

【智慧医疗】

AI能力驱动的急诊教学分层模型构建与实证研究： 基于ChatGPT与DeepSeek的双模系统设计

马宏飞 张杨 杨杰 杨新文*

新疆医科大学第一附属医院

摘要：本研究构建了基于ChatGPT与DeepSeek双模系统的急诊医学分层教学模型，结合建构主义学习理论与布鲁姆教育目标分类学，设计了学习者画像与自适应教学策略。系统采用混合云架构，ChatGPT负责通用教学与交互，DeepSeek用于复杂临床推理。USMLE模拟测试准确率达87.6%，医学推理准确率达93%。一项为期3个月的随机对照试验（N=64）显示，实验组在知识掌握度、临床推理能力和学习满意度方面分别提高27%、35%和16个百分点（94%vs78%）。结果表明，该双模系统在急诊医学教学中具有显著效果，为AI辅助医学教育提供了可行的模型框架。

关键词：人工智能；急诊医学；分层教学；个性化学习；双模系统；随机对照试验

DOI：10.65976/3078-8137.2026.01.018

急诊医学具有时间敏感性强、病例复杂、决策节奏快等特点，对医学人才培养提出了更高要求。传统教学模式在师资配置、个体差异响应及评价方式等方面存在局限^[1-2]。近年来，大语言模型在医学知识与推理方面展现出潜力，例如ChatGPT在USMLE模拟考试中达到87.6%的准确率，DeepSeek在医学推理任务中的准确率达93%^[3]。但单一模型难以同时兼顾教学交互性与临床推理深度，因此有必要构建多模型协同的教学系统。基于此，本研究提出ChatGPT与DeepSeek双模协同的急诊教学系统，以实现分层化与个性化教学支持。

1 理论基础与文献综述

分层教学基于个体差异与认知负荷理论，通过调

节教学复杂度以提升学习效率。布鲁姆教育目标分类学为急诊医学能力分级提供了结构化框架^[4]。建构主义学习理论强调情境互动与反思建构，与病例驱动的急诊教学高度契合^[5]。现有研究表明，AI在医学教育中的应用多集中于虚拟病例、自适应学习与评估，但单模型架构难以同时满足交互性与专业推理需求^[6-10]，因此多模型协同具有现实必要性。

不同认知层次的教学任务可与相应的AI模块功能相对应，如表1所示。

2 双模系统架构设计

2.1 ChatGPT模块

ChatGPT用于通用医学教学、对话交互与情境模拟，可承担病史采集、基础知识讲解与沟通训练等任务^[1]。

表1 AI模型功能与布鲁姆认知层次的映射

布鲁姆认知层次	急诊医学应用场景	主要AI模型	主要功能类型
记忆 (Remembering)	基础解剖知识回顾、药理剂量查询	ChatGPT	通用知识检索与定义
理解 (Understanding)	病理生理机制解释、临床指南总结	ChatGPT	教学内容生成与交互
应用 (Applying)	标准操作流程模拟、简单病例实践	ChatGPT+DeepSeek 协同	基础案例演练与指导
分析 (Analyzing)	疾病鉴别诊断、复杂模式识别	DeepSeek	深度医学推理与模式识别
评估 (Evaluating)	治疗方案效果评价、风险效益分析	DeepSeek	批判性思维支持与决策分析
创造 (Creating)	创新诊疗方案设计、资源调配优化	DeepSeek 协同 (高复杂度)	复杂问题解决与方案创新

基金项目：本项目受到新疆维吾尔自治区青年科学基金：lncRNA-MALAT1在失血性休克缺血再灌注损伤中大鼠肠粘膜屏障作用的研究的支持（2022D01C764）。

作者简介：马宏飞（1987—），男，硕士，主治医师，研究方向为急腹症与多发伤。

通讯作者：杨新文，男，主任医师，研究方向为急腹症、腹部创伤等。

2.2 DeepSeek 模块

DeepSeek 用于高复杂度临床推理与案例分析，支持鉴别诊断与治疗决策。经医学语料微调后，其在医学推理任务中的准确率可达 93%，并可输出可解释的推理链^[11-13]。

2.3 双模协同机制

系统根据教学任务复杂度动态调用两种模型：一般知识与沟通类任务由 ChatGPT 处理，涉及诊断与决策的复杂问题由 DeepSeek 承担；高复杂度情形下两者协同工作，由 DeepSeek 提供结构化推理、ChatGPT 生成可读教学反馈^[3,14]。

双模系统采用分层架构设计（如图 1 所示），包括前端接口层、服务编排层、AI 模型层、数据管理层

和基础设施层。

如表 2 所示，双模系统的智能路由与推理融合流程包括从前端接收查询到最终反馈生成的各主要步骤。

3 分层教学模型构建

系统构建学习者画像（认知、技能、行为与情感维度），据此自适应调整教学内容与难度。评估体系涵盖知识、技能与态度，并提供即时与发展性反馈^[8-9]。

4 实证研究结果

采用随机对照试验（3 个月，N=64，实验组 / 对照组各 32 名住培医师）。具体对比结果见表 3。

5 系统优势与挑战分析

双模协同兼顾教学交互性与医学推理准确性，优

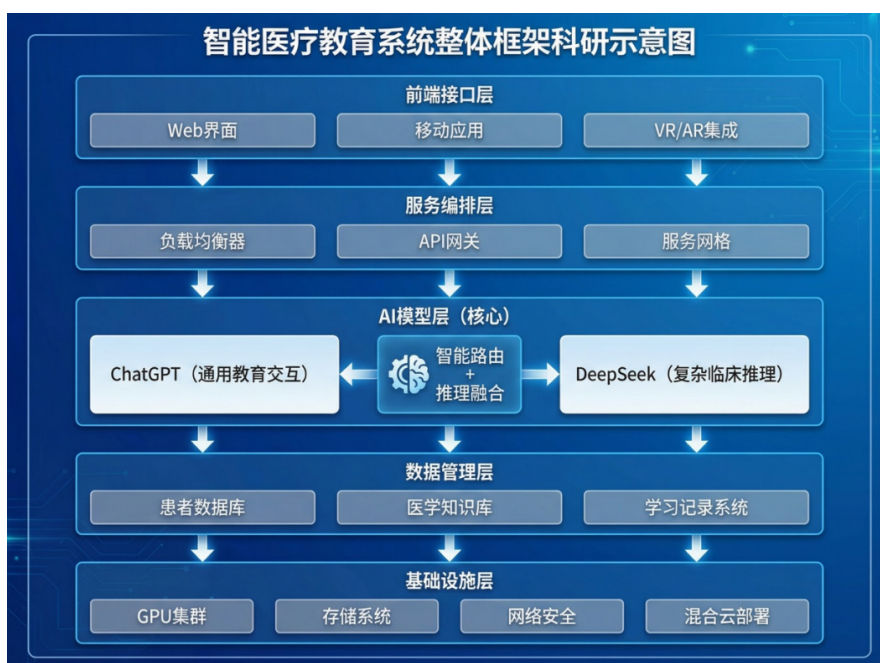


图 1 双模系统采用分层架构设计

表 2 双模系统智能路由与推理融合流程架构

流程节点	描述	AI 模型 / 机制
前端查询接收	接收学员请求（文本输入 / 语音指令 / VR 交互）	前端接口层
查询特征分析	评估查询类型（知识回顾 vs. 临床决策）， 计算查询复杂度分值 CS	服务编排层 / 知识图谱
智能路由决策 1	若 $CS \leq 0.5$ ：路由至通用医学教育模块，优化交互体验	ChatGPT 模块
智能路由决策 2	若 $0.5 < CS \leq 0.8$ ：路由至 DeepSeek 模块	DeepSeek 模块（高精度推理）
智能路由决策 3（协作）	若 $CS > 0.8$ ：触发高复杂度任务，启动 ChatGPT 与 DeepSeek 协同工作	DeepSeek+ 推理融合
推理融合与风险校准	采用加权集成、低置信度结果双模型验证等策略； DeepSeek 结果占较高权重	推理融合机制 / 服务网格
反馈生成	将最终验证后的结果转化为个性化、可解释的 反馈和指导	ChatGPT 模块（用于交互输出）
学习者画像更新	记录交互结果、常见错误模式及能力发展水平	数据管理层（学习记录系统）

表3 随机对照试验主要结果指标对比 (实验组 vs. 对照组)

结果指标	实验组较对照组提升幅度	统计显著性	效应量 (Cohen's d)	效应分类
知识掌握度	↑ 27%	$P < 0.01$	0.75	中等到大效应
临床推理能力	↑ 35%	$P < 0.001$	1.10	大效应
技能操作评分	↑ 22%	$P < 0.05$	0.55	中等效应
学习满意度	实验组 94% (对照组 78%)	$P < 0.01$	-	极高用户接受度
学习效率	完成任务时间减少 40%	-	-	高效率提升
自我效能感	↑ 31%	-	0.82	大效应
学习动机	↑ 28%	-	0.78	中等到大效应

于单模型^[3,12]; DeepSeek 支持本地化部署,有助于保障数据安全^[13]。主要风险在于生成式模型可能产生不准确输出,系统通过双模型交叉校验与教师监督降低风险,并遵循数据隐私与伦理规范,明确 AI 仅为教学辅助^[14-16]。

6 结论

基于 ChatGPT 与 DeepSeek 的双模分层教学模型在随机对照试验中显著提升了急诊住培医师的知识、推理与满意度,具有较高的可行性与推广价值。

参考文献:

[1] BASNAWI A, KOSHAK A. Application of Artificial Intelligence in Advanced Training and Education of Emergency Medicine Doctors: A Narrative Review[J]. *Emergency Care and Medicine*, 2024, 1(03): 247-59.

[2] NATESAN S, BAILITZ J, KING A, et al. Clinical Teaching: An Evidence-based Guide to Best Practices from the Council of Emergency Medicine Residency Directors[J]. *West J Emerg Med*, 2020, 21(04): 985-98.

[3] MOELL B, SAND ARONSSON F, AKBAR S. Medical reasoning in LLMs: an in-depth analysis of DeepSeek R1[J]. *Front Artif Intell*, 2025, 8: 1616145.

[4] DONG H, LIO J, SHERER R, et al. Some Learning Theories for Medical Educators[J]. *Med Sci Educ*, 2021, 31(03): 1157-72.

[5] TAN S, MILLS G. Designing Chinese hospital emergency departments to leverage artificial intelligence—a systematic literature review on the challenges and opportunities[J]. *Front Med Technol*, 2024, 6: 1307625.

[6] ASTER A, LAUPICHLER M C, ROCKWELL-KOLLMANN T, et al. ChatGPT and Other Large Language Models in Medical Education—Scoping Literature Review[J]. *Med Sci Educ*, 2025, 35(01): 555-67.

[7] FEIGERLOVA E, HANI H, HOTHERSALL-DAVIES E. A systematic review of the impact of artificial intelligence on educational outcomes in health professions education[J]. *BMC Med Educ*, 2025, 25(01): 129.

[8] MANSOOR M, IBRAHIM A, HAMIDE A. Performance of DeepSeek and GPT Models on Pediatric Board Preparation Questions: Comparative Evaluation[J]. *JMIR AI*, 2025, 4: e76056.

[9] SALMAN H, POWELL L, ALSUWAIDI L, et al. Profiling of Learners in Medical Schools as a Move Toward Precision Education: Protocol for a Scoping Review[J]. *JMIR Res Protoc*, 2022, 11(10): e41828.

[10] 杨海燕, 李涛. ChatGPT 教学应用: 场景、局限与突破策略[J]. *中国教育信息化*, 2023, 29(6): 26-34.

[11] SOENKSEN L R, MA Y, ZENG C, et al. Integrated multimodal artificial intelligence framework for healthcare applications[J]. *NPJ Digit Med*, 2022, 5(01): 149.

[12] TEMSAH A, ALHASAN K, ALTAMIMI I, et al. DeepSeek in Healthcare: Revealing Opportunities and Steering Challenges of a New Open-Source Artificial Intelligence Frontier[J]. *Cureus*, 2025, 17(02): e79221.

[13] DAVENPORT T, KALAKOTA R. The potential for artificial intelligence in healthcare[J]. *Future Healthc J*, 2019, 6(02): 94-8.

[14] PETERSSON L, LARSSON I, NYGREN J M, et al. Challenges to implementing artificial intelligence in healthcare: a qualitative interview study with healthcare leaders in Sweden[J]. *BMC Health Serv Res*, 2022, 22(01): 850.

[15] VAN LANKVELD T, SCHOONENBOOM J, KUSURKAR R A, et al. Integrating the teaching role into one's identity: a qualitative study of beginning undergraduate medical teachers[J]. *Adv Health Sci Educ Theory Pract*, 2017, 22(03): 601-22.

[16] RINCON E H H, JIMENEZ D, AGUILAR L A C, et al. Mapping the use of artificial intelligence in medical education: a scoping review[J]. *BMC Med Educ*, 2025, 25(01): 526.