

• 专 论 •

中医药大语言模型的关键技术与构建策略

萧文科¹, 宋 驰^{1,2}, 陈士林^{1,2}, 陈 伟^{1,2*}

1. 成都中医药大学 中医药创新研究院/交叉学科研究院, 四川 成都 611137

2. 成都中医药大学 本草基因组学研究院, 四川 成都 611137

摘要: 大语言模型 (large language model, LLM) 通过处理和理解自然语言数据, 实现高质量的信息检索、知识提取等功能, 为中医药研究提供了新机遇。基于中医药大模型发展现状, 梳理了 LLM 开发过程中的数据存储与处理方法, 概述了检索增强生成、混合专家模型、人类反馈强化学习、知识蒸馏等人工智能方法, 归纳了 LLM 训练微调与性能评价方法。针对中医药数据的特点, 从高质量数据集构建、多领域专家系统融合、信息快速提取、训练与调优等方面入手, 提出了中医药 LLM 的构建策略, 并分析了 LLM 在中医药领域的具体应用场景, 为中医药领域 LLM 的构建和应用提供参考, 推动中医药现代化和智能化发展。

关键词: 中医药; 大语言模型; 混合专家系统; 检索增强生成; 人类反馈强化学习; 知识蒸馏

中图分类号: R28; TP18 文献标志码: A 文章编号: 0253-2670(2024)17-5747-10

DOI: 10.7501/j.issn.0253-2670.2024.17.001

Key technologies and construction strategies of large language models for traditional Chinese medicine

XIAO Wenke¹, SONG Chi^{1,2}, CHEN Shilin^{1,2}, CHEN Wei^{1,2}

1. Innovative Institute of Chinese Medicine and Pharmacy/Academy for Interdiscipline, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, China

2. Institute of Herbgonomics, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, China

Abstract: By processing and understanding natural language data, large language models (LLM) enable the high-quality information retrieval, knowledge extraction, etc., and provide new opportunities for traditional Chinese medicine (TCM) research. Based on recent developments of LLM in TCM, the present work summarizes the data storage and processing algorithms, as well as artificial intelligence methods, such as retrieval-augmented generation, mixture of experts, reinforcement learning from human feedback, and knowledge distillation for developing LLM. It also summarizes methods for training fine-tuning and performance evaluation of LLM. In response to the characteristics of TCM data, strategies for developing LLM for TCM are proposed, which focuses on developing high-quality datasets, integrating mixture of experts, rapid information extraction, and model training and optimization. Additionally, it outlines specific application scenarios of LLM in TCM. The aim of this work is to provide insights for the development and application of LLM in TCM, promoting the modernization and intelligent development of TCM.

Key words: traditional Chinese medicine; large language models; mixture of experts; retrieval-augmented generation; reinforcement learning from human feedback; knowledge distillation

中药产业既是传统的优势产业, 也是国家重点扶持的战略性新兴产业。国家中医药管理局发布的《“十四五”中医药信息化发展规划》强调要在中医药领域开展云计算、大数据、物联网、人工智能、

收稿日期: 2024-06-14

基金项目: 成都中医药大学引进人才项目 (030041225)

作者简介: 萧文科, 助理实验师, 从事中医药信息学研究。E-mail: xiaowenke@cdutcm.edu.cn

*通信作者: 陈 伟, 博士生导师, 教授, 从事生物信息与计算生物学研究。E-mail: greatchen@ncst.edu.cn

5G、区块链、智能感知等新一代信息技术的集成应用研究，探索一批中医药数字化应用场景建设^[1]。当前，中医药产业转型升级迫在眉睫，“数字化”作为当下经济发展的重要引擎，为中医药发展提供了新思路。云计算和大数据技术为中医药数据的收集、存储、分析提供了良好的技术支持和基础架构。然而，海量的中医药数据并非易于直接解读，传统的人工分析方法不仅耗时耗力，而且容易受到主观因素的影响，难以挖掘其中深层次规律和潜在关联。如何利用人工智能和大数据分析技术，推动中医药的转型升级，成为了一个亟待解决的问题。

人工智能是一种模拟和扩展人类智能的技术，其通过计算机程序来模拟通常需要人类智能才能完成的任务，如感知、学习、推理和决策。人工智能已经成为推动社会发展、提升人类生活质量的重要力量，并已逐步融入中医药体系的多个环节，在中医四诊^[2-6]、健康管理、名老中医传承^[7]、配方创新等多方面都得到应用。近年来，以生成式人工智能（artificial intelligence generated content, AIGC）为代表的新一代人工智能技术，加速了人机互动进程。在这一领域，大语言模型（large language models, LLM）已成

为研究热点，其在多种自然语言处理任务上已取得显著成果^[8-9]，并表现出了良好的知识涌现能力。

中医药 LLM 是数字化、智能化革命的代表，将传统中医理论与现代科技相结合，有望为中医药的研究、临床应用和发展提供强大的支持。LLM 在中医药领域中的应用主要体现在通过先进的人工智能技术，对中医药的古籍文献、临床数据和药物成分进行深入分析和挖掘，从而辅助新药发现、优化中药配方、提高诊疗精度，促进中医药知识的传承与创新，为中医药的现代化和全球化提供强有力的技术支持。自 2022 年 11 月聊天生成预训练转换器（chat generative pre-trained transformer, ChatGPT）发布以来，中医药领域已有多个聚焦于中医药智能问诊、中医辅助诊断、中医养生咨询、中医药知识问答、中药产业链的中医药 LLM 问世（图 1）。这些中医药 LLM 在提升智能化诊断准确性、优化问诊流程、增强用户体验和促进中药产业链数字化等方面取得了显著进展。但由于数据资源匮乏、模型训练数据不足、核心算法适配性有限及实际应用场景复杂等原因，这些模型还存在泛化能力不足、数据融合难度大、模型可靠性有待提高等问题。

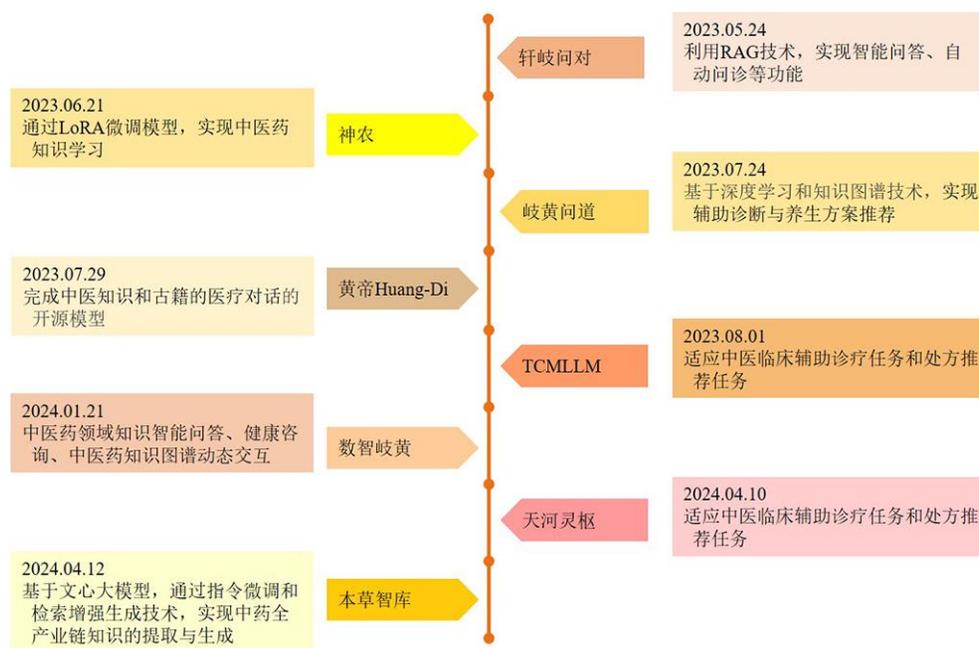


图 1 现有中医药 LLM 介绍

Fig. 1 Introduction of existing LLM for traditional Chinese medicine

中医药作为中国传统医学的重要组成部分，其独特的治疗理念和丰富的药材资源为世界医学做出了巨大贡献。然而，中医药智能化发展底层核心技术相对薄弱，制约了中医药现代化高速发展的进程。

为加强中医药科技创新体系建设，提升传承创新能力，中药现代化的发展急需融入各行业前沿技术。通过构建中医药领域的专有数据 LLM，促进中医药丰富的理论、经验和实践知识有机整合，有助于挖

掘中医药的潜在价值，提升诊疗的精准度和效率，促进个性化医疗服务的发展。

LLM 技术的融入为中医药的传统诊疗方法带来了创新，但在数据质量、模型准确性、应用场景等方面仍面临挑战。因此，本文旨在探讨开发中医药 LLM 的关键技术方法、中医药 LLM 构建策略及应用场景，为推动人工智能辅助中医药 LLM 的技术探索和科技创新提供参考。

1 中医药 LLM 的关键技术与方法

1.1 数据存储

中医药领域涉及大量的数据，在中医药 LLM 的构建过程中，数据存储技术具有重要作用，其不仅确保了数据的安全存储，还为模型的高效训练和深度分析提供了强有力的支持。为了有效支持 LLM 的训练需求，需要采用多种高效的数据存储和管理技术，如分布式文件系统（Hadoop 的 HDFS、Google

的 GFS，以其高吞吐量的数据访问能力，可为大规模数据集提供高效的存取解决方案）、对象存储系统（Amazon S3、Google Cloud Storage 具有极高的扩展性和全球访问能力，适用于大量静态数据的存储）。

1.2 数据处理

为构建功能强大的中医药大模型，研究者需从多元化的数据源中收集大量数据，然后通过数据清洗、词元化及数据调度等步骤进行数据处理，具体流程见图 2。数据清洗通过去除数据集中的噪声、不相关信息、敏感内容及重复条目，确保训练数据的纯净度和准确性，实现数据质量的严格控制。词元化过程将文本数据分解为更小的单元，提升模型对文本的理解和处理效率，增强模型捕捉语言细微差别的能力。数据调度环节至关重要，涉及高效合理地安排数据处理的时间和顺序，确保模型能够快速吸收和学习数据中的关键信息。



图 2 LLM 数据处理流程

Fig. 2 Data processing flow of LLM

1.3 检索增强生成 (retrieval-augmented generation, RAG)

RAG 作为一种结合检索和生成的技术，为提高模型在特定任务上的表现和效率提供了新视角^[10]。RAG 采用向量化技术将文档库中的文档转化为向量，在这些向量上构建高效索引系统^[11]，具体构建流程见图 3。RAG 技术通过有效融合检索和生成过程，能够在回答复杂问题时提供更深入、更具体的信息，提升了内容的相关性与准确性^[12]。

在中医药领域，RAG 技术通过结合现代信息检索系统和自然语言处理技术，可以高效的从海量中医药文献和临床数据中提取与特定病证或药物相关的信息，进而促进中医药病例分析、药物配伍及治疗效果评估等领域的发展^[13-15]。然而，RAG 对检索系统的依赖性较强，模型的表现很大程度上依赖于底层检索系统的效果，且 RAG 系统在训练和推理阶段计算成本较高，其在实际应用中的广泛性和实时性会受到影响^[16]。

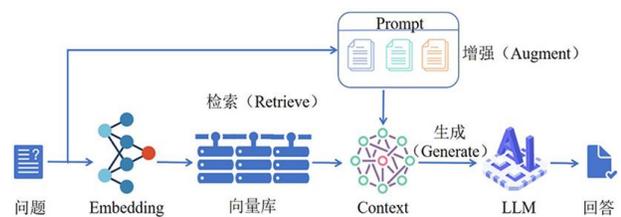


图 3 RAG 原理示意图

Fig. 3 Schematic diagram of RAG

1.4 人类反馈强化学习 (reinforcement learning from human feedback, RLHF)

RLHF 是一种结合了传统强化学习算法和人类反馈的技术，用于改进 LLM 的行为和决策过程^[17-18]。通过数据收集、奖励设计、模型训练、策略迭代，RLHF 可有效提升模型的性能，其工作原理见图 4。在中医药领域，名老中医的丰富经验与用户反馈被整合进 RLHF，进一步增强了 LLM 在智能问诊、知识问答、药方推荐及养生建议方面的准确性和实用性。这种方法不仅提升了中医药 LLM 的应用效果，

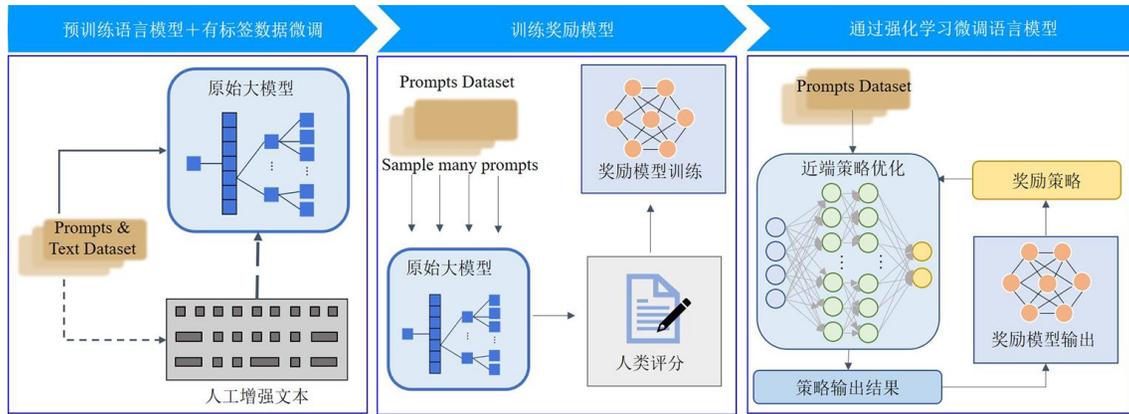


图 4 RLHF 原理图

Fig. 4 Schematic diagram of RLHF

也保证了其在复杂临床场景中的适应性和有效性。具体而言，RLHF 整合人类直觉和偏好优化语言模型，使模型能够适应复杂和不确定的任务，更好地反映人类的价值观和期望，从而提高模型的适应性。为了增强 LLM 与人类价值观的一致性，RLHF 通过利用人类反馈数据来指导模型微调，从而在多个标准上实现与人类的对齐。该过程首先需要收集人类对不同模型输出的偏好，然后使用这些反馈数据训练奖励模型，最后基于奖励模型运用强化学习算法对 LLM 进行微调。RLHF 支持细粒度的优化，允许模型在执行特定任务时，如文本生成和对话交互，更精确地调整行为以满足用户特定偏好^[19]。这种方法不仅提升了模型的实用性，而且为模型的定制化和迭代开发提供了极大的灵活性和可扩展性。

然而，RLHF 在使语言模型与人类意图保持一致的过程中也面临挑战。RLHF 依赖的数据及其标注的复杂性引入了多种主观因素，如标注者偏好和研究设计等，可能导致模型输出与实际用户需求之间存在偏差^[20]。因此，尽管 RLHF 在提高模型适应性方面具有显著优势，但仍需谨慎处理这些潜在的偏差问题，以确保模型的输出能够真正符合用户需求。

1.5 混合专家系统 (mixture of experts, MoE)

MoE 通过整合规则推理、神经网络、机器学习和模糊逻辑等技术，形成能够处理复杂和不确定问题的系统^[21]，具体原理见图 5。MoE 通常包括知识处理模块、推理机制模块和学习模块。每个模块专注于处理特定类型的任务，通过动态调度协同工作，以实现最优决策。MoE 的自我学习和适应性使其能够从新的数据和经验中不断学习和优化，从而提高决策的效率和质量^[22]。在中医药 LLM 中，MoE 通

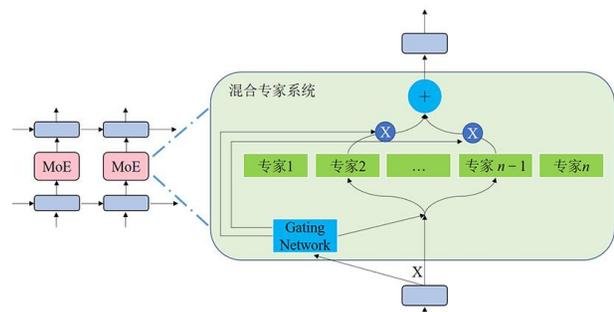


图 5 MoE 原理示意图

Fig. 5 Schematic diagram of MoE

过整合传统中医知识和现代医学研究结果，能够处理诊断和治疗中的不确定性问题，并吸纳新的临床研究结果和患者反馈，不断优化和更新治疗策略，使中医药 LLM 的应用更加精准、高效。具体而言，MoE 可以有效整合不同的决策支持技术来增强模型的理解和生成能力^[23]，在处理自然语言的模糊性和复杂性问题时提供更加精准和鲁性棒的决策支持^[24]，从而提高模型在复杂语境中的应用效率，增强模型对新情况的适应能力和学习速度^[25]。

1.6 知识蒸馏

知识蒸馏是一种模型压缩技术，其基本原理是将大型、复杂的教师模型学到的知识传递给小型、结构简单的学生模型^[26]，具体原理见图 6。在知识蒸馏过程中，通常使用教师模型的输出指导学生模型的训练，从而使学生模型能够模仿教师模型的行为。通过这种方式，学生模型可以从教师模型中蒸馏出丰富的知识，包括数据分布、模式和规律等，从而达到在学生模型上实现与教师模型相似甚至更好性能的效果。

在 LLM 开发领域，知识蒸馏主要用于减少模

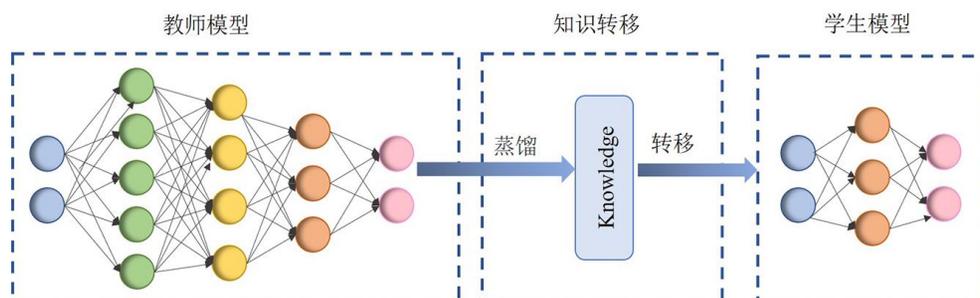


图6 知识蒸馏原理示意图

Fig. 6 Schematic diagram of knowledge distillation

型的规模、算力需求和能耗,同时尽量保持模型的性能^[27]。如通过从1个几十亿参数的LLM中蒸馏出只有几百万参数的模型,可以显著降低部署成本和提高运行速度,同时在任务性能上保持接近原始LLM的效果^[28]。

中医药系统包含大量的复杂知识和临床实践数据,传统的LLM往往需要庞大的计算资源来处理这些信息。通过知识蒸馏技术,可以有效地将复杂的中医药知识和临床经验从大型教师模型转移到更小、更高效的学生模型中,有助于在有限的计算资源下提供快速、准确的中医药诊断和治疗建议,促进中医药知识的更广泛传播和应用。此外,知识蒸馏还有助于解决中医药数据稀疏性和不均匀性问题,通过从LLM中提取和压缩关键信息,学生模型能够更好地处理和理解分散的中医药数据,从而提高诊断的准确性和治疗的个性化推荐。这种方法不仅提高了中医药LLM的实用性和访问性,还增强了其在实际临床环境中的应用效率和效果。

1.7 LLM训练微调方法

1.7.1 预训练 LLM的预训练策略主要通过广泛的语料库上进行无监督学习,使模型掌握语言的通用模式和结构。这一过程通常涉及预训练和微调2个阶段。在预训练阶段,LLM通过大量文本进行训练,以学习语言的基本语法、语义和上下文关系。掩码语言建模^[29]和下一句预测^[30]是2种在自然语言处理领域中广泛使用的预训练任务,能够帮助模型捕捉词之间的依赖关系和句子间的逻辑关系。

1.7.2 微调与对齐 大型语言模型的微调是针对特定任务调整预训练模型的过程,旨在通过细微调整模型的参数来适应具体应用,从而提高任务相关性^[31]。微调通常在较小的、特定任务的数据集上进行,使用与预训练阶段相同的或修改后的损失函数,加快模型的训练速度。

监督微调是LLM微调中一种关键的优化策略。监督微调在预训练阶段后,利用带标签的数据集进行的定向训练,通过反向传播和梯度下降方法调整模型权重,以减小预测输出与实际标签间的误差^[32]。监督微调充分利用了预训练模型的语言处理能力,提升模型在具体任务上的专业适应性和准确性。在实际应用中,监督微调能够实现LLM的高效性能提升,为深入研究LLM在各领域中的应用提供了重要的技术支持^[33-34]。

在LLM构建中,指令微调是常用的微调方法^[35]。指令微调需要首先收集或构建指令化的实例,然后通过有监督的方式对LLM的参数进行微调。经过指令微调后,LLM能够展现出较强的指令遵循能力,能够通过零样本学习方式解决多种下游任务^[36]。

低秩适配器(low-rank adaptation, LoRA)作为一种参数高效的微调技术,在LLM微调中受到广泛关注。LoRA通过在预训练模型的参数矩阵中引入低秩分解矩阵,实现对模型各层参数的近似更新,维持模型性能并且显著降低了针对特定下游任务的训练参数量^[37]。与传统的全参数微调方法相比,LoRA提供了一种成本效益较高的替代方案,被广泛用于LLM的高效微调^[38]。

人类对齐是在微调过程中确保学习的表示与特定任务需求保持一致的过程,旨在保证LLM的行为与人类期望和价值观一致。在LLM的预训练和监督微调的过程中,主要训练目标是根据上下文内容来预测下一个词元。但是,这一过程并未充分考虑人类的价值观或偏好,可能导致LLM从数据中学习不符合人类期望的生成模式^[39]。人类对齐确保了不同语言之间的语义和语法结构能够相互转换,保持表征的有效性和一致性。这些过程使得大型模型在各种应用中都能达到较高的精度和适应

性，同时保持了预训练阶段获得的丰富知识。

1.8 LLM 评估方法

在评估 LLM 的能力时，研究者需综合多个维度考察其性能，包括通过多种任务来测试模型的泛化能力，并且根据不同的应用场景，选择恰当的评价指标以确保性能的精确测量。表 1 展示了不同分类任务的常用评价指标，这些指标确保了评估过程的严谨性和模型性能评价的有效性。

2 中医药 LLM 的构建框架

中医药数据具有类型多样性与逻辑复杂性的特点，为实现中医药数据的有机融合，提高信息处理的效率和模型的智能性，本文从高质量数据集构建、多领域专家系统融合、信息快速提取、训练与调优等方面入手，结合 RAG、MoE、LoRA、监督微调技术，提出了构建中医药 LLM 的框架，见图 7。

表 1 LLM 评估指标

Table 1 Evaluation metrics for LLM

评价任务	评价指标	介绍
分类任务	精确率	计算模型预测为正例的样本中真正例的比例
	召回率	计算真正例的样本中被模型正确预测的比例
	F1 分数	综合衡量模型输出的精确率和召回率
语言建模任务	困惑度	衡量模型对参考文本的建模概率
文本生成任务	BLEU	评价机器翻译与参考翻译之间的重合度
	ROUGE	评价机器摘要与参考摘要的重合度
回答任务	准确率	衡量模型预测准确性的指标
执行类任务	成功率	衡量模型成功完成任务的比例
	Pass@k	表征模型生成的 k 个方案中至少包含 1 个正确答案的比率
偏好排序类任务	Elo 等级系统	衡量模型在候选者中的相对水平

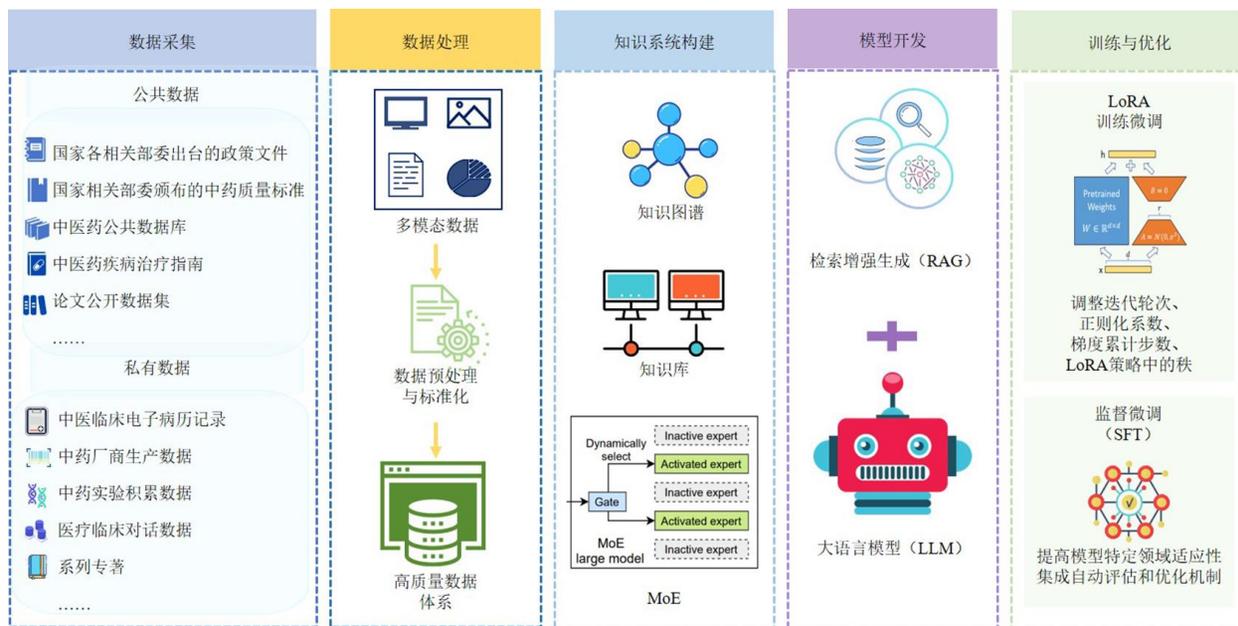


图 7 中医药 LLM 架构示意图

Fig. 7 Schematic diagram for architecture of traditional Chinese medicine LLM

2.1 中医药高质量数据集

高质量数据集的构建是实现 LLM 开发的基础。中医药数据集通常包含公开数据和私有数据 2 部分。在收集公开数据集时，必须对数据来源进行严格筛选，优先考虑经过同行评审的学术期刊、权威中医药机构发布的数据，及政府和相关机构的公开数据库。这类数据的可靠性高，有益于 LLM 从中提取正确的知识。对于私有数据，如医院的临床记录、个人健康记录等，需确保在遵守相关法规和国际规范的前提下进行收集和使用。此外，数据还需经过

严格的隐私保护措施，如采用高级数据脱敏技术和加密协议，确保数据在分析和模型训练过程中的安全性和隐私性。

数据整合和标准化是构建高质量数据集的关键。通过制定统一的数据模型和编码系统，将不同来源和格式的数据转化为统一的格式，可以促进数据的整合、存储和分析。如采用国际疾病分类系统和中医病症分类标准，可以标准化诊断信息，增强数据的国际化兼容性和扩展性。广泛采集多样化的数据，进行标准化和结构化处理，实施严格的质量

控制,同时确保数据的伦理性和隐私保护,可构建出一个高质量的数据集,为中医药 LLM 的开发提供坚实的基础。

2.2 中医药多领域专家系统构建

中医临床决策支持系统作为中医现代化发展的重要组成部分,正日益受到广大医务工作者的关注和重视^[40]。中医药涵盖内科、外科、妇科、儿科等多个领域,各个领域都有其独特的诊疗方法和理论体系。中医药多领域专家系统构建有助于中医临床决策支持系统的快速发展,利用 MoE 技术构建各领域的子模型,通过集成所有子模型能够保证各领域模型精度,应对不同领域的需求,提高模型的综合性能。

2.3 中医药文本信息快速提取

中医药文本浩如烟海,其中包含大量宝贵的知识,但如何有效地从中获取并利用这些知识是一个关键问题。在中医药 LLM 设计中融入 RAG 技术,将知识检索与生成相结合,通过 RAG 模型,可以从海量中医药文献中精准提取所需信息,并生成相应的回答,有助于中医药文本数据中隐性知识的发现。

中医药学术传承是中医药服务高质量发展的基础^[41]。中医古籍是中华传世古籍的重要组成部分,在中医药传承创新中具有重要地位^[42]。然而中医古籍文本的语言晦涩、数据结构非标准化^[43],理解这些古籍内容的难度相对较大。为构建兼容性强的中医古籍 LLM,研究者可以利用自然语言处理技术对古籍文本进行深入的语义分析和结构解析,提取关键术语和句式结构。然后,通过继续预训练方法,将大量中医药文本输入模型,使其在学习通用语言的同时,增强对中医术语的理解和应用能力。此外,融合有监督学习策略,利用专家标注的高质量数据集进行微调,确保模型输出的准确性和可靠性。通过以上方法,LLM 能解读中医古籍文本,为研究者和实践者提供精准、易于理解的知识服务,从而有效克服传统中医文本处理的难题。

2.4 训练与调优

中医药 LLM 需要在多种不同的任务和数据集上进行适应和微调,以满足不同的应用需求。LoRA 和监督微调是常用的 2 种微调策略。LoRA 技术通过引入低秩适应机制,在保持模型参数量不变的情况下,实现对不同任务的高效适应。在通用 LLM 基础上,运用 LoRA 技术对模型进行微调,使其能够更好地适应不同的中医药任务,如辨证论治、药物

配伍等任务,有助于实现较低算力成本下的模型性能提升。中医药 LLM 的构建需要不断进行优化和调整,以适应不断变化的应用需求和数据环境。监督微调技术通过在特定任务的数据集上进行进一步训练,使预训练模型能够更好地适应特定任务的需求,并通过标注数据集进行监督训练,提高中医药 LLM 在特定任务中的性能。

2.5 计算资源优化

在中医药 LLM 的开发过程中,算法效率和计算资源的优化是实现模型高效运行的关键因素。首先,针对中医药领域的特殊需求,可通过分布式计算框架如 Apache Spark 或 TensorFlow,将大规模的数据处理和模型训练任务分散到多个计算节点上,提高计算效率,并减少单个任务的处理时间。同时,还可利用 GPU 加速技术对模型训练过程进行加速,缩短训练周期。其次,为了优化模型的存储和计算性能,可采用知识蒸馏技术,减小模型规模,降低计算资源消耗,将复杂的模型简化,保持模型性能,使其更适合在资源受限的环境中部署和运行。此外,考虑到中医药数据特有的层次性和非结构性,还应及时监控模型的性能和计算负载,动态调整算法参数和计算资源分配。

3 中医药 LLM 的应用场景

3.1 中医药学术传承

LLM 能够处理大规模文本数据,提高信息检索的精确度和速度,张君冬等^[44]通过构建的中医古籍生成式对话 LLM,生成与古籍内容相关的深层次、动态的解答和解析,增强了知识的可访问性和实用性。结合自然语言处理技术对古籍文本进行语义分析和结构解析,构建中医药专业领域数据集,通过 LLM 的自然语言交互能力、知识理解和学习能力,深入挖掘和融合中医古籍资源,可为中医药学术传承提供有力支撑^[45]。

3.2 中药智能制造

智能制造通过融合尖端制造技术和先进的信息技术,已成为提升中药制造质量控制水平的前沿方法^[46],其对中药产业的数字化转型至关重要^[47]。中药智能制造的核心在于对数据的深度挖掘和智能分析,从原材料的采购、加工、配方、生产到最终成品的出库,每一个环节都伴随着大量数据的生成。构建中药智能制造 LLM,需要将分散在供应链各环节的数据进行有效整合,AIGC 技术在这一过程中发挥着至关重要的作用,能够模拟专家的决

策过程,识别数据中的模式和趋势,预测潜在的问题,并提出解决方案^[48]。利用 AIGC 技术,可以提高非结构化数据的利用效率。如在生产过程中,可以实时监控生产参数,如温度、湿度等,并根据实时数据调整生产条件,以确保生产过程的稳定性和产品质量的一致性。中药智能制造 LLM,不仅能够提升药品的质量控制水平,还能够推动中药制造行业的现代化和智能化发展。通过不断的技术创新和优化,中药智能制造 LLM 有望成为中药制造业的重要支撑,为传统医药的传承与发展注入新的活力。

3.3 中药全产业链智能决策

在前述中医药 LLM 构建框架的基础上,课题组前期融合千种本草基因组数据库等具有独立知识产权的多源异构海量数据,建立了含 2 000 余万个实体和超 20 亿个关系对的知识图谱;提出了覆盖中药全产业链各环节的知识学习及发现方法;采用 MoE 技术,开发了集中药知识获取、管理与服务于一体的中药全产业链 LLM——本草智库。多模态 LLM 对中药全产业链的优化和提升具有重要的战略意义。通过图像和文本的联合分析,有助于理解中药材的特性和作用机制,从而指导更精准的药物配伍和中药质量控制。因此,利用多模态 LLM,实现中药研究底层核心数据与中药全产业链关键环节的有机结合,将是团队在 LLM 开发领域的一个关键研究方向。

3.4 临床决策支持

LLM 在处理中医数据方面具有显著的优势,特别是在中医临床决策支持领域,其应用前景广阔^[49]。中医疾病诊断依赖于较为抽象的“辨证论治”理论体系,为辅助 LLM 深入理解“辨证论治”,利用 Transformers 模型学习和模拟中医诊疗过程中的逻辑和推理,从而提炼出疾病诊断、治疗方法与药物配伍的深层次知识。通过在特定的中医临床数据集上进行微调,增强对中医独特诊疗体系的理解,包括辨证论治原则、药物配伍禁忌等,LLM 能够为临床决策提供个性化的治疗建议。此外,还可以通过构建中医药知识图谱,将药物、疾病、症状及治疗原则等要素系统化,利用图神经网络理解和推理要素间复杂的关系,模拟中医药中的复杂逻辑和推理过程,增强 LLM 分析患者的症状描述、体质特点及历史治疗响应能力,综合生成符合中医理论的治疗方案,有效提升诊疗效率和精准度^[50]。

4 结语与展望

自 ChatGPT 发布以来,全球对此类模型的研究兴趣日益浓厚。众多研究机构和公司随之开发了类似的开源 LLM,如 LLaMA、ChatGLM-6B 和 Gemini 等,这些模型主要利用公开的学术文献和网络数据进行训练,应用广泛,但在特定行业的应用和精确度方面还存在限制。在中医药领域,尽管人工智能技术已被用于整理中医古籍、临床数据和药物信息等,但数据标准化和客观化程度不足,限制了人工智能技术在中医药领域的应用效果。因此,如何构建高质量的中医药数据集以提升模型训练效果,成为了中医药 LLM 开发领域的一个重要研究方向。

在中医药领域,随着 LLM 的广泛应用,确保数据安全和保护隐私变得尤为关键。为有效保护中医药数据的安全和隐私,可以采用分类与分层管理的策略。首先将数据根据敏感性和重要性进行分类,如将患者个人信息、临床试验数据和药物配方等高敏感性数据划分为最高保护等级。然后根据数据的分类结果实施分层管理,不同级别的数据采取不同级别的保护措施。随着技术的进步和法规的完善,中医药领域的数据安全管理将变得更加精细化和智能化,这不仅能够保障数据的安全性,也能提高数据的利用效率,为中医药领域的持续发展和创新提供支撑。

在中医药垂直领域 LLM 构建中,架构优化是提升其应用效果和精确度的关键。当前,模型架构主要围绕提高对中医药专业术语和治疗方法的理解能力进行优化。如通过增加与中医药相关的预训练任务,如古文解析和医案分类,可以显著增强模型的领域适应性。此外,多学科交叉融合为中医药 LLM 构建提供了新的机会。通过采用迁移学习和微调技术,将在其他领域训练好的模型适配到中医药特定任务上,能够在提升模型训练效率的同时,保证模型输出的专业性和准确性。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中华人民共和国国务院. 国务院办公厅关于印发“十四五”中医药发展规划的通知 [EB/OL]. (2022-03-29) [2024-06-15] https://www.gov.cn/zhengce/content/2022-03/29/content_5682255.htm.
- [2] 宫爱民,曹玉,董秀娟,等. 120例肝纤维化患者舌面象特征分析 [J]. 世界科学技术—中医药现代化, 2016, 18(10): 1646-1651.
- [3] Shi Y L, Liu J Y, Hu X J, et al. A new method for syndrome

- classification of non-small-cell lung cancer based on data of tongue and pulse with machine learning [J]. *Biomed Res Int*, 2021, 2021: 1337558.
- [4] Zhuo L, Zhang J, Dong P, *et al.* An SA-GA-BP neural network-based color correction algorithm for TCM tongue images [J]. *Neurocomputing*, 2014, 134: 111-116.
- [5] Yuan L, Yang L, Zhang S C, *et al.* Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: A prospective multicentre clinical cohort study [J]. *EClinicalMedicine*, 2023, 57: 101834.
- [6] Zhang Q, Zhou J H, Zhang B. Graph based multichannel feature fusion for wrist pulse diagnosis [J]. *IEEE J Biomed Health Inform*, 2021, 25(10): 3732-3743.
- [7] 王桂彬, 庞博. 名老中医隐性知识发现与医案解构模式研究 [J]. *中华中医药杂志*, 2023, 38(5): 2230-2234.
- [8] 刘莉, 邵波. 生成式 AI 赋能智慧图书馆的融合路径探析——以扎耶德大学图书馆为例 [J]. *图书馆学研究*, 2023, (12): 34-43.
- [9] 武子豪, 吴礼滨, 洪伟, 等. 机器学习在生态环境损害鉴定评估领域的应用前景 [J]. *农业环境科学学报*, 2023, 42(12): 2860-2868.
- [10] Lewis P S H. *Improving Neural Question Answering with Retrieval and Generation* [D]. London: University College London, 2022.
- [11] Lewis P, Perez E, Piktus A, *et al.* Retrieval-augmented generation for knowledge-intensive NLP tasks [EB/OL]. (2021-04-12) [2024-06-15]. <http://arxiv.org/abs/2005.11401>.
- [12] Ram O, Levine Y, Dalmedigos I, *et al.* In-context retrieval-augmented language models [J]. *Trans Assoc Comput Linguist*, 2023, 11: 1316-1331.
- [13] Kang B, Kim J, Yun T R, *et al.* Prompt-RAG: Pioneering vector embedding-free retrieval-augmented generation in niche domains, exemplified by Korean medicine [EB/OL]. (2024-1-20) [2024-06-15]. <http://arxiv.org/abs/2401.11246>.
- [14] Li J Q, Wang X D, Wu X B, *et al.* Huatuo-26M, a large-scale Chinese medical QA dataset [EB/OL]. (2023-05-02) [2024-06-15]. <http://arxiv.org/abs/2305.01526>.
- [15] Zhao Z J, Ren X T, Song K, *et al.* PreGenerator: TCM prescription recommendation model based on retrieval and generation method [J]. *IEEE Access*, 2023, 11:103679-103692.
- [16] Izacard G, Grave E. Leveraging passage retrieval with generative models for open domain question answering [EB/OL]. (2020-07-02) [2024-06-15]. <https://arxiv.org/abs/2007.01282>.
- [17] Lin J Y, Ma Z, Gomez R, *et al.* A review on interactive reinforcement learning from human social feedback [J]. *IEEE Access*, 2020, 8: 120757-120765.
- [18] Ahmadian A, Cremer C, Gallé M, *et al.* Back to basics: Revisiting REINFORCE style optimization for learning from human feedback in LLMs [EB/OL]. (2024-02-22) [2024-06-15]. <http://arxiv.org/abs/2402.14740>.
- [19] Hu J, Wu X B, Wang W X, *et al.* OpenRLHF: An easy-to-use, scalable and high-performance RLHF framework [EB/OL]. (2024-05-20) [2024-06-15]. <http://arxiv.org/abs/2405.11143>.
- [20] McIntosh T R, Susnjak T, Liu T, *et al.* The inadequacy of reinforcement learning from human feedback - radicalizing large language models via semantic vulnerabilities [J]. *IEEE Trans Cogn Dev Syst*, 2024, doi: 10.1109/TCDS.2024.337744.
- [21] Jacobs R A, Jordan M I, Nowlan S J, *et al.* Adaptive mixtures of local experts [J]. *Neural Computation*, 1991, 3(1): 79-87.
- [22] Eigen D, Ranzato M, Sutskever I. Learning factored representations in a deep mixture of experts [EB/OL]. (2013-12-16) [2024-06-15]. <http://arxiv.org/abs/1312.4314>.
- [23] Du N, Huang Y P, Dai A M, *et al.* GLaM: Efficient scaling of language models with mixture-of-experts [EB/OL]. (2021-12-13) [2024-06-15]. <http://arxiv.org/abs/2112.06905>.
- [24] Shen S, Hou L, Zhou Y Q, *et al.* Mixture-of-experts meets instruction tuning: A winning combination for large language models [EB/OL]. (2023-05-24) [2024-06-15]. <http://arxiv.org/abs/2305.14705>.
- [25] Sukhbaatar S, Golovneva O, Sharma V, *et al.* Branch-train-MiX: Mixing expert LLMs into a mixture-of-experts LLM [EB/OL]. (2024-05-12) [2024-06-15]. <http://arxiv.org/abs/2403.07816>.
- [26] Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network [EB/OL]. (2015-03-09) [2024-06-15]. <http://arxiv.org/abs/1503.02531>.
- [27] Sanh V, Debut L, Chaumond J, *et al.* DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter [EB/OL]. (2019-05-09) [2024-06-15]. <http://arxiv.org/abs/1910.01108>.
- [28] Jiao X Q, Yin Y C, Shang L F, *et al.* TinyBERT: Distilling BERT for natural language understanding [EB/OL]. (2019-09-23) [2024-06-15]. <http://arxiv.org/abs/1909.10351>.
- [29] Wu J L, Chung W Y. Sentiment-based masked language modeling for improving sentence-level valence-arousal prediction [J]. *Appl Intell*, 2022, 52(14): 16353-16369.
- [30] Li L, Wang Q P, Zhao B H, *et al.* Pre-training and fine-tuning with next sentence prediction for multimodal entity linking [J]. *Electronics*, 2022, 11(14): 2134.
- [31] Xia L Q, Li C X, Zhang C B, *et al.* Leveraging error-assisted fine-tuning large language models for

- manufacturing excellence [J]. *Robot Comput Integr Manuf*, 2024, 88: 102728.
- [32] Shen M. Rethinking data selection for supervised fine-tuning [EB/OL]. (2024-02-08) [2024-06-15]. <https://arxiv.org/abs/2402.06094>.
- [33] Xu H Y, Zhan R Z, Wong D F, *et al.* Let's focus on neuron: Neuron-level supervised fine-tuning for large language model [EB/OL]. (2024-05-18) [2024-06-15]. <http://arxiv.org/abs/2403.11621>.
- [34] Zhou H T, Liu T K, Ma Q L, *et al.* LoBaSS: Gauging learnability in supervised fine-tuning data [EB/OL]. (2023-10-16) [2024-06-15]. <http://arxiv.org/abs/2310.13008>.
- [35] Singhal K, Azizi S, Tu T, *et al.* Large language models encode clinical knowledge [J]. *Nature*, 2023, 620(7972): 172-180.
- [36] Han J M, Zhang R R, Shao W Q, *et al.* ImageBind-LLM: Multi-modality instruction tuning [EB/OL]. (2023-09-07) [2024-06-15]. <http://arxiv.org/abs/2309.03905>.
- [37] Hu E J, Shen Y L, Wallis P, *et al.* LoRA: Low-rank adaptation of large language models [EB/OL]. (2021-06-17) [2024-06-15]. <http://arxiv.org/abs/2106.09685>.
- [38] Qin H T, Ma X D, Zheng X Y, *et al.* Accurate LoRA-finetuning quantization of LLMs via information retention [EB/OL]. (2024-02-08) [2024-06-15]. <http://arxiv.org/abs/2402.05445>.
- [39] Ouyang L, Wu J, Xu J, *et al.* Training language models to follow instructions with human feedback [EB/OL]. (2022-05-04) [2024-06-15]. <http://arxiv.org/abs/2203.02155>.
- [40] 陈联忠, 邓超, 孟庆伟. 中医临床决策支持系统技术发展论述 [J]. *中国数字医学*, 2024, 19(5): 14-19.
- [41] 谷晓红, 于河. 名老中医经验传承与创新的问题及策略 [J]. *北京中医药大学学报*, 2022, 45(9): 919-924.
- [42] 李盼飞, 张楚楚, 李海燕. 科技赋能中医古籍精华传承与创新应用 [J]. *中医杂志*, 2023, 64(15): 1519-1524.
- [43] 曹霞, 裴丽. 中医古籍数字化建设实践分析与应对策略 [J]. *图书馆学研究*, 2016, 13: 42-44+66.
- [44] 张君冬, 杨松桦, 刘江峰, 等. AIGC 赋能中医古籍活化: Huang-Di 大模型的构建 [J/OL]. *图书馆论坛*, [2024-01-24]. <http://kns.cnki.net/kcms/detail/44.1306.G2.20240124.1341.002.html>.
- [45] 陈炫婷, 叶俊杰, 祖璨, 等. GPT 系列大语言模型在自然语言处理任务中的鲁棒性 [J]. *计算机研究与发展*, 2024, 61(5): 1128-1142.
- [46] 张磊, 孟昭平, 岳洪水, 等. 中药工业智能制造转型模式及监管问题探讨 [J]. *中国医药导刊*, 2024, 26(2): 129-138.
- [47] 中华人民共和国国务院. 中共中央国务院印发《“健康中国2030”规划纲要》[EB/OL]. (2016-10-25) [2024-06-15]. <https://www.rmzxb.com.cn/c/2016-10-25/1102812.shtml>.
- [48] 熊皓舒, 王璧璇, 侯健, 等. 生成式人工智能 (AI) 在中药智能制造及供应链中的应用场景设计与展望 [J]. *中国中药杂志*, 2024, 49(14): 3963-3970.
- [49] Zhang H Y, Wang X, Meng Z P, *et al.* Qibo: A large language model for traditional Chinese medicine [EB/OL]. (2024-03-24) [2024-06-15]. <http://arxiv.org/abs/2403.16056>.
- [50] Yang S H, Zhao H J, Zhu S B, *et al.* Zhongjing: Enhancing the Chinese medical capabilities of large language model through expert feedback and real-world multi-turn dialogue [J]. *Proceedings AAAI Conference Artificial Intelligence*, 2024, 38(17): 19368-19376.

[责任编辑 赵慧亮]