沟通和支持有助于增强患者的健康管理信心和能。

2.8 辅助电子健康记录(EHR)的解读与生成 医学数字人 GPT 在辅助 EHR 的解读与生成方面发挥着重要作用。(1)自动化解读:医学数字人 GPT 能够自动解读和分析 EHR 中大量数据,包括健康记录和病史、诊断结果、用药记录等。GPT 能够提取出关键信息,帮助医生更快速地理解患者的病情和治疗过程。(2)结构化生成:GPT 可以根据医生的需求和指示,自动生成结构化的 EHR。其中包括患者的症状、体征和检查结果等信息,方便医生进行查阅和共享。这种结构化的生成方式可以提高 EHR 的一

致性和可读性^[9]。(3)智能提醒与辅助:在记录过程中,GPT 还能根据已有的医学知识和经验,为医生提供智能提醒和辅助。例如,当医生输入某个症状时,GPT 可以自动显示与该症状相关的疾病、检查项目或治疗建议,帮助医生更全面地考虑和记录患者的病情^[26]。(4)数据整合与可视化:GPT 还能将 EHR 与其他医疗信息系统进行整合,实现数据的共享和交换。同时,GPT 还可以将复杂的数据以可视化的方式呈现出来,如生成图表、报告等,帮助医生更直观地了解患者的病情和治疗效果^[19]。医学数字人 GPT 在辅助 EHR 的解读与生成方面的应用,将极大地提高医疗服务的效率和准确性,同时,GPT 的自动化和智能化功能也可以减轻医生的工作负担,提升医疗服务的整体质量。

2.9 赋能患者关怀和健康教育 医学数字人 GPT 还能够为患者提供更加个性化、高效和便捷的医疗服务。首先,在患者关怀方面,医学数字人 GPT 可以通过 NLP 技术,与患者进行实时交流,了解患者的需求和病情。GPT 能够根据患者的情绪状态和病情变化,提供及时的关怀和支持,帮助患者缓解焦虑和恐惧情绪,增强治疗信心^[27-28]。其次,在健康教育方面,医学数字人 GPT 可以根据患者的病情和治疗方案,生成个性化的健康教育材料,包括疾病知识、治疗方案、药物使用说明、饮食和运动建议等,帮助患者更好地理解和管理自己的病情。GPT 还可以根据患者的反馈和需求,不断调整和优化健康教育内容,确保其针对性和有效性。此外,医学数字人 GPT 还可以通过智能语音技术,为视力障碍或阅读困难的患者提供语音健康教育服务^[27]。这种服务方式可以打破传统文字材料的限制,让患者更加方便地获取健康知识。这种创新的医疗服务模式,有望为更多患者带来更好的医疗体验和健康保障。同时,GPT 的应用也有助于推动医疗行业的数字化转型和智能化升级。

2.10 辅助跨语言医学交流与翻译 医学数字人 GPT 在辅助跨语言医学交流与翻译方面展现出强大的潜力^[9-10],包括(1)实时翻译:医学数字人 GPT 能够实时将医学术语、病历记录、研究报告等内容从 1 种语言翻译成其他语言。这种即时翻译的能力极大地促进了不同语言背景的医生、研究人员和患者之间的交流。(2)术语准确性:GPT 经过大量医学文献和术语的训练,能够确保在翻译过程中使用准确、专业的医学术语。这对于保持医学信息的准确性和一致性至关重要。(3)文化敏感性:医学数字人 GPT 还具备一定程度的文化敏感性,能够在翻译时

考虑到不同文化背景下的表达方式和理解习惯,利于确保跨语言医学交流的文化恰当性。(4)多模态交流:除了文本翻译外,医学数字人 GPT 还可以辅助图像、视频等多媒体内容的跨语言解读,利于医学知识、手术演示等内容的传播更加直观和高效。(5)教育与研究:GPT 的跨语言翻译功能也极大地促进了医学教育和研究的国际化。医生和研究人员可以更容易地获取和分享全球范围内的医学知识和最新研究成果,这对于推动全球医学领域的合作与进步具有重要意义。

2.11 辅助医学研究与文献分析 医学数字人 GPT 在辅助医学研究与文献分析方面拥有巨大的应用潜力。(1)文献筛选与整理:研究人员常常需要花费大量时间进行文献的筛选和整理。GPT 可以帮助研究人员快速筛选出与研究主题相关的文献,并按照研究目的进行分类整理,从而提高文献检索和阅读的效率。(2)内容提炼与总结:医学文献往往篇幅冗长、专业术语众多,不利于非专业领域的研究人员的理解。GPT 可以对文献进行深度分析,提炼出关键信息并进行总结,帮助研究人员更快地掌握文献的核心内容和研究成果。(3)辅助实验设计:GPT 可以根据已有文献和研究目标,为研究人员提供实验设计的思路和建议,从而辅助他们制定出更加科学合理的实验方案。(4)数据分析与挖掘:通过对医学文献和实验数据进行深度挖掘和分析,GPT 可以帮助研究人员发现隐藏在数据中的关联和规律,从而推动医学研究的新发现和新突破。(5)写作与发表辅助:GPT 可以帮助研究人员在写作过程中进行语法检查、句式优化等工作,提高论文的质量和可读性。此外,GPT 还可以为研究人员提供投稿建议和策略,辅助他们更好地将研究成果发表在国际一流的学术期刊上。

3 医学数字人 GPT 的限制及原因

3.1 数据偏差和局限性 医学数字人 GPT 在数据偏差和局限性方面可能存在一些问题^[1-2,5]。(1)训练数据的不完整性:GPT 的训练数据主要来源于已有的医学文献和病例报告等,数据可能存在不完整的情况。例如,某些罕见疾病或新兴疾病的案例可能较少,导致 GPT 在这方面的知识储备有限。(2)数据质量和标注问题:医学数据的质量和标注准确性对 GPT 的性能至关重要。如果训练数据中存在错误、冗余或不一致的标注,GPT 可能会学习到错误的信息,从而影响其判断的准确性。(3)文化和地域差异:医学实践和文化背景密切相关,不同地域和文

化的医学数据可能存在差异。GPT 在训练过程中可能主要基于某一特定文化或地域的数据,导致其与其他文化或地域的应用中存在一定的局限性。(4)技术更新和医学进展:医学领域的知识和技术不断更新,新的研究成果和治疗方法不断涌现。GPT 的训练数据可能无法及时涵盖最新的医学进展,导致其在这方面的信息滞后。(5)隐私和伦理问题:医学数据涉及患者隐私和伦理问题,因此在收集和使用过程中需要严格遵守相关法律法规和伦理规范。这可能导致可用于训练 GPT 的医学数据受到限制,进而影响其性能和应用范围。

3.2 缺乏临床实践经验 医学数字人 GPT 缺乏临床实践经验的原因主要有以下几方面^[2-4,5-8]。(1)数据来源限制:GPT 主要基于大量的文本数据进行训练,虽然包含了丰富的医学知识和信息,但是往往缺乏真实的临床实践数据。临床实践经验的获取需要直接参与患者的诊疗过程,观察疾病的发展变化,以及了解患者的反馈和治疗效果,这是单纯的文本数据无法提供的。(2)模型本质:GPT 作为一种语言模型,其核心能力在于理解和生成自然语言文本。虽然它可以通过学习医学文献和病例报告等间接获取一定的医学知识,但这些知识往往停留在理论层面,缺乏实际的临床应用和实践验证。(3)技术挑战:将 GPT 应用于临床实践需要解决一系列技术挑战,如数据的准确性和完整性、模型的可靠性和稳定性,以及与现有医疗系统的集成等。这些技术问题的解决需要时间和持续的研究投入。(4)伦理和法规限制:医学领域涉及患者的生命安全和隐私保护等敏感问题,因此 AI 技术在医学领域的应用需要严格遵守相关的伦理和法规要求。这在一定程度上限制了 GPT 等 AI 技术在临床实践中的应用范围和自主性。

3.3 隐私和安全问题 医学数字人 GPT 在隐私和安全方面存在的问题^[29-31]主要包括,(1)数据泄露风险:GPT 的训练需要大量的数据,这些数据往往包含用户的用户信息和敏感信息。如果这些数据在收集、存储和处理过程中没有得到充分的保护,就可能发生数据泄露事件,导致用户的隐私被侵犯。(2)模型漏洞:GPT 作为一种复杂的机器学习模型,可能存在一些未知的漏洞和缺陷。这些漏洞和缺陷可能被恶意攻击者应用,通过输入特定的数据来窃取用户的隐私信息或破坏系统的正常运行。(3)不当使用:医学数字人 GPT 的使用者如果没有经过充分的培训和指导,可能会在使用过程中违反隐私保护原则,例如将患者的隐私信息泄露给无关人员或用

于非法目的。(4)法律和监管缺失:目前,对于医学数字人 GPT 等人工智能技术在隐私和安全方面的监管和法律规范还相对滞后。这可能导致一些不法分子利用法律漏洞进行隐私侵犯和非法活动。

3.4 无法替代自然人医生的判断和情感支持 医学数字人 GPT 虽然在医学领域具有广泛的应用潜力,但仍无法完全替代自然人医生的判断和情感支持,主要原因如下^[2-3]:(1)缺乏临床实践经验。医学数字人 GPT 主要基于大量的医学文献和数据进行训练,但缺乏真实的临床实践经验。自然人医生通过长期的临床实践积累了丰富的经验,能够更准确地判断病情和制定治疗方案。(2)情感支持的缺失。医学不仅仅是科学和技术,还涉及到患者的情感和心理需求。自然人医生能够与患者建立信任和共情,提供情感支持和安慰,这是医学数字人 GPT 无法替代的。(3)伦理和法规限制。医学领域涉及患者的生命安全和隐私保护等敏感问题。虽然医学数字人 GPT 可以处理和分析医学数据,但在涉及伦理决策和法规遵守方面,仍需要自然人医生的参与和判断。(4)技术局限性。尽管 GPT 等人工智能技术取得了显著进展,但仍存在一定的技术局限性。例如,对于复杂病例和罕见疾病的诊断,医学数字人 GPT 可能无法提供准确和可靠的判断。(5)人机协作的重要性。医学数字人 GPT 可以作为自然人医生的辅助工具,提供信息支持和分析建议,但在决策过程中仍需要自然人医生的最终判断和决策。人机协作能够结合人工智能的优势和自然人医生的经验,提高医疗服务的质量和效率。

3.5 技术更新和维护成本 医学数字人 GPT 在技术更新和维护成本方面存在的问题主要有,(1)技术更新迅速:人工智能领域的技术发展日新月异,GPT 等 NLP 模型也在不断更新迭代。为了保持医学数字人 GPT 的先进性和竞争力,需要不断进行技术更新和升级,这涉及到研发团队的投入、新技术的引入以及模型的重新训练等,都会产生相应的成本。(2)数据维护成本高:随着医学知识的不断更新和新增病例的出现,需要持续收集、整理和更新医学文献、病例报告等训练数据,以确保 GPT 能够学习到最新的医学知识和经验。同时,数据的存储、管理和保护也需要投入相应的成本。(3)模型优化和调整:在实际应用过程中,医学数字人 GPT 可能会遇到各种挑战和问题,如数据偏差、模型泛化能力不足等。为了解决这些问题,需要对模型进行优化和调整,包括改进算法、调整模型参数等。这些工作都需要专业的研发团队和技术支持,因此会产生相

应的成本。(4)安全和隐私保护:医学数字人 GPT 处理的数据涉及患者的隐私和敏感信息,因此在技术更新和维护过程中需要严格遵守相关的安全和隐私保护要求。这包括加强数据加密、访问控制等安全措施,以及定期进行安全漏洞检测和修复等,都会产生额外的成本。(5)培训和推广:随着医学数字人 GPT 功能的不断扩展和升级,需要对使用者进行持续的培训和推广,以确保他们能够充分了解和掌握新功能的使用方法。这也需要投入一定的人力和物力成本。

4 如何克服医学数字人 GPT 的局限性

4.1 提高数据来源的多样性和权威性 提高医学数字人 GPT 数据的质量和可用性是解决其局限性问题的主要方法^[1-3,5,31]。(1)数据验证与清洗:建立有效的数据验证机制,对采集到的医学数据进行验证和清洗。通过剔除错误、冗余和不完整的数据,确保数据的准确性和完整性。这可以通过使用自动化工具和算法,结合人工审核来实现。(2)数据标准化:制定统一的数据标准,如使用标准诊断编码系统、标准数据格式等。标准化有助于确保数据的一致性和可比性,使得不同来源和格式的数据能够整合和分析。(3)数据扩充与多样性:积极寻找和整合多个数据来源,增加数据的多样性和代表性。这包括医学文献、病例报告、临床试验数据等。通过扩充数据集,可以提高 GPT 模型的泛化能力和适应性。(4)持续学习与更新:建立医学数字人 GPT 的持续学习机制,使其能够不断从新的医学数据和知识中学习和更新。这可以通过定期重新训练模型、引入新的训练数据和技术来实现。(5)数据备份与恢复:建立定期的数据备份机制,并进行数据恢复测试。这可以确保在意外数据损坏或丢失的情况下,能够及时恢复数据,保障数据的可靠性和可用性。(6)用户反馈与迭代改进:鼓励用户在使用医学数字人 GPT 时提供反馈,特别是关于数据质量和可用性的问题。通过收集和分析用户反馈,可以及时发现和修复数据质量问题,并进行迭代改进。(7)数据隐私与安全保护:加强数据隐私和安全保护措施,确保医学数据的机密性和完整性。采用加密技术、访问控制和数据脱敏等手段,防止数据泄露和滥用。

4.2 提高数据清洗和预处理效率 提高医学数字人 GPT 数据清洗和预处理效率是确保数据质量和可用性的重要环节。(1)自动化工具和脚本:应用自动化工具和脚本进行数据清洗和预处理,可以显著提高效率。例如,使用 Python 等编程语言编写脚本,

自动化执行数据去重、缺失值填充、异常值检测等常规任务。(2)并行处理和分布式计算:如果处理的数据量非常大,可以考虑使用并行处理和分布式计算技术来加速数据清洗和预处理过程。通过将任务拆分成多个子任务,并同时多个计算节点上执行,可以显著提高整体处理速度。(3)数据管道和流程自动化:建立数据管道和流程自动化,将数据清洗和预处理步骤整合到1个连贯的流程中。这样可以确保数据从原始状态到最终处理状态的转换过程更加高效和可靠,减少手动干预和错误。(4)数据质量监控和校验:在数据清洗和预处理过程中,建立数据质量监控和校验机制。通过设置合适的指标和校验规则,可以及时发现和处理数据质量问题,避免在后续分析和建模中出现错误。(5)增量处理和更新:对于持续更新的医学数据,采用增量处理和更新的方式,而不是每次都重新处理整个数据集。这样可以节省计算资源和时间,提高数据清洗和预处理的效率。(6)优化算法和模型:针对特定的医学数据清洗和预处理任务,优化算法和模型的选择。例如,选择合适的缺失值填充算法、异常值检测算法等,可以提高处理速度和准确性。(7)使用专用工具和平台:考虑使用专用的数据清洗和预处理工具或平台,这些工具通常提供了丰富的功能和优化算法,可以简化数据清洗和预处理的流程,并提高处理效率。

4.3 使用标准化术语和编码系统 医学数字人 GPT 在使用标准化术语和编码系统时,会首先确保数据的准确性和一致性。通过深度学习技术,GPT 能够理解和识别各种医学术语,包括疾病名称、药物名称、手术名称等。同时,GPT 还可以应用编码系统,如 ICD-10、SNOMED-CT 等,对医学概念进行标准化处理,从而实现不同系统之间的数据交换和共享。此外,GPT 还具备 NLP 能力,可以与医生进行自然语言交互,理解医生意图和需求,并提供相应医学知识和建议。医生可以更加便捷地获取准确的医学信息,提高诊疗效率和质量。

4.4 提高数据标注和验证水平 医学数字人 GPT 在进行数据标注和验证时,需要采取一系列严谨步骤确保数据的准确性和可靠性^[8,32]。首先,在数据标注方面,需应用专业的医学知识和经验,对医学文本进行精确的标注,包括疾病名称、症状、治疗方案等关键信息。同时,还会采用多种标注方法,如实体标注、关系标注等,以捕捉文本中的复杂医学关系。在数据验证方面,我们会采用交叉验证、盲测试等多种方法来评估标注数据的准确性和一致性。

此外,还需应用先进的 NLP 技术,如深度学习模型,对标注数据进行自动验证和纠错,以进一步提高数据的准确性。

4.5 持续更新和维护 医学数字人 GPT 的持续更新和维护是1个重要的环节,以确保其能够持续提供准确、可靠的医学知识和建议:(1)数据更新。医学领域的知识和技术在不断进步和更新,因此,医学数字人 GPT 需要定期更新其训练数据,以纳入最新的医学研究成果、临床实践指南等。这可以通过与医学数据库、学术出版社等合作,获取最新的医学文献和数据资源。(2)模型优化。随着深度学习技术的不断发展,医学数字人 GPT 的模型也需要不断优化和改进。这可以通过改进模型架构、增加模型深度、调整超参数等方式实现,以提高模型的准确性、泛化能力和效率。(3)用户反馈。用户反馈是医学数字人 GPT 持续更新和维护的重要参考。通过收集用户的反馈和建议,可以及时发现和解决模型存在的问题和不足,从而不断完善和优化医学数字人 GPT 的功能和性能。(4)安全性保障。在更新和维护过程中,需要确保医学数字人 GPT 的安全性和稳定性。这包括对数据的安全性进行保护,防止数据泄露和滥用;对模型的稳定性进行监测和维护,确保模型能够持续、稳定地提供服务。总之,医学数字人 GPT 的持续更新和维护是1个综合性的工作,需要多方面的协作和努力。通过不断更新数据、优化模型、收集用户反馈和保障安全性等措施,可以确保医学数字人 GPT 始终能够提供最新、最准确的医学知识和建议。

4.6 提高数据隐私和安全保护水平 医学数字人 GPT 在实施数据隐私和安全保护时,会采取一系列严格的技术和管理措施。首先,在数据采集、存储和处理过程中,医学数字人 GPT 会严格遵守相关的隐私政策和法规,如 HIPAA(健康保险移植性和责任法案)等,确保患者的用户信息得到充分的保护。同时,采用加密技术对敏感数据进行加密处理,防止数据在传输和存储过程中被窃取或篡改。其次,医学数字人 GPT 会建立完善的数据访问控制机制,根据用户的角色和权限,对其访问的数据进行严格的限制和监控。只有经过授权的人员才能访问相应的数据,确保数据的机密性和完整性。此外,医学数字人 GPT 还会定期进行安全漏洞扫描和风险评估,及时发现和修复潜在的安全隐患。同时,建立完善的安全日志和监控机制,对系统的使用和操作情况进行实时记录和监控,以便及时发现和处理异常行为和安全事件。最后,医学数字人 GPT 还会与专业的

安全机构合作,共同应对网络安全威胁和挑战,确保系统的安全性和稳定性。

4.7 保护隐私和满足伦理要求 解决医学数字人 GPT 在隐私保护和伦理方面的问题,需要采取一系列综合措施^[29-31]。(1)立法和监管:制定严格的隐私法律和医疗数据保护法规,明确数据收集、使用、共享和存储的合法性和透明度要求。设立专门的监管机构来监督医学数字人 GPT 的开发和应用,确保其符合伦理规范和法律要求。(2)数据加密和安全存储:采用高级加密技术来保护医疗数据,确保在传输和存储过程中的安全。使用安全的云存储解决方案或本地加密存储存放数据,防止未经授权的访问和数据泄露。(3)访问控制和审计跟踪:实施严格的访问控制策略,只有经过授权的人员才能访问敏感数据。建立详细的审计跟踪系统,记录所有数据的访问、修改和删除操作,以便在必要时进行审查和追责。(4)匿名化和伪名化:在使用医学数据进行 GPT 训练时,应尽量采用匿名化或伪名化技术,去除或替换可以直接识别个体身份的信息。对于无法完全匿名化的数据,应使用脱敏技术来降低隐私泄露的风险。(5)透明度和用户授权:在使用医学数字人 GPT 时,应向用户清晰地说明数据的收集、使用和处理目的,并获得用户的明确授权。鼓励医疗机构和技术提供商建立用户友好的隐私政策和数据使用协议,提高用户对隐私保护的认知。(6)伦理审查和评估:在医学数字人 GPT 研究和应用的各个阶段,应进行全面的伦理审查和评估。建立多学科组成的伦理委员会,对 GPT 的潜在伦理风险进行讨论和决策,确保其应用符合广泛接受的伦理原则。(7)持续教育和培训:加强对医疗从业人员和技术开发者的隐私和伦理教育,提高他们的隐私保护意识和伦理责任感。鼓励行业组织和学术机构定期举办关于医学数字人 GPT 隐私和伦理问题的研讨会和培训活动。(8)建立患者权益保护机制:确保患者有权随时查看、更正和删除自己的医疗数据。提供方便的渠道,使患者能够报告隐私侵犯或伦理违规行为,并及时得到处理和回应。

5 应用前景

医学数字人 GPT 的发展前景非常广阔,随着人工智能技术的不断进步和医学领域的需求不断增长,它将在多个方面发挥越来越重要的作用。在医学研究和教育方面,医学数字人 GPT 可以应用其强大的 NLP 能力,帮助研究人员快速获取和分析大量的医学文献和数据,促进医学科研的进一步创新。

它还可以作为 1 种在线教育工具,为医学生和医生提供相关的知识和培训,模拟各种病例情况,提供虚拟患者供实践操作,从而提高医学教育的质量和效率。在临床诊断和治疗方面,医学数字人 GPT 可以通过与患者进行智能对话,了解他们的病情和需求,并给予专业的建议和指导,提高患者的满意度和治疗效果。它还可以分析患者的基因、生理和病理生理指标等信息,为医生提供个性化的诊断和治疗方案,推动精准医疗的发展。另外,在药物研发方面,医学数字人 GPT 也有着广阔的应用前景。它可以分析药物分子结构、药效数据等大量信息,预测药物的疗效和不良反应,提供针对性的药物设计方案,加快新药的研发进程,为人类提供更好的治疗手段。然而,随着医学数字人 GPT 的发展和应用,我们也需要关注一些挑战和问题。例如,如何确保数据的隐私和安全,如何避免算法偏见和错误等问题。因此,在未来的发展中,我们需要继续加强技术研发和监管,推动医学数字人 GPT 的健康发展。医学数字人 GPT 的发展前景非常广阔,它将在医学领域发挥越来越重要的作用,为人类健康事业做出更大的贡献。

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 韦球:论文撰写;蒋维芑:论文修改;杨超勉:论文修改;白春学:文章选题。

参考文献

- [1] GHASSEMI M, BIRHANE A, BILAL M, et al. ChatGPT one year on: who is using it, how and why? [J]. Nature, 2023, 624: 39-41.
- [2] HAUG C J, DRAZEN J M. Artificial intelligence and machine learning in clinical medicine, 2023 [J]. N Engl J Med, 2023, 388(13): 1201-1208.
- [3] RAJPURKAR P, CHEN E, BANERJEE O, et al. AI in health and medicine [J]. Nat Med, 2022, 28(1): 31-38.
- [4] SINGHAL K, AZIZI S, TU T, et al. Large language models encode clinical knowledge [J]. Nature, 2023, 620 (7972): 172-180.
- [5] SCHWALBE N, WAHL B. Artificial intelligence and the future of global health [J]. Lancet, 2020, 395(10236): 1579-1586.
- [6] ARONSON J K. When I use a word.. ChatGPT: a differential diagnosis [J]. BMJ, 2023, 382: 1862.
- [7] THIRUNAVUKARASU A J, TING D S J, ELANGOVA N K, et al. Large language models in medicine [J]. Nat Med, 2023, 29 (8): 1930-1940.
- [8] ESTEVA A, ROBICQUET A, RAMSUNDAR B, et al. A guide to deep learning in healthcare [J]. Nat Med, 2019, 25 (1): 24-29.

- [9] AYERS J W, POLIAK A, DREDZE M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum [J]. *JAMA Intern Med*, 2023, 183(6): 589–596.
- [10] 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): 971–973.
- [11] BOSCARDIN C K, GIN B, GOLDE P B, et al. ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity[J]. *Acad Med*, 2024, 99(1): 22–27.
- [12] HAMAN M, ŠKOLNÍK M. Using ChatGPT to conduct a literature review[J]. *Account Res*, 2023: 1–3.
- [13] WEI C H, ALLOT A, LAI P T, et al. PubTator 3.0: an AI-powered literature resource for unlocking biomedical knowledge [J]. *Nucleic Acids Res*, 2024: gkae235.
- [14] NOY S, ZHANG W. Experimental evidence on the productivity effects of generative artificial intelligence[J]. *Science*, 2023, 381(6654): 187–192.
- [15] ZHENG Z L, ZHANG O F, BORGS C, et al. ChatGPT chemistry assistant for text mining and the prediction of MOF synthesis[J]. *J Am Chem Soc*, 2023, 145(32): 18048–18062.
- [16] FINK M A, BISCHOFF A, FINK C A, et al. Potential of ChatGPT and GPT-4 for data mining of free-text CT reports on lung cancer[J]. *Radiology*, 2023, 308(3): e231362.
- [17] HAJ M E, BOUTOLEAU-BRETONNIÈRE C, CHAPELET G. ChatGPT's dance with neuropsychological data: a case study in Alzheimer's disease[J]. *Ageing Res Rev*, 2023, 92: 102117.
- [18] KERBAGE A, KASSAB J, EL DAHDAH J, et al. Accuracy of ChatGPT in common gastrointestinal diseases: impact for patients and providers [J]. *Clin Gastroenterol Hepatol*, 2023; S1542–3565(23)00946–1.
- [19] BHAYANA R, KRISHNA S, BLEAKNEY R R. Performance of ChatGPT on a radiology board-style examination: insights into current strengths and limitations [J]. *Radiology*, 2023, 307(5): e230582.
- [20] XUE Z W, ZHANG Y M, GAN W Y, et al. Quality and dependability of ChatGPT and DingXiangYuan forums for remote orthopedic consultations: comparative analysis [J]. *J Med Internet Res*, 2024, 26: e50882.
- [21] GOODMAN R S, PATRINELY J R, STONE C A JR, et al. Accuracy and reliability of chatbot responses to physician questions[J]. *JAMA Netw Open*, 2023, 6(10): e2336483.
- [22] VAISHYA R, MISRA A, VAISH A. ChatGPT: is this version good for healthcare and research?[J]. *Diabetes Metab Syndr*, 2023, 17(4): 102744.
- [23] VARGHESE J, CHAPIRO J. ChatGPT: the transformative influence of generative AI on science and healthcare [J]. *J Hepatol*, 2023: S0168–S8278(23)05039–0.
- [24] HAVER H L, AMBINDER E B, BAHL M, et al. Appropriateness of breast cancer prevention and screening recommendations provided by ChatGPT [J]. *Radiology*, 2023, 307(4): e230424.
- [25] WU X W, ZHANG B. ChatGPT promotes healthcare: current applications and potential challenges[J]. *Int J Surg*, 2024, 110(1): 606–608.
- [26] BERG H T, VAN BAKEL B, VAN DE WOUW L, et al. ChatGPT and generating a differential diagnosis early in an emergency department presentation [J]. *Ann Emerg Med*, 2024, 83(1): 83–86.
- [27] CHENG S W, CHANG C W, CHANG W J, et al. The now and future of ChatGPT and GPT in psychiatry [J]. *Psychiatry Clin Neurosci*, 2023, 77(11): 592–596.
- [28] LUYKX J J, GERRITSE F, HABETS P C, et al. The performance of ChatGPT in generating answers to clinical questions in psychiatry: a two-layer assessment [J]. *World Psychiatry*, 2023, 22(3): 479–480.
- [29] MARKS M, HAUPT C E. AI chatbots, health privacy, and challenges to HIPAA compliance [J]. *JAMA*, 2023, 330(4): 309–310.
- [30] WANG C Y, LIU S R, YANG H, et al. Ethical considerations of using ChatGPT in health care [J]. *J Med Internet Res*, 2023, 25: e48009.
- [31] MINSSEN T, VAYENA E, COHEN I G. The challenges for regulating medical use of ChatGPT and other large language models[J]. *JAMA*, 2023, 330(4): 315–316.
- [32] MENZ B D, KUDERER N M, BACCHI S, et al. Current safeguards, risk mitigation, and transparency measures of large language models against the generation of health disinformation: repeated cross sectional analysis [J]. *BMJ*, 2024, 384: e078538.

引用本文

韦 球, 蒋维芃, 杨超勉, 等. 医学数字人 GPT 的研究现状及展望[J]. 元宇宙医学, 2024, 1(1): 43–51.

WEI Q, JIANG W P, YANG C M, et al. Current status and prospects of digital human GPT in medicine [J]. *Metaverse Med*, 2024, 1(1): 43–51.