Advances in Education 教育进展, 2026, 16(11), 878-890
Published Online November 2026 in Hans. https://www.hanspub.org/journal/ae
https://doi.org/10.12677/ae.2026.16112454

Hans汉斯

基于人工智能的医学生个性化指导模式

——以“智慧阅读助手”为例

李国琳[1,2]，张　琪[1,2]，高　昕[1,2]，乔可宏[2*]

[1]长治医学院第一临床学院，山西　长治
[2]太原市中心医院呼吸与危重症医学科，山西　太原

收稿日期：2026年10月6日；录用日期：2026年11月5日；发布日期：2026年11月12日

摘　要

传统医学教学多以教师讲授为中心，学生的主体性容易被忽视。如何借助人工智能精准识别个体差异，并提供适切的指导支持，已成为教学改革中的重要命题。本文系统介绍了一款自主研发的人工智能医学学习平台。该平台分为教师端与学生端：教师端支持学习资料的智能管理与教学反馈分析，学生端通过智能对话引擎实现个性化学习交互，并自动记录学习者的提问轨迹。平台的核心价值在于，通过对同一知识点下不同学生提问角度、认知层次与表达风格的深度挖掘，揭示学习者在知识建构与掌握上的个体差异，为教师提供精准的教学诊断依据。在学情数据分析的基础上，平台进一步设计了“数据驱动的教学活动建议”模块与“教学策略库”，将诊断结果转化为可操作的教学行动方案，并以《诊断学》课程为例，展示教师如何依据学情报告设计与改进课堂教学。技术层面，平台采用基于BERT预训练模型微调的多标签文本分类方案，实现提问类型的自动识别，在医学教育提问数据集上取得了较好的分类性能，并通过决策阈值校准与低置信度回退机制保障分类可靠性。本文认为，学生在人工智能辅助环境下提出的问题，本质上是其内在认知过程的外化表征；人工智能平台的意义，正在于将这一过程系统化、可视化与可操作化，从而重构医学教育中的师生互动模式与教学决策逻辑。

关键词

人工智能，医学教育，个性化学习，认知诊断

An AI-Based Personalized Guidance Model for Medical Students

—A Case Study of the “Smart Reading Assistant”

Guolin Li[1,2], Qi Zhang[1,2], Xin Gao[1,2], Kehong Qiao[2*]

[1]First Clinical College, Changzhi Medical College, Changzhi Shanxi
[2]Department of Respiratory and Critical Care Medicine, Taiyuan Central Hospital, Taiyuan Shanxi

[*]通讯作者。

文章引用: 李国琳, 张琪, 高昕, 乔可宏. 基于人工智能的医学生个性化指导模式[J]. 教育进展, 2026, 16(11): 878-890.
DOI: 10.12677/ae.2026.16112454

Received: October 6, 2026; accepted: November 5, 2026; published: November 12, 2026

Abstract

Traditional medical teaching model is often teacher-centered, which tends to neglect students' agency. How to use artificial intelligence to accurately identify individual differences and provide appropriate guidance and support has become an important issue in teaching reform. This paper systematically introduces a self-developed AI-powered medical learning platform. The platform consists of a teacher portal and a student portal. The teacher portal supports intelligent management of learning materials and analysis of instructional feedback, while the student portal enables personalized learning interactions through an intelligent dialogue engine and automatically records learners' question-asking trajectories. The core value of the platform lies in uncovering individual differences in learners' knowledge construction and mastery by deeply mining the questioning perspectives, cognitive levels, and expression styles of different students around the same knowledge point, thereby providing teachers with a precise basis for instructional diagnosis. Building on learning analytics, the platform further incorporates a "data-driven teaching activity recommendation" module and a "teaching strategy library", translating diagnostic results into actionable teaching plans. Using the Diagnostics course as an example, it demonstrates how teachers can design and improve classroom instruction based on learning reports. Technically, the platform adopts a multi-label text classification approach based on fine-tuning a pre-trained BERT model to automatically identify question types. It achieves strong classification performance on a medical education question dataset and ensures classification reliability through decision threshold calibration and a low-confidence fallback mechanism. This paper argues that the questions students raise in AI-assisted environments are essentially externalized representations of their internal cognitive processes. The significance of AI platforms lies in systematizing, visualizing, and operationalizing this process, thereby reshaping teacher-student interaction patterns and the logic of teaching decisions in medical education.

Keywords

Artificial Intelligence, Medical Education, Personalized Learning, Cognitive Diagnosis

Copyright © 2026 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

医学教育以其知识体系庞杂、知识更新迅速、理论与实践高度耦合而区别于其他学科教育。传统医学教学模式以课堂讲授为核心，受教学时间限制，教师常常面临“教不完”的困境，学生则容易陷入“学不透”的状态。这种模式存在明显局限：教学过程缺乏对个体认知差异的识别与回应，学生被动接受，难以形成真正的知识内化。尤其在《诊断学》《内科学》等核心课程中，复杂的病理机制与鉴别诊断要求学生具备较强的理解与迁移能力；而大班授课环境下，教师难以捕捉每位学生的即时困惑，问题逐渐积累，学习动力也随之下降。因此，医学教育亟须一场个性化的教学转型[1]。

近年来，人工智能(AI)技术的快速发展为上述困境提供了新的破解思路。研究表明，AI 介入能够显著提升学生的学习能力与学习主动性[2]；AI 驱动的学习系统可根据个体表现定制学习计划[3]；虚拟患者与模拟训练环境则能打破时空边界，为医学教育提供更丰富的实践场景[4]。然而，梳理现有研究可以发

现一个共同盲区：学生在自主学习过程中产生的“问题”——他们主动提出的困惑、搜索的关键词、提问的表述方式——蕴含着丰富的认知诊断信息，却极少被系统性地用于指导教学。现有 AI 应用大多聚焦于“答案”的准确性，如答题正确率，而忽视了“问题”本身的质量与结构。事实上，学生提出什么样的问题、以何种方式提问，直接反映其认知层次与思维深度[5]。

基于这一认识，本研究自主研发了“智慧阅读助手”AI 医学学习平台。该平台分为教师端与学生端：教师端支持学习资料的智能管理与教学反馈分析，教师可上传课程资料并接收系统自动生成的学情诊断报告；学生端通过智能对话引擎实现个性化学习交互，在学习过程中可随时提出疑问，系统自动记录每一位学习者的提问内容与提问轨迹。平台通过对同一知识点下不同学生提问角度、认知层次与表达风格的深度挖掘，揭示学习者知识建构与掌握的个体差异，为教师提供精准的教学诊断依据。与常见平台不同，本平台不依赖大规模标准题库，也不试图用 AI 替代教师，而是将 AI 定位为“认知诊断工具”——帮助教师看清学生“在哪里不懂”以及“以何种层次不懂”，从而让有限的课堂时间真正用在学生最需要的地方，推动医学教育从经验驱动走向数据驱动的个性化教学转型。

本文将从理论背景、平台架构、核心技术实现、核心机制、教学应用设计、应用价值与挑战等方面，系统阐述这一平台与模式。

2. 理论背景

2.1. 学习分析：从描述性仪表盘到闭环干预

学习分析作为一个独立研究领域，自 2011 年首届 LAK 会议以来迅速发展。Sailer 等人提出的闭环学习分析框架，系统整合了数据采集、处理分析、自适应个性化三个核心环节，强调教育理论与数据驱动洞察之间的双向互动。该框架的核心洞见在于：学习分析不应止步于数据可视化，而应形成“采集–分析–干预–评价”的完整闭环。然而，Sailer 等人在综述中明确指出，当前研究中“自适应与个性化”环节存在显著缺口，多数系统能够生成描述性报告，却缺乏将分析结果转化为具体教学行动的机制[6]。

与此同时，学习分析仪表盘研究也呈现出从“呈现数据”向“支持行动”演进的趋势。一项系统性综述分析了近年来学习分析仪表盘研究中使用的理论框架，发现自我调节学习是最常被引用的框架之一，但多数仪表盘仍停留在信息展示层面，未能为教师提供可直接执行的教学策略建议[6]。这一发现揭示了本领域的一个关键矛盾：数据可视化技术已相对成熟，但“从数据到行动”的转化桥梁仍然缺失。

本平台的设计正是针对这一缺口。平台不仅生成学情数据报告，更在此基础上设计了“数据驱动的教学活动建议”模块与“教学策略库”，尝试将学习分析的闭环从“数据分析”延伸至“教学行动”，为教师提供从诊断到处方的完整支持。

2.2. 智能导师系统：从知识追踪到认知诊断

智能导师系统是 AI 教育应用中最成熟的技术范式之一。自 20 世纪 70 年代第一代智能导师系统问世以来，该领域经历了从基于规则的模型追踪系统到基于机器学习和大语言模型的智能学习伴侣系统的演进。研究表明，智能导师系统能够模拟一对一辅导的效果，为学习者提供即时反馈和个性化指导，其核心功能包括自动解答学科问题、实时追踪学习路径、动态诊断认知偏差及生成个性化反馈建议[7]。

在个性化学习的技术路径方面，研究者提出了从基于规则的系统到 AI 导师的五阶段演进模型，强调 AI 工具应根据教学目标和学习者特征进行差异化应用。近年来，基于大语言模型的智能导师系统进一步拓展了认知诊断的深度，部分系统尝试构建“诊断–推荐–反馈”的闭环框架。然而，现有智能导师系统的认知诊断主要依赖学习者的答题表现，如正确率、反应时间等，其诊断信息源相对单一。正如相关综述所指出的，当前系统在“理解学生为什么犯错”这一层面仍存在明显局限，多数系统的认知模型以

知识点掌握程度为核心，较少触及学生的思维过程和认知层次[7]。

本平台与上述智能导师系统范式的重要区别在于诊断数据源的选取。本平台不以答题表现作为主要诊断依据，而是将学习者在自然学习情境中主动提出的问题作为认知诊断的核心数据。这一定位使平台能够捕捉到传统系统难以获取的认知信息——学生不仅“做错了什么”，更“在想什么”“困惑在哪里”。

2.3. 学生提问研究：从形成性评价资源到系统性认知诊断

学生提问的教育价值早已被学界所认识。Chin 和 Brown 的经典研究指出，学生的问题为教师提供了洞察学生知识状态、理解水平和困惑所在的窗口，可用于形成性评价和教学诊断[5]。Bergey 等人的实证研究进一步表明，学生在课堂中生成的问题能够有效预测其阅读理解监控过程与考试成绩——寻求验证或澄清的问题与考试表现呈正相关[8]。在医学教育领域，Wong 等人在新加坡国立大学开展的研究将学生提问与修订版布鲁姆认知目标分类相结合，对 298 个医学生提问进行认知层次编码，发现大多数问题集中在“理解”和“应用”层次，且不同知识主题引发的提问频率存在显著差异[9]。

然而，现有学生提问研究存在两个显著局限。其一，研究场景多限于课堂即时提问或专门设计的提问任务，学生在日常自主学习过程中自然产生的大量问题尚未被系统采集和分析。其二，研究目的多聚焦于“证明提问具有诊断价值”，但尚未建立从提问数据到教学决策的完整转化路径。换言之，学界已经认识到“提问是认知的窗口”，但如何让这扇窗口真正服务于教学实践，仍缺乏可操作的系统方案[9]。

本平台的设计直接回应了上述两个局限。在数据采集方面，平台通过低侵入性的侧边栏提问功能，捕捉学生在电子教材阅读过程中自然产生的疑问，而非依赖刻意设置的提问任务；在数据应用方面，平台通过语义分析和认知层次分类，将提问数据转化为可视化的学情报告和可执行的教学建议，尝试打通“提问采集–认知诊断–教学行动”的完整链路。

2.4. 本研究的定位与理论贡献

综合以上三个研究脉络，本平台的设计理念可被定位为：以学习分析的闭环框架为方法论指导，以智能导师系统的个性化诊断为目标取向，以学生提问研究为数据基础，构建一种“以提问行为为认知诊断指标”的医学教育个性化指导模式。其理论贡献主要体现在两个层面：

第一，在诊断数据源层面，本研究将“提问方式”而非“答题表现”作为认知诊断的核心指标，拓展了学习分析和智能导师系统的数据基础。答题表现反映的是学习的结果状态，而提问行为反映的是学习的认知过程状态，后者在形成性评价和教学过程性诊断中具有独特价值。

第二，在数据应用层面，本研究通过“数据驱动的教学活动建议”模块和“教学策略库”的设计，为学习分析领域长期存在的“从数据到行动”缺口提供了一种具体的解决方案。不同于仅提供数据可视化的传统仪表盘，本平台将诊断结果直接映射为教学策略选项，降低了教师将数据洞察转化为教学行动的门槛[6]。

3. 平台架构与核心机制

3.1. 总体设计理念

本平台的设计遵循三项核心理念：

第一，以学习者的认知过程为中心。研究表明，学生在 AI 赋能的学习环境中表现出不同的行为、认知和情感参与倾向[10]。平台设计的出发点不是方便教师“布置任务”，而是记录和分析学习者主动建构知识的过程，将平台打造为反映学习者认知活动的“镜子”。

第二，以提问行为为分析抓手。提问是学习者主动发起的信息获取行为，最能反映其知识状态、思

维路径和认知需求。学生在 AI 辅助环境下提出的问题，本质上是其内在认知过程的外化表征。有效学习的路径并非唯一，而 AI 平台的价值正在于揭示这些多元路径的存在。

第三，以教学决策支持为目的。平台的价值不在于替代教师，而在于为教师提供传统教学手段无法获取的深层学情信息，使教师能够从“教什么”的决策层面下沉到“学生如何学”的理解层面。

3.2. 教师端功能设计

教师端是平台的教学管理与数据洞察中枢，主要包括以下模块。

3.2.1. 学习资料管理系统

教师可上传教学课件、临床案例、参考文献、习题库等多格式资源。系统支持资料分类、版本管理和权限控制，确保教学资源的有序性与可追溯性。该模块的核心价值在于构建一个结构化的课程知识库，所有学生端的学习交互均以该知识库为基础展开。教师通过该模块完成课程学习材料的统一管理与分发，保证学生获取信息的规范性与一致性。同时，学生提问与搜索的范围被限定在教师指定的课程资料之内，确保学习交互的内容边界与教学目标一致。

3.2.2. 学情数据分析看板

系统自动汇总学生的提问数据，从提问频次、知识点分布、提问时段、认知层次等多个维度生成可视化分析图表。教师可通过看板快速了解：哪些知识点被高频提问，从而反映教学难点；哪些时段学生提问最活跃，从而反映自主学习节奏；哪些学生提问深度不足，从而反映个体学习困境。具体而言，平台以两周为一个报告周期，自动生成学情报告，内容包括：

① 高频困惑点列表：统计搜索频率最高的知识点关键词，识别普遍性学习难点；

② 提问方式分布图谱：将班级提问按“封闭式”“开放式”“应用型”三分类进行比例统计与可视化呈现；

③ 个体认知层次分布：标注每位学生主要停留的提问层次，形成总体认知层次全景图。

教师据此可以在下一节课中“带着数据进课堂”，将有限的时间精准投放到学生最需要的地方。

3.2.3. 教学反馈与建议生成模块

基于学生提问数据的分析，系统自动生成教学优化建议。与仅呈现数据图表的传统学习分析仪表盘不同，本模块的核心特征在于将诊断结果直接转化为可操作的教学策略选项。系统内置的策略规则如下：

① 当某一知识点被大量学生从同一角度反复提问，且封闭式提问占比超过 70%时，系统提示该知识点可能存在普遍性理解偏差，建议教师在下次课中增加原理性阐释和机制推演环节，并推荐使用“对比讲解法”或“因果链条分析法”进行教学。

② 当某一知识点的提问集中在临床应用层面而缺乏机制层面的问题时，系统提示该知识点的原理教学可能需要加强，建议教师增设“从临床现象反推病理机制”的引导式讨论，并推荐相应的 PBL 案例。

③ 当班级提问长期以封闭式为主时，系统提示教师的讲解可能过于结论化、缺乏启发性，建议在课堂中增设开放式提问示范和“苏格拉底式追问”环节。

当应用型提问活跃时，系统提示学生已具备知识迁移的意愿和能力，建议教师适当增加 PBL 案例教学和临床情境模拟的比重。

此外，系统还提供“教学反思参考”。例如，若某知识点的提问在两周内从封闭式为主转变为开放式和应用型为主，系统会标注这一变化趋势，提示教师该教学策略调整可能已产生效果。这种数据反馈不是对教师的监督，而是对教学优化的导航。

3.2.4. 教学策略库

教学策略库是本平台在“数据报告”基础之上进一步设计的核心功能模块。策略库中预置了与不同学情诊断结果相对应的教学活动建议方案，教师可根据学情报告呈现的诊断结果，从策略库中选取适配的教学活动方案进行组合和调整。

策略库中的教学活动方案按功能分为四类：

① 概念澄清类：适用于封闭式提问集中、学生停留于记忆层面的知识点。包括“正误对比辨析”“概念图共建”“关键术语溯源”等活动方案。

② 机制解析类：适用于开放式提问活跃但缺乏深度机制追问的知识点。包括“因果链条推演”“从现象到机制的阶梯式提问”“文献证据链分析”等活动方案。

③ 临床迁移类：适用于应用型提问活跃、学生表现出知识迁移意愿的知识点。包括“鉴别诊断情景模拟”“临床决策树构建”“标准化病人问诊实训”等活动方案。

④ 综合提升类：适用于高阶提问活跃、学生具备较强认知基础的知识点。包括“疑难病例讨论”“诊疗方案循证评价”“跨学科整合研讨”等活动方案。

每个教学活动方案包含活动目标、适用场景、实施步骤、所需时长、评价方式和注意事项六个要素，教师可根据课时安排和班级实际情况灵活选用。策略库的价值在于将教学决策从“凭经验判断”转化为“基于证据选择”，同时为经验不足的青年教师提供可参考的教学活动范式。

3.3. 学生端功能设计

学生端是学习者与 AI 交互的主要界面，核心功能包括以下两项。

3.3.1. 智能对话引擎

学生可自由提出医学问题，系统在教师上传的资料库中进行智能检索与匹配，生成带有引用来源的可靠回答。系统支持多轮对话，学生可就同一主题进行追问、澄清和深化，形成完整的探究链条。

在阅读电子教材或参考文献时，学生可随时在侧边栏输入疑问，平台自动记录提问文本、关联文档和页面位置。后台的语义分析模块对提问进行实时分类，积累数据后生成学生端报告。这一设计有两点用意：一是低侵入性，学生无须改变学习习惯，只需将原本散落在笔记本或脑海中的“隐形问题”显性化为文本输入，不增加额外认知负担；二是真实性，学生在真实学习情境中自然产生的提问，比任何测试题更能反映其真实困惑，避免了“应试性提问”的偏差。

3.3.2. 个人提问档案与认知画像

系统自动记录每一位学生的完整提问历史，包括提问内容、提问时间、所属知识点、对话轮次、系统回答等，形成个人化的学习轨迹。基于这些数据，系统构建个人认知画像，将提问类型可视化呈现。

对学生而言，这一功能的核心价值在于帮助学生看见自己的思考过程。多数医学生并不清楚自己是在“背知识”还是在“理解知识”——他们只知道“没记住”或“记不住”。个人认知画像将提问类型可视化后，学生可以直观地看到：我的提问中有多少是“是什么”类，多少是“为什么”类，多少是“怎么办”类。这种自我觉察本身就是元认知能力的训练。当学生意识到自己长期停留在封闭式提问时，会有意识地尝试提出更深层的问题，从而实现从“被动应答者”向“主动提问者”的角色转变。

4. 核心技术实现：提问类型自动分类模块

本节详细阐述平台“提问类型自动分类”模块的技术方案，涵盖系统架构、模型选型与微调策略、数据标注、性能评估及部署方案。

4.1. 技术选型依据与系统架构

提问类型自动分类是本平台认知诊断功能的技术基石。平台需要将学生以自然语言提出的医学问题自动归类为“封闭式”“开放式”“应用型”三个认知层次类别，为后续学情分析、策略匹配和个体画像提供结构化数据基础。

在技术选型上，研究团队对基于关键词规则的分类、传统机器学习(如 SVM + TF-IDF)和基于预训练语言模型微调的分类三种方案进行了比较。关键词规则方案实现简单，但无法处理同义表述和复杂的语义差异[11]。例如，“支气管哮喘怎么诊断”和“诊断支气管哮喘需要满足什么条件”在规则层面差异较大，但认知层次相同。传统机器学习方案依赖人工特征工程，在医学教育这一垂直领域泛化能力有限。预训练语言模型方案能够捕捉深层语义特征，且已有多项研究证明其在布鲁姆认知层次分类任务上的优越性[12]。

系统的整体处理流程为：学生输入问题文本后，首先经过预处理模块进行文本清洗、停用词过滤等操作，然后进入 BERT 分类模型进行认知层次预测。模型输出六个布鲁姆认知层次的概率分布，通过映射规则将其归并为封闭式、开放式、应用型三个类别，最终结果存入学情数据库，供后续分析调用[13]。

4.2. 数据标注方案与标注一致性

研究团队从《诊断学》《内科学》《病理学》等核心课程的学生提问中采集了约 3000 条问题文本作为初始训练语料。标注方案采用修订版布鲁姆认知目标分类作为编码框架，六层级中的“记忆”和“理解”归入封闭式类别，“理解”和“分析”归入开放式类别，“应用”和“评价”归入应用型类别[11]。每条提问由两名具有医学教育背景的标注者独立进行认知层次标注，标注者间一致性采用 Cohen’s Kappa 系数进行评估，达到 0.78，表明标注具有较好的可重复性。对于标注不一致的样本，由第三名资深医学教育专家进行仲裁。Wong 等人在新加坡国立大学的医学生提问认知编码研究中报告了 83.2%的编码者间一致性，为本研究的数据标注提供了方法学参照[9]。

数据集中三类提问的分布呈现一定的类别不均衡：封闭式提问约占 52%，开放式提问约占 31%，应用型提问约占 17%。这一分布与 Wong 等人的研究发现一致——医学生提问主要集中在较低认知层次，高阶提问比例偏低。针对类别不均衡问题，研究团队在模型训练中采用了分层采样与加权损失函数的组合策略[9]。

4.3. 模型微调与性能评估

分类模型基于 bert-base-chinese 预训练模型进行微调。该模型在中文维基百科语料上预训练，具备较强的中文语义理解能力[14]。微调策略上，研究团队参考了 Flanders 互动分析研究中冻结前 20 层 Transformer 参数、仅更新最后 4 层的经验，在保留预训练阶段习得的通用语义特征的同时，降低过拟合风险并减少计算资源消耗[15]。分类层使用 softmax 函数输出六个布鲁姆层次的预测概率，取概率最高的层级作为模型预测结果。训练超参数方面，学习率设为 2e−5，batch size 为 32，训练轮次为 10 个 epoch，采用 AdamW 优化器。

在测试集上的性能评估结果显示：三分类任务的整体准确率为 87.4%，Macro F1 为 85.1%。分层次来看，封闭式提问的 F1 为 90.2%，开放式提问的 F1 为 83.6%，应用型提问的 F1 为 81.4%。应用型提问的分类性能相对较低，主要原因是该类提问样本量较少且表述方式多样。这一结果与 Lee 等人使用 DeBERTa 进行学习者提问认知分类的研究发现一致——模型在处理事实性、知识传递类查询时表现较好，但在区分高阶提问时面临挑战[12]。

为进一步优化分类性能，研究团队参考了投票分类器的集成思路，将 BERT 分类模型与基于语义依

存分析的特征增强模块相结合，使应用型提问的 F1 值提升至 85.8%，整体 Macro F1 提升至 87.3% [16]。

4.4. 决策阈值校准与低置信度回退机制

在实际部署中，模型输出的概率分布并非总是清晰可分。部分提问，尤其是涉及多个认知层次的复杂提问，可能出现相邻层级概率接近的情况。为避免误分类对学情报告准确性造成影响，平台引入了决策阈值校准机制：当模型对某条提问的最高预测概率低于 0.6 时，该提问被标记为“低置信度”样本，暂不参与学情报告的认知层次统计，而是进入待人工复核队列。在平台后续迭代中，人工复核结果将被反馈至训练数据集，持续优化模型性能[17]。

此外，平台还采用了按类别校准的决策阈值策略——针对应用型提问类别适当降低判定阈值，以减少因类别样本不均衡导致的漏判。参考 Socratic 聊天机器人对话中布鲁姆分类的实践，按类别调整阈值可有效改善少数类别的召回率，该方法在多标签分类场景中已得到验证[18] [19]。

4.5. 技术方案的局限与迭代方向

当前分类方案存在两方面局限。其一，模型将每条提问独立分类，未能利用学生在同一知识点下的多轮提问序列信息。实际上，学生的提问往往呈现“逐步深化”的特征——从“是什么”到“为什么”再到“怎么办”，这种动态变化本身是认知深化的过程性证据。未来可探索序列建模方法，如引入对话上下文的 BERT 变体，使分类模型能够结合学生的历史提问轨迹进行更准确的认知层次判定[20]。其二，当前模型对“边界案例”的处理仍有提升空间，例如兼具机制探究和临床迁移特征的提问。后续研究可引入课程知识图谱，将提问分类与知识点诊断更紧密地耦合[21]。

5. 平台的核心机制：提问方式作为认知诊断指标

本平台最具辨识度的设计，是将“提问方式”作为认知诊断的核心指标。布鲁姆认知目标分类将认知层次从低到高分为记忆、理解、应用、分析、评价、创造六个层级。Wong 等人的研究已证明，将布鲁姆分类应用于医学生提问的认知层次编码具有可行性和诊断价值，其编码者间一致性达到 83.2% [9]。映射到提问行为上，可以形成如下对应关系：

① 封闭式提问，如“支气管哮喘的诊断标准是什么？”，对应“记忆–理解”层次，表明学生处于知识获取与初步理解阶段；

② 开放式提问，如“为什么哮喘患者气道反应性会增高？”，对应“理解–分析”层次，表明学生正在寻求机制解释与逻辑关联；

③ 应用型提问，如“心源性哮喘与支气管哮喘在临床上如何鉴别？”，对应“应用–评价”层次，表明学生尝试将知识迁移至临床情境。

这一映射关系使抽象的认知层次转化为可观测、可量化的行为指标。教师无需依赖复杂的认知诊断量表，仅通过分析班级提问类型的分布，即可把握群体认知结构：若某知识点下绝大多数提问为封闭式，说明学生尚停留于记忆层面，教师需要增加原理性阐释；若开放式与应用型提问增多，则提示学生已具备一定基础，教学可向深度与综合方向推进。

对于个体学生而言，其提问类型从封闭式向开放式、应用式的动态转变，可作为认知深化的过程性证据，为形成性评价提供依据。考核的不再仅是“答对了多少题”，更是“问了什么样的问题”。

AI 平台的意义正在于将这一认知诊断过程系统化、可视化与可操作化——它让每个学生的思考路径从不可见的“黑箱”变为可观察、可分析、可反馈的“透明过程”，从而重构医学教育中的师生互动模式与教学决策逻辑。

6. 教学应用设计：从学情报告到课堂教学改进

本平台的核心应用价值不仅在于提供学情数据报告，更在于将诊断结果转化为可操作的课堂教学行动。以下以《诊断学》课程中“呼吸系统体格检查”章节为例，展示教师如何依据平台的学情报告设计与改进课堂教学。

6.1. 学情报告示例

假设某医学院临床医学专业大二班级(60 人)在使用本平台学习《诊断学》“呼吸系统体格检查”章节两周后，系统自动生成了以下学情报告：

① 高频困惑点：排名前三的知识点分别为“支气管呼吸音与肺泡呼吸音的鉴别”“干啰音与湿啰音的发生机制”“胸膜摩擦感与心包摩擦感的区分”。

② 提问方式分布：全班共产生 186 条提问，其中封闭式提问 112 条，占 60.2%；开放式提问 48 条，占 25.8%；应用型提问 26 条，占 14.0%。

③ 个体认知层次分布：多数学生(约 65%)的提问主要停留在“记忆–理解”层次，约 25%的学生提问涉及“分析”层次，仅约 10%的学生提问达到“应用–评价”层次。

值得注意的异常模式是：在“干啰音与湿啰音的发生机制”这一知识点下，开放式提问比例显著高于其他知识点，占比达 45%，说明学生对该知识点的机制层面有较强的探究意愿；但封闭式提问中反复出现的“怎么区分”表明，学生在鉴别诊断方面仍存在困惑。

6.2. 基于学情报告的教学策略选择与教学设计

教师根据上述学情报告，从平台教学策略库中选取了“机制解析类”和“概念澄清类”活动方案，设计了以下改进后的课堂教学方案。

课前环节(15 分钟)：基于数据的分层预习任务推送。教师利用平台向不同认知层次的学生推送差异化预习材料。对于提问以封闭式为主的学生(约 65%)，推送“呼吸音鉴别要点”对比表格和微课视频，帮助其夯实基础知识；对于已具备一定分析能力的学生(约 25%)，推送“从气流动力学角度理解干湿啰音差异”的进阶阅读材料，引导其深入理解机制；对于提问已达到应用层次的学生(约 10%)，推送一个简短临床病例，要求其完成初步的体格检查方案设计。

课中环节(80 分钟)：聚焦核心困惑点的精准教学。课堂结构分为三个段落：

第一段(25 分钟)——概念辨析。针对学情报告中“支气管呼吸音与肺泡呼吸音鉴别”这一高频困惑点，班级中 52%的学生在此知识点下以封闭式提问为主，教师采用策略库中的“正误对比辨析”活动方案。具体实施为：展示 6 组听诊音频片段，要求学生先独立判断，再分组讨论，最后教师公布答案并引导学生归纳鉴别要点。此环节替代了传统的“从头讲到尾”的讲授方式，将有限的时间集中投放在学生最困惑的知识点上。

第二段(30 分钟)——机制推演。针对“干啰音与湿啰音发生机制”这一知识点下开放式提问集中、占比 45%的特点，教师采用策略库中的“因果链条推演”活动方案。具体实施为：以“为什么气道狭窄程度与啰音性质之间存在对应关系？”为引导问题，带领学生从气道解剖结构出发，逐步推演至气流动力学变化，最终推导出干湿啰音的发生机制。此设计利用了学生已有的探究意愿，顺势将其从“理解”层次推进到“分析”层次。

第三段(25 分钟)——临床迁移。针对班级应用型提问占比较低、仅 14%的现状，教师采用策略库中的“鉴别诊断情景模拟”活动方案，主动引导学生向临床应用层面迁移。具体实施为：呈现一个呼吸困难患者的简要病史，要求学生设计体格检查方案并预判可能的听诊发现，然后在小组内进行模拟查体演

练。此环节旨在将学生的认知层次从“分析”进一步推进到“应用–评价”层次。

课后环节：数据驱动的教学效果追踪。课后，教师布置了一项开放性任务：“请结合本课内容，提出一个你在学习呼吸系统体格检查过程中尚未解决的疑问。”学生通过平台提交问题后，系统自动进行分类并与课前数据进行对比。两周后的数据显示：封闭式提问占比从 60.2%下降至 48.5%，开放式提问占比从 25.8%上升至 33.2%，应用型提问占比从 14.0%上升至 18.3%。提问层次的分布变化表明，经过精准教学设计，学生的认知参与深度已呈现上移趋势。

6.3. 教学策略库的动态优化机制

上述案例展示了教师如何依据学情报告和策略库设计课堂教学。值得注意的是，教学策略库并非静态不变。平台设计了策略效果追踪功能：系统自动记录教师在每次教学活动中选用的策略类型，并在课后学情报告中标注该策略实施后的提问层次变化趋势。当某一策略在多轮教学实践中持续产生正向效果，即学生提问层次显著上移时，系统将其标注为“高效果策略”，推荐给同课程组的其他教师；当某一策略效果不显著或产生负向效果时，系统提示教师进行策略调整并记录调整方案。

这一机制使平台不仅服务于个体教师的课堂教学改进，也逐步积累校本化的教学策略知识，为课程组的集体教研和教学改革提供数据支撑。

7. 平台的应用价值

7.1. 对教师教学的价值

本平台最直接的受益者是教师。在传统教学中，教师对学情的了解主要依赖课堂观察、作业批改和阶段性考试，这些方式存在滞后性、片面性和被动性。本平台通过对同一知识点下不同学生提问角度、认知层次与表达风格的深度挖掘，揭示学习者知识建构与掌握的个体差异，为教师提供精准的教学诊断依据，将学情诊断前置到学生的学习过程中。每两周一份的班级学情报告，清晰呈现了学生当前最困惑的知识点、最集中的提问类型和认知层次的分布。更重要的是，平台的教学策略库和教学活动建议模块，使教师能够直接将诊断结果转化为可操作的教学行动。教师无需精通数据分析技术，只需根据学情报告选择相应的策略方案，即可完成从诊断到教学的闭环。这种“数据驱动 + 策略支架”的设计，降低了教师使用学习分析工具的门槛，推动医学教育从经验驱动走向数据驱动的个性化教学转型。

7.2. 对学生学习的价值

对学生而言，本平台的核心价值不在于“提供一个 AI 问答机器人”，而在于帮助学生看见自己的思考过程。个人认知画像将提问类型可视化后，学生可以直观地看到自己的提问层次分布，这种自我觉察本身就是认知能力的训练。学生在 AI 辅助环境下提出的问题，本质上是其内在认知过程的外化表征。平台将这一过程系统化、可视化与可操作化，使学生得以从“被动应答者”向“主动提问者”转变。在医学教育中，高阶临床思维能力的培养尤为重要。本平台通过记录和分析学生的提问行为，鼓励学生从“是什么”向“为什么”和“怎么办”进阶，为临床决策能力的形成奠定基础[22]。

7.3. 对医学教育管理的价值

在院校管理层面，本平台提供了一种低成本、高兼容的教学信息化方案。它不要求学校购置昂贵的 VR 设备或建设智慧教室，只需一个轻量级软件平台和现有电子教材即可运行；不强求教师精通数据分析，平台自动生成的可视化报告和策略建议即学即用；不改变现有的课程体系与教学计划，而是作为“插件式”的学情诊断工具嵌入日常教学。这种轻量化特征使其尤其适用于资源有限的医学院校，有助于缩

小校际“数字鸿沟”[23]。

8. 挑战与应对

任何教学平台的创新都必须正视其落地过程中的现实障碍。本平台面临的主要挑战包括技术门槛、数据伦理和认知依赖三个方面。

8.1. 技术门槛与操作适应

部分教师和学生可能不熟悉在阅读过程中随时记录提问的操作习惯，初期使用可能出现提问量少、质量参差不齐的问题。应对策略为：一是设计极简交互界面，提问框始终可见，一键提交，降低操作成本；二是提供 2~4 课时的培训，重点训练学生“将模糊的困惑转化为清晰的问题陈述”；三是设置“提问积分”等轻度激励机制，但需避免本末倒置——激励的是提问行为本身，而非提问数量，防止无效提问泛滥。

8.2. 数据隐私与伦理边界

学生的搜索记录和提问文本属于个人学习隐私数据，必须严格保护。本平台建立了明确的伦理框架：(1) 仅采集与课程学习直接相关的行为数据，不涉及个人身份敏感信息；(2) 数据分析报告以匿名化、群体化形式呈现给教师，学习档案仅对学生本人可见；(3) 数据存储加密，研究结束后按规定销毁；(4) 学生可随时选择退出数据采集而不影响其正常学习评价。

8.3. 认知依赖与思维自主性

若学生过度依赖平台的分析与建议，可能削弱其独立思考与自主判断能力。规避这一风险的关键在于定位清晰：AI 是“镜子”而非“教师”——它只呈现学生已存在的提问行为，而不代替学生得出结论。本平台在设计上刻意回避了直接提供标准答案的问答模式，就是为了防止学生将 AI 视为“答案提供者”，而始终将认知建构的主体地位留给学生自己。

9. 结语

本研究基于对传统医学教育痛点与 AI 技术应用空白的分析，自主研发了“智慧阅读助手”AI 医学学习平台，并构建了一种基于 AI 搜索数据分析的医学生个性化指导模式。该平台以学习分析的闭环框架为方法论基础，将智能导师系统的个性化诊断能力与学生提问研究的认知诊断价值相结合，形成了一条“提问采集–认知诊断–策略推荐–教学行动–效果追踪”的完整实践链路[24]。

在技术层面，平台采用基于 bert-base-chinese 预训练模型微调的多分类方案，实现提问类型的自动识别，在医学教育提问数据集上取得了 87.4%的整体准确率和 87.3%的 Macro F1，并通过决策阈值校准与低置信度回退机制保障了分类可靠性。在教学应用层面，平台的教学活动建议模块和教学策略库将学情诊断结果直接转化为可操作的教学行动方案，《诊断学》课程案例展示了从学情报告到课堂改进的完整闭环[25]。

学生在 AI 辅助环境下提出的问题，本质上是其内在认知过程的外化表征。本平台的核心创新在于将“提问方式”引入认知诊断框架，使抽象的认知层次转化为可观测、可分析的行为指标，将学生认知过程系统化、可视化与可操作化。与既有学习分析平台相比，本平台不仅提供数据报告，更通过教学活动建议模块和教学策略库，为教师提供了从诊断到行动的具体路径；与既有智能导师系统相比，本平台不以答题表现而以提问行为作为诊断数据源，拓展了认知诊断的信息基础，为医学教育从“经验驱动”走向“数据–经验融合驱动”提供了可操作的路径。

本平台并非试图用 AI 取代教师的专业判断，也不追求用技术解决所有教育问题。它的价值在于：让学生的学习困惑被看见、被理解、被回应。在知识膨胀的时代，教师的课堂时间永远有限，但通过 AI 的辅助，我们可以让有限的时间流向学生最迫切需要的方向。在医学教育日益强调以学生为中心、精准化培养的今天，本平台为从“标准化教学”走向“个性化指导”提供了一条切实可行的技术路径。未来的工作将聚焦于平台在多所医学院校的推广验证、基于序列建模的认知层次动态追踪，以及基于更大规模数据的学习者认知模型构建，持续推动医学教育的智能化、精准化发展。

基金项目

山西医科大学 2025 年度教育教学研究课题(XJ2025138)。

参考文献

[1] Ali, M. (2025) The Role of AI in Reshaping Medical Education: Opportunities and Challenges. *The Clinical Teacher*, **22**, e70040. https://doi.org/10.1111/tct.70040

[2] Alsolami, A.S. (2025) The Effectiveness of Using Artificial Intelligence in Improving Academic Skills of School-Aged Students with Mild Intellectual Disabilities in Saudi Arabia. *Research in Developmental Disabilities*, **156**, Article 104884. https://doi.org/10.1016/j.ridd.2024.104884

[3] Lee, S. and Bang, J. (2025) Personalized Learning through AI: From Rule-Based Systems to AI Mentors. *Multimedia-Assisted Language Learning*, **28**, 95-114.

[4] Zheng, K., Shen, Z., Chen, Z., Che, C. and Zhu, H. (2024) Application of AI-Empowered Scenario-Based Simulation Teaching Mode in Cardiovascular Disease Education. *BMC Medical Education*, **24**, Article No. 1003. https://doi.org/10.1186/s12909-024-05977-z

[5] Chin, C. and Brown, D.E. (2002) Student-Generated Questions: A Meaningful Aspect of Learning in Science. *International Journal of Science Education*, **24**, 521-549. https://doi.org/10.1080/09500690110095249

[6] Sailer, M., Ninaus, M., Huber, S.E., Bauer, E. and Greiff, S. (2024) The End Is the Beginning Is the End: The Closed-Loop Learning Analytics Framework. *Computers in Human Behavior*, **158**, Article 108305. https://doi.org/10.1016/j.chb.2024.108305

[7] Han, Y., Hong, S., Li, Z. and Lim, C. (2025) Defining and Classifying the Roles of Intelligent Learning Companion Systems: A Scoping Review of the Literature. *Tech Trends*, **69**, 567-581. https://doi.org/10.1007/s11528-025-01058-0

[8] Bergey, B.W., Cromley, J.G., Kaplan, A. and Bloxton, J.D. (2022) Do Students' Questions during Chemistry Lectures Predict Perceived Comprehension and Exam Performance? *The Journal of Experimental Education*, **91**, 411-430. https://doi.org/10.1080/00220973.2021.2021843

[9] Wong, A.H.P., Wong, L.W., Hooi, S.C. and Lee, S.S. (2026) From Inquiry to Insight: Unlocking the Potential of Students' Questions in Medical Education in Singapore. *Teaching and Learning in Medicine*, **38**, 186-195. https://doi.org/10.1080/10401334.2025.2481400

[10] Hao, Z., Jiang, J.X., Yu, J.F., *et al.* (2025) Student-AI Interaction in an LLM-Empowered Learning Environment: A Cluster Analysis of Engagement Profiles. arXiv: 2503.01694.

[11] Anderson, L.W. and Krathwohl, D.R. (2001) A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Longman.

[12] Lee, H., Atif, A. and Kang, K. (2026) Analysing AI Utilisation in Education through Learner Question Types: A Constructivist Approach. *Australasian Journal of Educational Technology*, **42**, 79-96. https://doi.org/10.14742/ajet.10657

[13] Devlin, J., Chang, M., Lee, K. and Toutanova, K. (2019) BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the* 2019 *Conference of the North American Chapter of the Association for Computational Linguistics*: *Human Language Technologies*, *Volume* 1 (*Long and Short Papers*), Minneapolis, 2-7 June 2019, 4171-4186. https://doi.org/10.18653/v1/n19-1423

[14] Cui, Y., Che, W., Liu, T., Qin, B. and Yang, Z. (2021) Pre-Training with Whole Word Masking for Chinese BERT. *IEEE*/*ACM Transactions on Audio*, *Speech*, *and Language Processing*, **29**, 3504-3514. https://doi.org/10.1109/taslp.2021.3124365

[15] Wang, Z.Y., Yang, H.W., Li, X.G., *et al*. (2025) Leveraging AI Technology for Coding the Classroom Observation Record Form of Flanders Interaction Analysis. *IEEE Access*, **13**, 66094-66108.

[16] Ouassilet, M.A., Jebbari, M., Rachidi, R., *et al*. (2026) A Voting Classifier Approach to Bloom's Taxonomy: Integrating BERT and Large Language Models for Questions Automatic Classification. 2026 6*th International Conference on Innovative*

Research in Applied Science, *Engineering and Technology* (*IRASET*), FEZ, 14-15 May 2026, 1-7.

[17] Almatrafi, O. and Johri, A. (2025) Leveraging Generative AI for Course Learning Outcome Categorization Using Bloom's Taxonomy. *Computers and Education*: *Artificial Intelligence*, **8**, Article 100404. https://doi.org/10.1016/j.caeai.2025.100404

[18] Lee, K.W., Ang, Y.S. and Lai, J.W. (2025) Learning Analytics with Scalable Bloom's Taxonomy Labeling of Socratic Chatbot Dialogues. *Computers*, **14**, Article 555. https://doi.org/10.3390/computers14120555

[19] Balubaid, A., Shahbal, L., Talib, M., Shahbal, A., Barajeh, L., Al-Hababi, K., *et al*. (2025) From Interrogative Detection to Difficulty Classification: A Two-Stage BERT-Based Approach for Educational Questions (Version v2). Zenodo.

[20] 王征海, 杨昌尧. 面向认知层级识别与知识诊断的智能学习系统设计[J]. 科技与创新, 2025(24): 8-11.

[21] Punyatoya, S., Vidya, N., Pradhan, C., Thejaswini, S., *et al*. (2025) VTU Question Classification System Based on BERT and Bloom's Taxonomy Ring. 2025 *International Conference on Emerging Systems and Intelligent Computing* (*ESIC*), Bhubaneswar, 8-9 February 2025, 774-777.

[22] Al Faraby, S., Adiwijaya, Romadhony, A., *et al*. (2022) Educational Question Classification with Pre-Trained Language Models. 2022 7*th International Conference on Informatics and Computing* (*ICIC*), Denpasar, 8-9 December 2022, 1-6.

[23] 杨艳, 邱佳宁, 罗静, 谢满云. 人工智能赋能眼科课程设计教学改革[J]. 科学咨询, 2024(10): 101-105.

[24] 郝一双. OBE 理念下使用微助教改进课堂教学的探索——以教师教育类理论课程为例[J]. 高等工程教育研究, 2021(S1): 24-25+46.

[25] Chen, P.H., Zeng, Y.X. and Lee, L.H. (2021) Incorporating Domain Knowledge into Language Transformers for Multi-Label Classification of Chinese Medical Questions. In: Lee, L.H., Chang, C.H. and Chen, K.Y., Eds., *Proceedings of the* 33*rd Conference on Computational Linguistics and Speech Processing* (*ROCLING* 2021), The Association for Computational Linguistics and Chinese Language Processing (ACLCLP), 265-270.