

DOI: 10.3969/j.issn.2096-045X.2026.04.006

· 教育教学 ·

AI驱动课后练习在系统解剖学重修教学中的实践与效果评估

贾宁 计胜峰 许杰华 孙钦儒*

(西安交通大学医学部, 西安 710061)

【摘要】目的 人工智能(artificial intelligence, AI)为自主学习提供了更便捷的途径、更精准的策略和更及时的反馈评价。本文旨在探究AI驱动课后练习在系统解剖学重修教学中的实践效果。**方法** 以系统解剖学课程的神经系统为例,采用历史对照研究设计,运用AI大模型(DeepSeek和Kimi)构建课后练习,运用金数据的快出题功能模块组成试卷,分析课后练习对自主学习的影响及对学习效果的作用。**结果** 97.83%的学生认为课后练习对知识理解和记忆有帮助;89.13%的学生认为课后练习对提高学习主动性有帮助;课后练习可以显著提高学生卷面得分,课后练习组学生的神经系统卷面得分为(9.97±6.99)分,显著高于对照组的(6.88±5.78)分($P<0.01$)。**结论** AI驱动课后练习有助于提高学生自主学习意愿及学习成绩。

【关键词】 人工智能; 系统解剖学; 自主学习; 课后练习; 教育理念; 教学改革

【中图分类号】 R4; G640

【文献标识码】 A

Practice and effect evaluation of AI-driven after-class exercises in the teaching of systematic anatomy retake

Jia Ning, Ji Shengfeng, Xu Jiehua, Sun Qinru*

(Xi'an Jiaotong University Health Science Center, Xi'an 710061, China)

【Abstract】Objective Artificial intelligence (AI) provides more convenient approaches, more accurate strategies, and more timely feedback and evaluation for self-learning. This study aims to explore practice and effectiveness evaluation of AI-driven after-class exercises in systematic anatomy retake teaching. **Methods** Adopting a historical controlled study design and taking the nervous system of the course of Systemic Anatomy as an example, using AI large models (DeepSeek and Kimi) to construct after-class exercises and to analyze the effectiveness of AI in empowering self-learning. **Results** The results showed that 97.83% of students believe that after-class exercises are helpful for understanding and memorizing knowledge, and 89.13% believe that after-class practices are helpful in improving learning initiative. Furthermore, using AI models to generate after-class exercises significantly improved students' test scores. The nervous system test score (9.97±6.99) points of the after-class exercises group was significantly more than that (6.66±5.78) points of the control group ($P<0.01$). **Conclusion** AI-driven after-school exercises can help improve students' willingness to learn independently and academic performance.

【Keywords】 artificial intelligence; systematic anatomy; self-learning; after-class exercise; educational philosophy; teaching reform

基金项目 西安交通大学2024年生成式人工智能赋能课程教学改革(“智课”)专题研究项目:人工智能赋能法医心理学教学研究(24ZK35Y)

第一作者 贾宁, 博士, 副教授, 研究方向: 基础医学教育与教学研究。Email: jianing@mail.xjtu.edu.cn

***通信作者** 孙钦儒, 博士, 副教授, 研究方向: 法医学、基础医学教育与教学研究。

Email: qinrusun@mail.xjtu.edu.cn

人工智能(artificial intelligence, AI)技术在教育领域的早期应用可以追溯至20世纪60年代,其中由伊利诺伊大学香槟分校开发的一款智能教学系统,称为自动化教学操作程序逻辑系统(programmed logic for automated teaching operations, PLATO),它允许学生通过图形用户界面与教育材料互动,最早实现了学生个性化在线学习^[1]。2018年,OpenAI发布首个GPT模型,标志着自然语言处理(NLP)从任务特定模型向通用预训练模型的转变,开启新一轮的信息革命^[2]。2019年2月,中共中央、国务院发布《中国教育现代化2035》^[3],提出利用现代技术加快推动人才培养模式改革,实现规模化教育与个性化培养的有机结合。AI已经开始深入影响到教育理念、文化和生态,以DeepSeek为代表的AI大模型在资源获取、课件生成、智能反馈与评价等方面已显现出强大的优势^[4-5]。

近年来,各级医学院校针对知识相对繁多、枯燥,学生自主学习的意识不强等问题,都在积极开展教育教学改革,提出一系列对策和措施,以期提高学生的自主学习能力,提供自主学习的环境^[6]。各种在线自适应学习系统可根据学习者掌握情况实时调整内容推送与学习进度,获得医学生的良好反馈^[7];哈佛大学使用AI驱动的学习分析系统实时监测医学生学习进度,并自动推荐补充个性化学习材料^[8]。

系统解剖学是医学生的第一门基础医学课程,学生普遍反映知识点多、概念抽象,在短期内无法形成有效记忆。据笔者统计的2022年秋—2024年秋西安交通大学系统解剖学初修学生期末试卷得分,不及格率分别为40.05%、32.91%、46.00%。分析这些学生不及格的可能原因,主要是系统解剖学专业词汇多,知识结构复杂,理解困难,无法形成有效记忆。在以语言、数字等形式的学习记忆中,信息的初次接触会以“刺激-反应”的感觉形式储存起来,但只会构成瞬时记忆。要形成长期记忆,必须经过多次重复,即信息的接触次数越多,越容易被识别出来;一个语言单位越频繁使用,越易于强化和习得^[9]。在此过程中,需要学生主动重复信息,再现信息,问题和答

案的多次联合刺激以帮助学生调取所学的知识,加深理解和记忆。这无疑为教学提出新的要求:如何让学生愿意主动学习,多次从大脑中提取信息,形成长期记忆:是教学改革的重点方向。因此,多次阶段考试的过程性评价成为目前教学改革的重要措施,但是多次命题和组卷、组织考试和改卷无疑严重增加了教师的教学工作量。而且在一个学期内多次组织考试,在空间和时间上也是难以达成的。AI大模型和技术工具的出现恰恰是解决这一问题的利器。AI大模型(DeepSeek和Kimi)擅长处理文本、提取信息、分析资料。教师可以通过对话形式将书本上的一段文字、PPT课件等上传给DeepSeek和Kimi,快速得到限定条件下的题目、题型和答案及解释。金数据是国内一款AI智能组卷工具,支持多格式文档解析、智能题库生成和全流程在线考试管理。金数据可以帮助教师快速组卷、发布、统计和分析结果。学生可以点击教师发布的链接,随时登录进行在线答题,参与“考试”的形式不拘泥于考场,不需要在固定时间,可以随时随地进行,提交后能快速直观地发现自己的错题,而且经过多轮答题,在有限的时间内纠正错误、强化记忆。这些AI大模型和工具不仅可以提高教师工作效率,轻松实现教学过程中的以考促学,而且可以极大地满足个性化需求,有利于提高学习兴趣和主动性。

综上所述,本研究为了提高学生学习的主动性、便捷性,提高系统解剖学知识点的高频再现,提高系统解剖学教学质量和效率,以《系统解剖学神经系统》章节为例,首先在初修成绩不及格的暑期重修生中,使用DeepSeek、Kimi生成试题,利用金数据组建课后练习,经过后台数据分析、问卷调查和考试得分,探索AI驱动系统解剖学自主学习的智能路径和方法,推进高等医学教育高质量发展。

1 对象与方法

1.1 研究对象

选取2024年和2025年暑期系统解剖学重修

班136名学生作为研究对象,课后练习和问卷调查的参与者自愿加入,以金数据后台数据为参与课后练习依据;问卷星数据为问卷调查依据。初修成绩或重修成绩有缺项的学生在后期分析成绩时剔除。以2025年暑期重修班(以下简称暑期班)学生作为观察对象,2024年暑期班学生、2025年暑期班未提交课后练习的学生作为对照组。

1.2 研究方法

本研究为历史对照教学干预研究。使用AI大模型(DeepSeek和Kimi)构建课后练习的试题,利用金数据快出题,分别针对“中枢神经(1)”“中枢神经(2)”“周围神经”“传导路、脑和脊髓被膜和血管”4个单元形成4份课后练习的试卷。根据学生课后练习的完成情况,将2025年暑期班学生再分为未提交课后练习组、完成课后练习1个或2个单元组、完成3个及以上单元组。通过问卷星平台调查2025年暑期班的学生,评价课后练习的教学效果。

1.2.1 题库构建和试卷生成

将教师授课使用的PPT课件分别上传给DeepSeek和Kimi,使用对话“根据文件内容生成20~25个单项选择题,备选答案为5个,语言为中文,并给出正确答案和解释”,教师对题目进行筛选,最后针对“中枢神经(1)”“中枢神经(2)”“周围神经”“传导路、脑和脊髓被膜和血管”4个单元整理成题库。按照金数据快出题的要求上传题库,生成4个单元的课后练习试卷。在每次课后,发布链接给学生,要求学生在结课考试前提交结果,可多次提交,不限次数。

1.2.2 问卷星题目设计

在结课考试后,2025年暑期班的学生完成问卷星调查。题目包括对AI的了解和使用情况、课后练习对课程知识理解的帮助、课后练习是否提高学习主动性,以及学生对课后练习的接受程度。

1.2.3 结课考试中神经系统得分比较

统计2024年和2025年各组学生卷面的神经系统部分得分,以均数 $\pm$ 标准差($\bar{x}\pm s$)表示;统计2024年和2025年卷面神经系统得分 ≥ 18.6 分(神

经系统总分为31分,18.6分占60%,定义为及格)的人数。

1.3 统计学方法

采用SPSS 27.0统计软件进行数据处理。使用Shapiro-Wilk检验进行正态性检验,符合正态分布的计量资料以均数 $\pm$ 标准差($\bar{x}\pm s$)表示,并进行方差齐性检验后进行单因素方差分析(One-way ANOVA)和LSD多重比较。对于2024年和2025年的卷面成绩,进行协方差分析;对卷面神经系统得分 ≥ 18.6 分的人数进行卡方检验或校正卡方检验。以 $P<0.05$ 为差异有统计学意义。

2 结果

2.1 学生基本情况

2025年暑期班有65名学生参加考试,其中护理学专业5名、法医学专业9名、临床医学(包括口腔)专业35名、基础医学专业5名、预防医学专业2名、药学及相关专业9名。初修成绩(26.65 ± 8.75)分,重修成绩(55.58 ± 18.85)分。其中,有46名2025年暑期班学生完成问卷星调查。

2024年暑期班有71名学生参加考试,其中护理学专业3名、法医学专业7名、临床医学(包括口腔)专业17名、基础医学专业8名、预防医学专业12名、药学及相关专业24名。2024年暑期班初修成绩是(26.87 ± 8.47)分,重修成绩是(40.26 ± 15.41)分。

2024年和2025年暑期班两组学生的年龄、性别差异均无统计学意义($P>0.05$)(表1)。初修卷面总分成绩之间无显著性差异,重修成绩之间有显著性差异($P<0.001$)(表2)。

表1 2024年和2025年暑期班学生基本情况

组别	性别(男/女)[n(%)]	年龄 ($\bar{x}\pm s$,岁)
2024年暑期班($n=71$)	44(61.97)/27(38.03)	20.25 $\pm$ 0.86
2025年暑期班($n=65$)	42(64.62)/23(35.38)	20.45 $\pm$ 0.95
χ^2 值/ t 值	0.102	-1.241
P 值	0.749	0.217

表2 2024年和2025年暑期班学生总成绩比较($\bar{x}\pm s$,分)

组别	初修成绩	重修成绩
2024年暑期班($n=71$)	26.87 $\pm$ 8.47	40.26 $\pm$ 15.41
2025年暑期班($n=65$)	26.65 $\pm$ 8.75	55.58 $\pm$ 18.85
t 值	0.159	-5.206
P 值	0.874	<0.001

2.2 学生使用AI的情况

问卷数据显示,在2025年暑期班的65名学生中有46名学生参与问卷调查,其中44人(95.65%)在平时学习或完成作业时会使用AI工具,以DeepSeek使用频次最高(60.87%),其次是豆包(15.22%)和Kimi(10.87%),2人不使用AI工具(4.35%),4人使用其他AI工具(8.70%)(图1)。

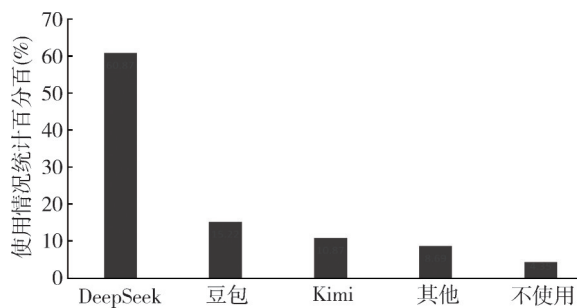


图1 学生使用AI情况

2.3 课后练习对学习和得分的影响

为分析初修成绩对重修成绩的影响,使用初修成绩作为协变量,结果显示,初修成绩与组别之间无交互作用($P=0.845$)。

问卷调查($n=46$)结果显示,有45人(97.83%)认为课后练习对知识理解和记忆有帮助,1人认为没有帮助。2025年暑期班学生参加结课考试,神经系统卷面总分为31分,学生平均得分(9.97 ± 6.99)分;2024年暑期班学生参加结课考试,神经系统卷面满分为31分,学生平均得分(6.66 ± 4.98)分,差异有统计学意义($P<0.01$),即2025年暑期班学生的神经系统得分高于2024年暑期班学生。2025年提交课后练习的49名学生,平均得分(10.59 ± 7.37)分;2025年未提交课后练习的16名学生,平均得分(6.88 ± 5.78)分,差

异无统计学意义($P=0.071$)。

2025年暑期班学生中,未提交课后练习的16名学生,神经系统得分均 <18.6 ;有49人提交课后练习,其中有6人神经系统得分 ≥ 18.6 ;2024年暑期班学生有1人神经系统得分 ≥ 18.6 。将2025年提交课后练习的学生与2024年学生的神经系统得分 ≥ 18.6 的人数,进行连续校正卡方检验,结果显示2025年提交课后练习学生的神经系统得分 ≥ 18.6 的人数显著多于2024年,差异有统计学意义($\chi^2=4.38, P<0.05$)。

2.4 学生课后练习的完成度对卷面得分的影响

2025年暑期班学生参加结课考试,49人(75.38%)至少完成了1个单元的课后练习,其中有20人仅完成1个单元,有12人完成了2个单元,4人完成了3个单元,13人完成了4个单元所有课后练习;16人(24.62%)未在平台提交课后练习结果,视为未参加课后练习。

完成3个及以上单元课后练习的17名学生神经系统的平均分为(12.59 ± 7.25)分,完成1个或2个单元的32名学生神经系统的平均分为(9.53 ± 7.31)分,未提交课后练习的16名学生神经系统的平均分为(6.88 ± 5.78)分。统计学分析显示,完成3个及以上单元课后练习的学生神经系统得分高于未提交组学生($P<0.05$);完成1个或2个单元组学生神经系统得分的均值虽然高于未提交组,但差异无统计学意义($P=0.21$),完成3个及以上单元组与完成1个或2个单元组之间差异亦无统计学意义($P=0.17$)。

2.5 课后练习对提高学生学习主动性的影响

调查问卷($n=46$)结果显示,有41人(89.13%)认为课后练习对提高学习主动性有帮助,5人(10.87%)认为对提高学习主动性的作用一般,没有人认为对提高学习主动性没有帮助(图2)。

2.6 学生对课后练习的接受程度

调查问卷($n=46$)结果显示,有35人(76.09%)表示喜欢课后练习,10人(21.74%)对课后练习一般喜欢,1人(2.17%)表示不喜欢课后练习(图3)。

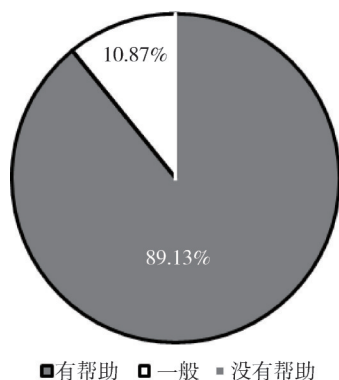


图2 课后练习对提高学习主动性影响的调查结果 (n=46)

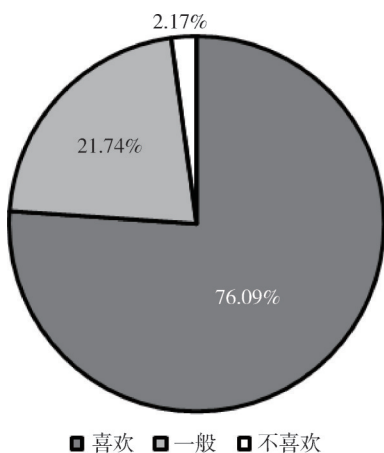


图3 学生对课后练习的接受程度的调查结果 (n=46)

3 讨论

近年来,随着AI技术的迅猛发展,“AI+X”正在更深层次地重塑多个领域的高质量发展。承载高等教育的高校作为科技创新与高素质人才培养的重要阵地,融合人工智能技术,促进科技创新、变革教育范式、提高人才培养质量势在必行。AI大模型从“感知时代”已发展到具有推理能力的“生成式阶段”,而且随着技术“周级迭代”的爆发态势,从通用基础模型衍生出各类专业模型。高校教师和学生应该积极学习、运用AI工具。已有研究证明,AI可以模拟医学环境,提供个性化学习,借助机器学习算法提高学习效率^[10];将人工智能技术应用到高校医学专业的教与学中,能有效提升学生学习兴趣及学习

成绩^[11]。

3.1 AI已成为日常学习工具

本研究中,参与问卷调查的46名学生中,44人(95.65%)均会在日常学习中使用AI工具,其中DeepSeek的使用人数最多,豆包位居第二。根据全球最大AI模型API(应用程序编程接口)聚合平台OpenRouter最新数据测算,全球AI大模型Token调用量排名前三中,DeepSeek-V4-Flash位居第一^[12]。自AI大模型问世以来,DeepSeek模型展现出与国际顶级模型比肩的实力,而且在日常对话、文本生成等高频需求上,用户面广且使用门槛低,受到国内广大用户的青睐。本次调查结果显示,超过60%的学生使用过DeepSeek,说明大多数学生熟悉AI,会使用AI工具。

3.2 课后练习显著提高卷面得分

自2022年秋季以来,系统解剖学考试的卷面成绩一直存在超过30%的不及格率,笔者所在学院针对此种局面,在2024年和2025年暑期小学期举办重修班,在结课后一周内考试,取得了较好的结果。但是在统计成绩中发现,2024年试卷的神经系统平均得分为6.66,仅占卷面31分的21.48%,远低于预期。分析学生神经系统得分低的原因,可能包括以下几点:①神经系统是人体解剖学课程中最具挑战性的部分^[13];②神经系统是课本最后的部分,相对来讲,不管是学习还是最后复习,到最后难免会缺乏耐心、注意力涣散;③神经系统概念多,而且大多神经结构无法肉眼观察,神经通路需要一定的想象力,甚至需要在脑中构建三维立体结构;④从开始学习到考试的时间较短,在多而难的知识冲击下,极易打击学习积极性,最终导致学生放弃神经系统的学习,总成绩中神经系统得分偏低,造成学业成绩偏低或者不及格。因此,教师在重修班的授课中需要适当地调整节奏,突出重点,攻坚克难,而且需要提高学生的兴趣,激发学生自主学习的主动性。因此,在2025年,我们利用DeepSeek、Kimi和金数据,将神经系统的内容生成课后练习试卷,在课后发布链接,督促学生学习。其优

点是学生可以随时、多次登录答题,提交后自评,以考促练、查漏补缺。数智时代已来,高等医学教育变革面临机遇和挑战,其中提升医学院校教师的数字素养是重中之重^[14]。作为高校教师,必须加速转变,提升数字素养,借助人工智能分析工具,帮助学生快速获取有价值的信息,利用学习分析系统,实时跟踪学生学习情况,及时调整教学策略,为学生提供个性化辅导^[15]。这些学习困难的学生尤其需要课后多练习,而AI赋能课后练习为其个性化的学习提供了可能。

本研究结果显示,在结课考试的卷面成绩方面,2025年暑期班学生神经系统平均得分9.67,显著高于2024年的6.66;从神经系统卷面分 ≥ 18.6 的人数来看,2025年组显著多于2024年组。在调查中发现,2025年暑期班中有16人未提交课后练习,因此在分析2025年成绩时,针对提交练习数量不同进行了组间比较。结果显示,2025年暑期班的学生中,参与课后练习的学生神经系统平均得分更高,为10.59;未提交课后练习的学生中,没有1人的神经系统卷面分 ≥ 18.6 。以上结果说明,课后练习有利于学生对知识的理解和记忆,对提高学生的考试成绩是有帮助的。

3.3 课后练习显著提高学生学习主动性

问卷调查结果显示,有97.83%的人认为课后练习对知识理解和记忆有帮助,89.13%的人认为课后练习对提高学习主动性有帮助,76.09%的人表示喜欢课后练习。结果说明,大部分学生认为课后练习对学习有帮助,喜欢这种形式。这说明课后练习的在线形式,从客观上延伸了学习的时间和场所,实现了自主性和个性化学习,得到多数学生的认可和喜爱。

3.4 研究局限性

本研究存在一定局限性:①研究对象仅有136人,参与问卷的学生46人,样本量偏小,可能对结果造成一定的偏倚,后期还需要更大的样本量和多中心合作夯实研究数据,支持研究结果的可重复性;②本研究的课后练习仅覆盖神经系统,后期可考虑课后练习覆盖更多章节,亦可以比较单一章节干预与多章节干预对学生卷面成

绩影响的差异,使研究结果具有更大的泛化性和推广性。

4 结 语

本研究针对的是系统解剖学课程重修的学生,这部分学生经常在学习中存在不同形式和不同程度的学习困难,教师必须在教学中重视这部分学生,而且AI大模型在自主学习中有天然优势。本研究结果显示,通过AI赋能显著提高学生系统解剖学自主学习主动性、显著提高学习成绩。下一步我们将在课程初修阶段,针对更多的章节,开展课后练习,满足个性化自主学习需求。

在AI工具的使用过程中,单一大模型容易产生不正确或非专业表述,另外还需要教师仔细检查、谨慎选题,才能严控试题质量。因此,本研究使用DeepSeek和Kimi两个AI大模型,提高对文本提取的准确率和效率,降低错误率,增加教师筛选题目的范围,提高试题质量。

AI在高等医学专业教学中具有广阔的应用前景,随着其不断地深度融合于医学教学,必将为实现医学教学的现代化和信息化发展插上翅膀,更精准地赋能医学卓越人才的培养,推进高等医学教育高质量发展。

利益冲突 所有作者声明不存在利益冲突。

作者贡献声明 贾宁:问卷调查,数据分析,论文撰写;计胜峰、许杰华:数据整理;孙钦儒:数据分析,论文修改。

参考文献

- [1] 罗玮琪. 人工智能在教育领域中的应用[J]. 汉字文化, 2025(8): 184-186.
- [2] 陈凌云, 姚宽达, 王茜, 等. ChatGPT: 研究进展、模型创新及医学信息研究应用场景优化[J]. 医学信息学杂志, 2023, 44(7): 18-23, 29.
- [3] 新华社. 中共中央、国务院印发《中国教育现代化2035》[EB/OL]. (2019-02-23)[2026-01-30]. https://www.gov.cn/xinwen/2019-02/23/content_5367987.htm.
- [4] 王茹, 李毅恒, 孙世仁, 等. 人工智能在医学教育

- 中的应用前景与挑战[J]. 中国医学教育技术, 2025, 39(3): 306-310.
- [5] 蒋小英, 宁启兰, 闫小飞, 等. 生成式人工智能在基础医学教学中的应用前景[J]. 基础医学教育, 2024, 26(6): 508-511.
- [6] 郭建如, 王维民. 新发展阶段我国高等医学教育改革探析[J]. 国家教育行政学院学报, 2022(7): 19-26, 47.
- [7] Tan C, Han R, Ye R, et al. Adaptive learning recommendation strategy based on deep Q-learning [J]. Appl Psychol Meas, 2020, 44(4): 251-266.
- [8] 邢园园, 钱玲. 美国一流大学应对人工智能教学应用的改革行动与反思[J]. 开放教育研究, 2025, 31(2): 24-35.
- [9] 李适. 学生从形式记忆到形象记忆的转向[J]. 教育理论与实践, 2024, 44(35): 38-41.
- [10] 余梦雪, 蔡昕昱, 王刚, 等. 人工智能在医学研究生教学中的应用现状与分析[J]. 中国继续医学教育, 2025, 17(5): 170-174.
- [11] 黄智若, 沈晓沛, 陈湖星, 等. “医学+人工智能”复合型创新人才培养模式探析[J]. 中国医学教育技术, 2024, 38(3): 271-275.
- [12] 宋欣悦. 中国 AI 大模型 Token 调用量持续领跑 [N]. 每日经济新闻, 2026-05-27(7).
- [13] 冯改丰, 周劲松, 靳辉, 等. 3D 打印在中枢神经系统解剖学教学中的应用研究[J]. 医学教育研究与实践, 2024, 32(1): 79-84.
- [14] 汪丽聪, 王宇, 王博, 等. 人工智能赋能医学院校师资建设的研究[J]. 中国医学教育技术, 2025, 39(4): 450-454.
- [15] 范建丽, 张新平. 大数据+智能时代的教师数智胜任力模型研究[J]. 远程教育杂志, 2022, 40(4): 65-74.
- (收稿日期: 2026-01-30, 修回日期: 2026-02-19)
(本文编辑: 闫红)
- 开放获取** 本文使用遵循知识共享 署名-非商业性-禁止演绎 4.0 协议(CC BY-NC-ND 4.0), 详细信息请访问 <https://creativecommons.org/licenses/by-nc-nd/4.0/>。
- OPEN ACCESS** This article is licensed for use under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Deed (CC BY-NC-ND 4.0). For more information, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>.