

人工智能与中医药前沿研究专栏

DOI: 10.16305/j.1007-1334.2025.z20241006001

ChatGPT 在中医药科研工作中的应用探讨

赵 涛¹, 韦昌法^{1,2}, 刘惠娜¹, 侯梦晗¹, 练 拓³

1. 湖南中医药大学(湖南 长沙 410208); 2. 湖南省智慧中医工程技术研究中心(湖南 长沙 410208);

3. 墨尔本皇家理工大学(澳大利亚 墨尔本 3000)

【摘要】 目的 探讨 ChatGPT 在中医药科研工作中的应用价值和模式, 并开展应用示范。方法 阐述 ChatGPT 及相关模型和应用的发展动态, 从通用工具调用、交互式对话、应用程序编程接口(API)集成与代码生成四个模式来分析 ChatGPT 在中医药科研工作中的应用模式, 并基于这些模式从医案结构化、命名实体识别、科研数据获取与分析、中医考研真题测评等方面开展了 ChatGPT 在中医药科研工作中的应用实践。结果 ChatGPT 在中医药科研工作中具有显著的应用成效, 其能够很好地完成医案结构化、命名实体识别、科研数据获取与分析等工作。但是, ChatGPT 解答《中医综合》考研真题等中医专业问题的能力还有待提升。结论 ChatGPT 在中医药领域展现出了巨大的应用潜力, 为科研工作者提供了高效、便捷的解决方案。未来, 随着技术的不断进步, ChatGPT 等大语言模型有望在中医药科研工作中发挥更加重要的作用。

【关键词】 人工智能; 大语言模型; 中医; 中药; 科研; 应用模式

Exploration on application of ChatGPT in scientific research work of traditional Chinese medicine

ZHAO Tao¹, WEI Changfa^{1,2}, LIU Huina¹, HOU Menghan¹, LIAN Tuo³

1. Hunan University of Chinese Medicine, Changsha, Hunan 410208, China; 2. Hunan Provincial Smart Traditional Chinese Medicine Engineering Technology Research Center, Changsha, Hunan 410208, China; 3. Royal Melbourne Institute of Technology (RMIT University), Melbourne 3000, Australia

Abstract: Objective To explore the application value and models of ChatGPT in scientific research of traditional Chinese medicine (TCM), and to demonstrate its practical applications. **Methods** This paper outlines the development trends of ChatGPT and related models, analyzing its application in TCM scientific research through four modes: general tool usage, interactive dialogue, API integration, and code generation. Based on these modes, practical applications of ChatGPT in TCM research were implemented in areas such as medical case structuring, named entity recognition, scientific data acquisition and analysis, and the assessment of TCM postgraduate exam questions. **Results** ChatGPT has shown significant effects in TCM research, excelling in tasks such as medical case structuring, named entity recognition, and scientific data acquisition and analysis. However, its ability to answer professional TCM questions, such as those in the past papers of the National Graduate Entrance Exam for *Comprehensive Traditional Chinese Medicine*, requires further improvement. **Conclusions** ChatGPT demonstrates enormous potential in the field of TCM, offering efficient and convenient solutions for researchers. With continuous advancements in technology, large language models like ChatGPT are expected to play an increasingly important role in TCM scientific research in the future.

Keywords: artificial intelligence; large language models; traditional Chinese medicine; Chinese materia medica; scientific research; application models

[基金项目] 湖南省自然科学基金项目(2025JJ90089); 湖南省卫生健康委健康科研课题(20254646, 20254730); 湖南省研究生科研创新项目(CX20240735); 湖南省研究生科研创新立项不资助项目(LXBZZ2024175); 湖南中医药大学研究生创新课题(2024CX084)

[作者简介] 赵涛, 男, 硕士研究生, 主要从事医学人工智能研究工作

[通信作者] 韦昌法, 教授, 硕士研究生导师; E-mail: weichangfa@hnuucm.edu.cn。刘惠娜, 讲师; E-mail: huinaliu@hnuucm.edu.cn

ChatGPT 是一种基于自然语言处理和深度学习的对话生成模型, 其能够根据用户的问题和对话语境, 给出符合语言逻辑和对话习惯的回答^[1]。自 2022 年 11 月由美国 OpenAI 公司发布以来, ChatGPT 迅速吸引了大量用户, 一度成为有史以来用户增长速度最快的应用^[2]。当前, ChatGPT 已经被广泛应用于许多研究领域, 其强大的语言理解和生成能力使其在中医药科研工作中也展现出了巨大的潜力和价值, 值得深入探讨。

以 ChatGPT 为代表的大语言模型的发展掀起了新一轮人工智能研究的热潮, ChatGPT 系列产品的发布和更新也促进了更多医药领域的大语言模型和应用的发展。许多研究团队应用 LLaMA、ChatGLM 等开源大模型研发了医药领域大模型, 典型的代表有医疗聊天模型 ChatDoctor^[3]、医疗健康大模型华佗 GPT^[4]、中文医疗大模型 DoctorGLM^[5]、本草 (BenTsao)^[6]、扁鹊 (BianQue)^[7]、明医 (MING)^[8] 等。上述大语言模型的出现, 能够促进这类技术在中医药领域发挥更大的作用。此外, 基于 ChatGPT 等大语言模型的应用软件正在呈井喷式发展, 这也为中医药科研工作的开展提供了极大的便利。

1 ChatGPT 在科研工作中的应用模式

近年来, ChatGPT 等大语言模型的应用模式逐渐多样化。有研究^[9-10]指出, ChatGPT 在科研工作中的应用至少包括 4 种模式: 第 1 种是使用基于 ChatGPT 的通用科研工具来辅助科研工作, 第 2 种是使用 ChatGPT 对话框来辅助科研工作, 第 3 种是调用 ChatGPT 的应用程序编程接口 (API) 来辅助科研工作, 第 4 种是指挥

ChatGPT 编写符合要求的程序来辅助科研工作、以实现特定的科研目标。这些不同的应用模式为科研工作提供了多样化的工具和方法, 提高了科研工作的效率。

1.1 使用基于 ChatGPT 的通用科研工具来辅助科研工作 如表 1 所示, ChatGPT 在科研中的应用广泛, 衍生了多款通用科研工具。其中, 国外研究者^[11]开发的 ChatGPT 应用 ChatPDF 能够轻松读取用户上传的 PDF 文献, 通过自然语言指令进行内容总结, 显著提高科研人员的文献阅读速度; 类似的工具还有 Chat2Doc、ChatExcel 和 ChatWeb 等。此外, GPT Academic 可以完成中英文学术润色、查找语法错误、进行中英文互译等工作, ChatPaper 可以帮助科研人员进行论文初筛, AutoResearcher 可以自动生成论文综述, ChatGenTitle 可以生成论文标题, ChatReviewer 可以进行论文预审稿, 它们都是学术论文写作的得力助手。对于无法访问 ChatGPT 的用户可以使用包括 ChatPDFs、Kimi 等在内的一些基于国产大语言模型开发的应用来完成相应工作。

表 1 基于 ChatGPT 的通用科研工具

工具	主要功能	网址
ChatPDFs	快速总结 PDF 内容, 并且可以再用指令与之对话	https://www.chatpdfs.cn
Chat2Doc	阅读、分析 PDF/Word 文档并回答用户的问题	https://chat2doc.cn
ChatExcel	可直接使用自然语言对表格中的数据信息进行查询、修改等操作	https://www.chatexcel.com
ChatWeb	可爬取任意网页并提取正文, 生成概要, 回答用户的问题	https://github.com/SkywalkerDarren/chatWeb
GPT Academic	学术优化: 中英文学术润色、查找语法错误、中英文互译等	https://github.com/binary-husky/gpt_academic
ChatPaper	通过 ChatGPT 实现对论文的总结, 帮助科研人员进行论文初筛	https://chatpaper.org
AutoResearcher	论文综述生成	https://github.com/eimenhmdt/autoresearcher
ChatGenTitle	论文标题生成	https://github.com/WangRongsheng/ChatGenTitle
ChatReviewer	论文预审稿	https://huggingface.co/spaces/ShiwenNi/ChatReviewer

1.2 使用 ChatGPT 对话框来辅助科研工作 ChatGPT 作为一种生成式预训练聊天机器人, 可以让用户通过对话框与其进行交互。如图 1 所示, 用户只需在对话框中提出自己的问题, ChatGPT 便会将用户输入的文本传入模型中进行处理, 利用先前的训练经验和上下文理解用户的意图并据此生成回复, 最后将回复结果输出到对话框中。值得注意的是, 给予 ChatGPT 特定的角色提示并编写高质量的提示词 (prompt) 可以大大提高 ChatGPT 回复内容的质量。

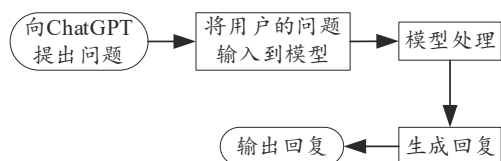


图 1 使用 ChatGPT 对话框辅助科研工作的流程

1.3 调用 ChatGPT 的 API 辅助科研工作 对于批量的问答和数据处理工作, 使用对话框进行逐次问答会比较

为繁琐而低效, 而调用 ChatGPT 的 API 则可以高效完成批量任务。如图 2 所示, 用户只需进行简单的操作即可在自己的程序中调用 ChatGPT 的 API, 以完成批量任务。

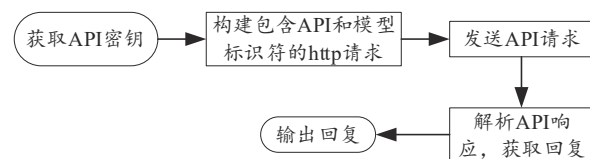


图 2 调用 ChatGPT 的应用程序编程接口 (API) 辅助科研工作的流程

1.4 指挥 ChatGPT 编写符合要求的程序来辅助科研工作 通常情况下, 较为复杂的科研工作任务往往无法通过使用 ChatGPT 对话框或调用 ChatGPT 的 API 来完成, 此时可以按照图 3 所示的流程指挥 ChatGPT 编写符合要求的程序来开展工作。通过与 ChatGPT 进行对话, 告知它程序需要实现的功能、需要处理的数据的结

构特点、需要采用的编程语言等信息,逐步指导它生成程序代码。需要注意的是,有时候 ChatGPT 生成的程序并不符合要求、需要进一步修改,此时可以将更多的需求细节或程序运行后产生的错误信息提供给 ChatGPT,要求它对程序进行修改和完善、直到符合用户的要求为止。

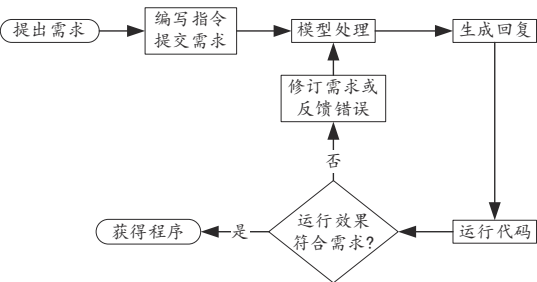


图3 指挥 ChatGPT 编写符合要求的程序来辅助科研工作的流程

2 ChatGPT 在中医药科研工作中的应用实践

本文将从医案结构化、命名实体识别、科研数据获取与分析、知识图谱构建以及中医考研真题测评等方面来介绍 ChatGPT 在中医药科研工作中的应用实践。鉴于 ChatGPT 的访问限制及法律考量,涉及 ChatGPT 的

实验环节均由具备访问权限的本团队海外合作者协助完成。

2.1 利用 ChatGPT 开展中医医案结构化工作 中医医案是中医临床经验的总结,是开展中医药研究的重要资料。中医医案结构化旨在对非结构化的医案文本进行处理,分别提取出患者基本信息、病情信息、诊断信息、治疗方案等各项具体内容^[12]。机器学习和深度学习方法曾被用于开展医案结构化,但是其中涉及的文本标注工作非常耗时,且需要进行大量的高维稀疏矩阵运算^[13]。ChatGPT 凭借其强大的自然语言理解能力,可以应用于医案结构化工作,并显著缩短时间和降低运算成本。

2.1.1 在 ChatGPT 对话框中进行医案文本结构化 图 4 展示了在 ChatGPT 对话框中进行医案文本结构化的示例,其中最重要的工作是根据用户需求编写提示词。为了使 ChatGPT 能够更精准地完成任务,在提示词中赋予它“中医药信息学专家”的角色,引导它基于角色完成任务。为了使 ChatGPT 能够更好地理解用户的输入内容并给出规范的输出,应在提示词中用符号指定待处理的医案文本及要求的输出格式。



你现在的身份是一名中医药信息学专家,请对后面的[]中的医案进行结构化处理并按照{}中的结构进行输出,只输出结构化处理后的结果,不要输出多余的信息。

{ (1) 患者姓氏: (2) 患者性别: (3) 患者年龄: (4) 病情信息: (5) 中医诊断: (6) 西医诊断: (7) 治疗原则(治则): (8) 治疗方法(治法): (9) 完整的处方信息: (10) 基本方: (11) 中药信息: (12) 剂量和用法: (13) 其他治疗措施: (14) 其他诊次信息: (15) 按语: } [张先生, 年龄45岁, 近期出现咳嗽、胸闷、气促等症状已持续两周, 伴有低热和乏力感。根据中医诊断, 患者被诊断为肺热咳嗽, 而西医诊断为急性支气管炎。治疗原则是清热化痰、润肺止咳, 并采用中药治疗的方式进行。完整的处方包括麻黄10g, 杏仁10g, 桔梗10g, 紫苏叶10g, 川贝母5g。基于此, 开具了麻杏石甘汤作为基本方。建议患者每次服用一小碗煎汤, 每天三次。此外, 还建议患者休息、保持室内空气流通, 并增加水的摄入量。三天后, 患者再次就诊时症状明显好转。综上所述, 根据清热化痰的治疗原则, 给予麻杏石甘汤治疗, 患者的病情有所缓解, 建议继续观察病情的变化。]



ChatGPT-
o4-mini

(1) 患者姓氏: 张; (2) 患者性别: 男性; (3) 患者年龄: 45岁; (4) 病情信息: 咳嗽、胸闷、气促, 持续两周, 伴低热和乏力感; (5) 中医诊断: 肺热咳嗽; (6) 西医诊断: 急性支气管炎; (7) 治疗原则(治则): 清热化痰, 润肺止咳; (8) 治疗方法(治法): 中药治疗; (9) 完整的处方信息: 麻黄10g, 杏仁10g, 桔梗10g, 紫苏叶10g, 川贝母5g; (10) 基本方: 麻杏石甘汤; (11) 中药信息: 每次服用一小碗煎汤, 每天三次; (12) 剂量和用法: 不详; (13) 其他治疗措施: 休息、保持室内空气流通, 并增加水的摄入量; (14) 其他诊次信息: 三天后再次就诊, 症状明显好转; (15) 按语: 根据清热化痰的治疗原则, 给予麻杏石甘汤治疗, 患者的病情有所缓解, 建议继续观察病情的变化。

图4 在 ChatGPT 对话框中进行医案文本结构化的示例

2.1.2 调用 ChatGPT 的 API 对单个医案文本进行结构化 医案文本结构化任务也可以通过在程序中调用

ChatGPT 的 API 来完成,其工作流程如图 2 所示,只需将图 4 所示的在 ChatGPT 对话框中输入的提示词指令以

字符串形式赋值给程序中的提示词变量即可。

2.1.3 调用 ChatGPT 的 API 批量进行医案文本结构化 为了完成医案文本结构化批量任务,可以指挥 ChatGPT 编写 Python 程序来完成工作,该程序首先读取一则医案文本,根据医案文本修改图 4 所示的提示词后调用 ChatGPT 的 API 来进行结构化处理,接收到 ChatGPT 返回的处理结果后存入 Excel 文件中,然后接着处理下一则医案,依此类推直到所有医案处理完毕。

无论是在 ChatGPT 对话框中提交提示词,还是调用 API 进行医案文本结构化处理,均可实现预期效果,这大大提高了中医医案文本结构化工作的效率。

2.2 利用 ChatGPT 开展中医药命名实体识别工作 命名实体识别是信息提取和信息检索中的关键任务,旨在从自然语言描述中识别、分类和标注命名实体^[14]。

2.2.1 在 ChatGPT 对话框中进行中医药命名实体识别 与中医医案结构化任务相似,可以在 ChatGPT 对话框中编写提示词来要求它执行命名实体识别任务,如图 5 所示。

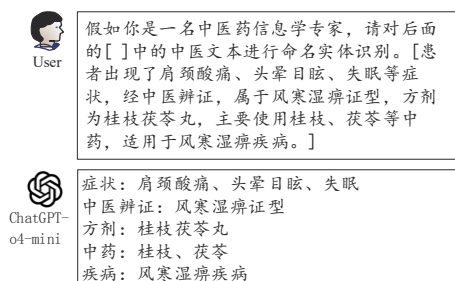


图5 在 ChatGPT 对话框中进行中医药命名实体识别的示例

2.2.2 调用 ChatGPT 的 API 批量进行中医药命名实体识别 当需要对大量中医文本开展命名实体识别任务时,可以通过程序循环调用 ChatGPT 的 API 来实现批量处理。本团队从中文医疗信息处理挑战榜 CBLUE (<https://tianchi.aliyun.com/dataset/95414>) 中获取中文医学命名实体识别数据集 CMeEE-V2,选取训练数据的前 100 条数据作为测试对象。使用 Python 程序读取 JSON 格式的 CMeEE-V2 数据,通过循环逐一处理每条数据,调用 ChatGPT 的 API 进行命名实体识别,并将结果结构化存储至新的 JSON 文件。为了对识别结果进行评估,以 Micro-F1 为主要评估指标,要求确保预测准确、实体的起始和结束索引以及类型精确匹配,CMeEE-V2 的训练数据提供了实体识别参考,可以直接作为准确度的参考。评估结果显示,针对上述 100 条测试数据,准确识别了 82 条数据,准确率为 82%。

上述实验结果显示了利用 ChatGPT 开展中医药命名实体识别的可行性以及调用 API 进行批量处理的高效性。其中 82% 的准确率表明 ChatGPT 具备较强的语

言理解能力,尤其在处理结构规范、语义清晰的医学文本时表现较为优异,但 18% 的错误识别集中于“肝风内动”与“肝阳上亢”等含混或多义词语,表明模型在中医语义区分能力上尚有不足。

2.3 利用 ChatGPT 获取中医药科研测试数据 由于 ChatGPT 是基于大量数据集训练的大语言模型,拥有丰富的内建数据库,因此可以利用 ChatGPT 获取中医药科研测试数据。

2.3.1 在 ChatGPT 对话框中获取中医药科研测试数据 如图 6 所示,可以在提示词中限定所需中医药科研测试数据的来源以及数据特点等信息,通过 ChatGPT 对话框发起请求,ChatGPT 便可依据自身的知识库返回符合需求的数据。

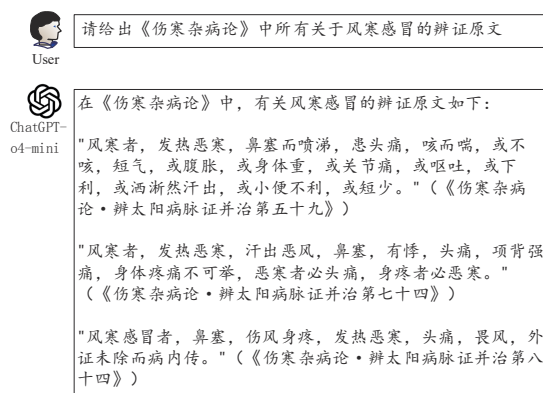


图6 在 ChatGPT 对话框中获取中医药科研测试数据的示例

2.3.2 调用 ChatGPT 的 API 批量获取中医药科研测试数据 当科研测试数据的需求量较大时,在 ChatGPT 对话框中获取数据会比较费时且繁琐,此时可以调用 ChatGPT 的 API 批量获取数据。本团队指挥 ChatGPT 编写了 Python 程序,循环调用 ChatGPT 的 API 获取中医典籍《黄帝内经》《金匱要略》《伤寒论》《温病条辨》中关于感冒的原文,获得《黄帝内经》原文 42 条、《金匱要略》原文 48 条、《伤寒论》原文 62 条、《温病条辨》原文 48 条。然而,ChatGPT 作为基于互联网语料训练的通用大语言模型,其生成数据的真实性尚需校验和考证。因此,通过该方法获取的数据仅适用于程序测试。同时,为确保通过 ChatGPT 生成的测试数据确实来源于《黄帝内经》等中医经典文献,可采用字符串匹配的方法对数据的真实性进行验证。

上述实验结果显示了利用 ChatGPT 获取中医药科研测试数据的可行性,缩短了获取数据的时间并降低经济成本。

2.4 利用 ChatGPT 开展中医药数据分析工作 在中医药数据分析工作中,传统的人工操作软件进行分析的方式较为耗时耗力,利用 ChatGPT 可以更加简单高效

地开展工作。对于相对简单的数据分析任务,用户只需在 ChatGPT 对话框中上传数据文件并告知分析目标,ChatGPT 即可完成分析任务并返回分析结果。如果用户出于数据安全的考虑,不希望将数据文件提供给 ChatGPT,或者数据分析任务较为复杂时,则可以把要分析的数据特点和分析目标告知 ChatGPT,按照图 3 所示的流程指挥 ChatGPT 编写出符合要求的数据分析程序。

本团队从国家自然科学基金大数据知识管理服务门户(<https://kd.nsfc.cn/>)获取了全国中医药院校 2017 年至 2021 年结题的国家自然科学基金项目数据,整理成 Excel 文件,指挥 ChatGPT 编写数据分析程序,对项目的批准年度、依托单位、项目类别和关键词等信息进行统计分析。下面将以研究热点统计分析为例,介绍如何利用 ChatGPT 进行数据分析并绘制热点词云图。

首先,根据数据分析需求制定提示词。本示例的提示词为:“请编写 Python 程序完成下列任务:①打开文件‘中医药院校国家自然科学基金项目信息.xlsx’。②读取‘项目关键词’的那一列数据;该列的每个单元格都存放着多个关键词,每个关键词之间用分号分隔;请提取各个关键词并进行频次统计;提取关键词时不要进行分词,直接根据分号区分不同关键词即可。③根据关键词及其频次数据绘制一个椭圆形的研究热点词云图,不同的词条用不同的颜色表示,并确保中英文等符号都能够正常显示、不出现乱码。”

接着,按照图 3 所示的流程与 ChatGPT 进行交互,引导 ChatGPT 生成并修改完善程序代码,直到完全满足需求。最后,运行最终代码,获得如图 7 所示的研究热点词云图。



图7 研究热点词云图

上述实验结果显示利用 ChatGPT 开展中医药数据分析工作的可行性,用户无需自己编写数据分析程序,大大降低了数据分析工作的门槛,提高了分析效率。

2.5 ChatGPT 辅助构建中医药知识图谱 中医药知识图谱是一种以结构化方式组织和表示中医药领域知识的图形数据结构。它通过将中医药相关的概念、实体、

关系以及事件进行关联,形成一个系统化的知识网络,从而为中医药的研究、诊断、治疗及智能化应用提供支持。然而中医药古籍文献的语义表达体系与现代医学存在显著差异,其特有的术语系统和理论框架使知识图谱的构建更加困难^[15]。在处理中医药文献时,常面临专业术语复杂、语义关系多样等挑战,尤其是在涉及“证候-治法-方药”这类复杂关系时,传统的知识图谱构建方法显示出诸多不足。

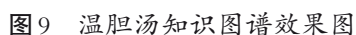
传统方法在处理中医“证候-治法-方药”关系时,往往依赖人工标注和规则匹配。人工标注不仅耗费大量人力和时间,而且由于中医知识的复杂性和主观性,不同标注者之间可能存在理解差异,导致标注结果的不一致性^[16]。规则匹配则需要预先定义大量复杂的规则,难以涵盖中医知识的多样性和灵活性,对于一些隐含的语义关系更是难以识别和抽取。例如,在面对复杂的中医证候描述时,传统方法很难准确地梳理出与之对应的治法和方药之间的逻辑关系。

ChatGPT 等大语言模型凭借其强大的自然语言处理能力,在抽取中医“证候-治法-方药”复杂关系方面展现出独特优势。它能够从海量中医药文献中自动抽取实体及其关系,来实现中医药知识图谱的构建^[17]。比如,ChatGPT 可以快速分析一篇描述多种证候及对应治法、方药的文献,准确识别出不同证候与相应治法、方药之间的关系,这种能力是传统方法难以企及的。相比传统方法,ChatGPT 等大语言模型能够更好地理解中医药领域的专业术语和复杂语义关系,从而提高知识抽取的准确性和效率^[18]。

2.5.1 使用 ChatGPT 辅助手动构建简单中医药知识图谱 对于数据量较少的简单知识图谱的构建,可以通过在 ChatGPT 的对话框中提交知识图谱构建的请求,然后根据 ChatGPT 返回的 Cypher 代码手动构建知识图谱。如图 8 所示,在提示词中给出构建知识图谱所用的温胆汤的原始描述文本以及构建知识图谱的要求,并要求 ChatGPT 返回完整的 Cypher 代码。ChatGPT 会分析温胆汤的描述文本,抽取出其中所有的实体及语义关系并给出构建知识图谱的详细步骤及每一步的 Cypher 代码。依据 ChatGPT 给出的步骤,将每一步的 Cypher 代码输入到 Neo4j Browser 或 Neo4j Desktop 里执行后就可以完成温胆汤知识图谱的构建,随后执行 ChatGPT 给出的最终可视化查询代码,就可以得到如图 9 所示的知识图谱效果图。

2.5.2 使用 ChatGPT 相关工具自动化构建中医药知识图谱 在构建文本量较大的中医药知识图谱时,若仅使用手动方式进行构建则显得效率过于低下。对此,如表 2 所示,为了更好地利用 ChatGPT 等大语言模型工具快速构建知识图谱,提高科研效率,研究者们开发出

图8 在 ChatGPT对话框中获取温胆汤的知识图谱构建完整代码



了包括 llmgraph、llm-graph-builder 等在内的知识图谱自动化构建工具。这些工具可以在 ChatGPT 等大语言模型的协助下将非结构化数据转换为存储在 Neo4j 中的知识图谱,从而实现知识图谱的自动化快速构建。

表 2 知识图谱自动化构建工具

工具	网址
llmgraph	https://github.com/dylanogg/llmgraph
SAC-KG	https://github.com/stair-lab/kg-gen
llm-graph-builder	https://github.com/neo4j-labs/llm-graph-builder
kg-gen	https://github.com/stair-lab/kg-gen

综上所述,ChatGPT 等大语言模型凭借其强大的自然语言处理能力,可高效解析海量中医药类非结构化文本,从而实现相关实体的精准识别与提取,并辨析出这些实体间的复杂关系,使得知识图谱的构建更为高效且精确、全面。与此同时,知识图谱也可以为大语言模型提供结构化的知识支撑,极大地提高了其回复的准确性与可靠性。知识图谱凭借自身丰富的语义关系和逻辑结构,为大语言模型注入具有上下文深度的语义层,使其能够以更细致入微的理解来处理语言。例如,在中医医疗问答系统里,知识图谱能够协助 ChatGPT 理解疾病、证候与症状之间的复杂关系,从而生成更准确的诊断建议。

2.6 ChatGPT 解答《中医综合》考研真题的能力测评

《中医综合》是为高等院校和科研院所招收中医学硕士研究生而设置的全国选拔性考试科目,其全部试题均为选择题,具有客观性。试题内容由教育专家精心设计,能够全面反映中医学核心知识体系。同时,该考试的数据易于获取,操作性强,具备较高的研究价值。因此,本团队选取《中医综合》考研真题作为评估 ChatGPT 解答中医专业问题能力的依据。

2.6.1 在 ChatGPT 对话框中进行测评 如图 10 所示,为了提升 ChatGPT 解答中医专业问题的准确率,在提示词中赋予它“中医专家”的角色,引导它基于角色解答问题。由于《中医综合》考研真题有多种题型,故在提示词中说明了题型信息及对应的作答要求,以帮助 ChatGPT 给出正确答案。此外,在提示词中通过特殊符号限定了题目信息的范围和输出格式,以便于 ChatGPT 准确理解题目并根据要求输出答案。

2.6.2 调用 ChatGPT 的 API 进行批量测评 为了全面评估 ChatGPT 解答中医专业问题的能力,本团队将 2021 年和 2022 年《中医综合》考研真题整理到 Excel 文件中,通过指挥 ChatGPT 编写 Python 程序、循环调用 ChatGPT 的 API 进行批量测评。程序依次从 Excel 文件中读取各个题目并提交给 ChatGPT 解答,将 ChatGPT 的答案与正确答案进行比对、确定每个题目的得分,并将

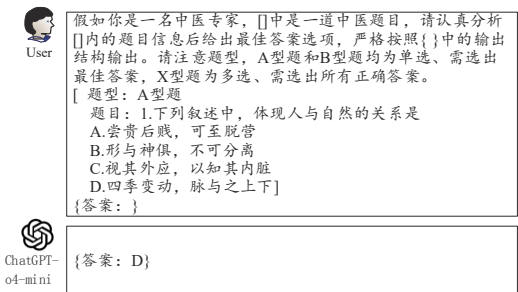


图 10 利用 ChatGPT 解答《中医综合》考研真题的示例
答案与得分存入 Excel 文件,最后对所有题目的作答情况进行统计分析。在 330 道《中医综合》考研真题中,ChatGPT 总答对 154 道,准确率为 46.67%。

上述测评结果中 46.67% 的准确率,暴露了 ChatGPT 在中医学复杂推理、概念体系及辨证思维等方面的不足。造成其准确率偏低的原因主要有以下三个方面:第一,ChatGPT 的训练数据以现代中文和英文等语料为主,对中医古文及专业教材覆盖不足;第二,模型主要依赖统计语言模式,缺乏“六淫致病”“气血津液”等中医理论的结构化知识支撑;第三,中医考研题型形式多样,包括单选、多选等类型,答题过程不仅依赖知识记忆,还需要综合推理与语境理解能力。针对上述问题,可采取以下改进策略:一方面,通过加载中医药相关的本地知识库,并结合 ChatGPT 的 API 调用,实现知识增强型问答系统,从而提升其在中医药领域的知识准确性;另一方面,也可利用中医古文和专业教材等高质量语料,对开源大模型进行有针对性的微调,以增强其对中医理论与语言体系的理解能力,从而弥补现有模型在该领域的不足。

3 ChatGPT 的局限性与优化方案

虽然 ChatGPT 凭借其强大的自然语言处理能力,可以大幅度提高中医药科研的效率,降低数据分析工作的门槛,但它仍存在一定的局限性,主要集中在数据滞后问题与“幻觉现象”这两个方面。

首先,ChatGPT 本身的知识来源于模型的训练数据,而由于模型的更新周期较长,其回答内容可能落后于最新的科研成果或社会动态。尤其是医学这种快速发展的领域,数据滞后问题尤为突出。为解决该问题,中医药科研工作者在使用 ChatGPT 等大语言模型时,可采用如下优化方案:一方面,整合联网搜索功能,让 ChatGPT 能够即时获取最新的科研动态和知识信息,进而提供更有时效性的回答;另一方面,使用检索增强生成(RAG)技术,RAG 框架可引入外部知识检索模块,在模型生成回答的过程中,实时从动态的知识源(如文档数据库、网页等)中检索相关信息,并将这些信息融入生成过程中,使模型能够生成更准确、更具针对性的回复。

其次,ChatGPT 在生成文本时可能产生“幻觉现

象”,也就是生成错误或者虚构的信息,在医学这种专业性较强的领域这种情况更为常见。比如在诊断、治疗建议和药物信息方面,ChatGPT 可能会提供不准确甚至有害的信息^[19]。这种现象可能会误导用户,降低模型的可靠性。针对此问题,可通过以下方式进行优化:其一,结合知识图谱加以验证,运用结构化的知识体系对生成内容进行事实核查,从而减少错误信息的产生;其二,强化提示词工程,通过设计更为精确的提示词来引导模型生成更符合事实与逻辑的回答,从而提高输出的准确性与可靠性。

此外,由于中医经典著作(如《伤寒论》)大多采用文言文书写,包含大量通假字和省略语法^[20],致使 ChatGPT 等大语言模型工具在中医药领域的应用中面临着独特的挑战。研究^[21-22]指出,ChatGPT 对现代医学问题的回答一致性较高,但对“肾阳虚”“营卫不和”等传统中医术语的解释常出现概念混淆等问题。

与此同时,由于数据脱敏不充分或者模型记忆效应等原因,在 ChatGPT 处理临床敏感信息(例如电子病历、影像诊断记录)时,可能会出现患者隐私泄露等问题,这一风险已经在医学教育界引起了广泛关注。调查^[23]显示,22% 的医学生因为其隐私隐患而明确反对将该技术应用于教学场景。另外,ChatGPT 所生成的内容可能包含未经授权的版权材料,存在一定的学术侵权风险^[24]。对于上述问题,在中医药科研活动中应用 ChatGPT 等大语言模型时可采取如下措施:首先,按照《人工智能伦理治理标准化指南》^[25],对输入给大语言模型的数据进行双重匿名化处理;其次,构建版权合规审查机制,通过与学术数据库进行比对来验证生成内容的原创性。

除了 ChatGPT 之外,逐渐涌现出的国产大语言模型也在逐步展现其在医学和中医药科研中的潜力。以 DeepSeek 为例,该模型支持本地部署使用,更符合科研工作中的数据隐私要求^[26]。有研究^[27]表明,在诊断推理任务中,DeepSeek-R1 的步骤准确性显著高于 ChatGPT4.0 和 LLaMA 3.1-405B(平均 Likert 评分分别为 3.61 和 3.22),但是在影像报告总结的全局质量上,DeepSeek 的表现略逊于 ChatGPT。因此,在实际应用中,应根据任务特点决定模型的选择,或尝试多模型集成策略,以达到更优的科研效果。

综上所述,综合运用联网搜索、本地知识库、知识图谱验证以及强化提示词工程等方法,能够有效减轻 ChatGPT 的数据滞后问题和幻觉现象,进一步提高其在实际应用中的性能和可靠性。

4 结语

ChatGPT 凭借其强大的语言理解和生成能力在中医药科研工作中展现出了巨大的潜力和价值,在中医

医案结构化、命名实体识别、科研数据获取与分析等方面都有不错的表现。可以预见,随着技术的不断进步,ChatGPT 等大语言模型将会在中医药科研工作中发挥更加重要的作用,广大科研工作者可以进一步深入探索 ChatGPT 等大语言模型在中医药科研工作中的应用模式和应用场景,以推进中医药事业的发展。对于无法访问 ChatGPT 的用户,也可以选择 DeepSeek、文心一言、通义千问、ChatGLM 等国产大语言模型来完成相应工作。由于这些模型同样属于通用大语言模型,因此本文所探讨的应用模式和使用方法同样适用于国产大模型,为用户提供了多样化的选择。

参考文献:

- [1] 卢经纬,郭超,戴星原,等. 问答 ChatGPT 之后:超大预训练模型的机遇和挑战[J]. 自动化学报, 2023, 49(4): 705-717.
- [2] 刘禹良,李鸿亮,白翔,等. 浅析 ChatGPT:历史沿革、应用现状及前景展望[J]. 中国图象图形学报, 2023, 28(4): 893-902.
- [3] LI Y X, LI Z H, ZHANG K, et al. ChatDoctor: a medical chat model fine-tuned on a large language model meta-AI (LLaMA) using medical domain knowledge[J]. Cureus, 2023, 15(6): e40895.
- [4] ZHANG H B, CHEN J Y, JIANG F, et al. HuatuoGPT, towards taming language model to be a doctor [J/OL]. arXiv, 2023 [2024-09-01]. <https://arxiv.org/pdf/2305.15075>.
- [5] XIONG H L, WANG S, ZHUY T, et al. DoctorGLM: fine-tuning your Chinese doctor is not a herculean task [J/OL]. arXiv, 2023 [2024-09-01]. <https://arxiv.org/abs/2304.01097>.
- [6] WANG H C, LIU C, XI N W, et al. HuaTuo: tuning LLaMA model with Chinese medical knowledge [J/OL]. arXiv, 2023 [2024-09-01]. <https://arxiv.org/abs/2304.06975>.
- [7] CHEN Y R, WANG Z Y, XING X F, et al. BianQue: balancing the questioning and suggestion ability of health LLMs with multi-turn health conversations polished by ChatGPT [J/OL]. arXiv, 2023 [2024-09-01]. <https://arxiv.org/abs/2310.15896>.
- [8] LIAO Y S, JIANG S Y, WANG Y, et al. MING-MOE: enhancing medical multi-task learning in large language models with sparse mixture of low-rank adapter experts [J/OL]. arXiv, 2024 [2024-09-01]. <https://arxiv.org/abs/2404.09027>.
- [9] LIU Y, LE-CONG T, WIDYASARI R, et al. Refining ChatGPT-generated code: characterizing and mitigating code quality issues [J/OL]. arXiv, 2023 [2024-09-01]. <https://arxiv.org/abs/2307.12596>.
- [10] LEE P Y, SALIM H, ABDULLAH A, et al. Use of ChatGPT in medical research and scientific writing [J]. Malays Fam Physician, 2023, 18: 58.
- [11] 桑基韬,于剑. 从 ChatGPT 看 AI 未来趋势和挑战[J]. 计算机研究与发展, 2023, 60(6): 1191-1201.
- [12] 韦昌法,罗丽琴,晏峻峰,等. 中医数字辨证配套医案智能采集与分析系统构建研究[J]. 湖南中医药大学学报, 2020, 40(1): 70-74.
- [13] 何晗. 自然语言处理入门[M]. 北京:人民邮电出版社, 2020: 8.
- [14] 高翔,王石,朱俊武,等. 命名实体识别任务综述[J]. 计算机科学, 2023, 50(S1): 26-33.

(下转第 16 页)