

元认知视域下 LLM 与 RAG 技术赋能 教学模式创新的路径探析

蒋帆

浙江同济科技职业学院

摘要：本研究以元认知理论为基石，深入剖析大型语言模型（LLM）与检索增强生成（RAG）技术在教学模式创新中的作用机制、应用路径及实践成效。研究发现，基于 LLM 和 RAG 技术的个性化学习支持系统，能够通过动态反馈和反思性任务设计，显著提升学生的元认知监控与调节能力；借助智能化教学内容设计与动态化教学过程调整，助力教师实现教学策略的迭代升级；凭借强大的资源整合能力与精准推送机制，推动教育资源的高效利用。然而，技术应用过程中面临着“幻觉”现象、检索效率瓶颈、高成本投入等技术难题，教师角色转型困境、学生学习习惯重塑挑战、评价体系滞后等教育教学问题，以及数据隐私泄露风险、学术诚信危机、技术滥用隐患等伦理安全困境。为此，本研究从技术优化创新、教育教学改革与教师培训、完善伦理规范与安全保障三个维度提出应对策略，为人工智能时代的教育模式转型提供实践方案。

关键词：元认知；大型语言模型（LLM）；检索增强生成（RAG）；教学模式创新；教育技术应用

引言

在人工智能技术迅猛发展的时代浪潮中，大型语言模型（Large Language Model, LLM）和检索增强生成（Retrieval-Augmented Generation, RAG）技术作为人工智能领域的重要突破，正深刻重塑教育教学的生态格局。元认知作为个体对自身认知过程的监控、调节与反思，在教育教学中发挥着关键作用。培养学生的元认知能力，有助于提升其学习自主性、批判性思维与问题解决能力。从元认知视域出发探究 LLM 与 RAG 技术赋能教学模式创新，旨在深入挖掘这些技术在促进学生认知管理、优化知识运用、推动教学模式变革等方面的潜力。本研究聚焦职业教育机电专业教学现状，通过设计基于 LLM 和 RAG 技术的学习工具，为教师构建促进学生元认知发展的教学模式提供切实可行的实现路径。

一、核心概念及理论基础

（一）元认知在教育学的积极作用

元认知概念由美国心理学家弗拉维尔（J.H.Flavell）于 20 世纪 70 年代正式提出，为认知心理学开辟了全新的研究方向。弗拉维尔指出，元认知是“对认知的认知”，涵盖个体对自身认知过程、结果及相关活动

的认识、监控与调节，强调个体对认知活动的觉察、理解与管理。

元认知包含元认知知识、元认知体验和元认知监控三个核心要素^[1]。在教育领域，众多研究表明，元认知能力对学生学习成效具有显著影响。善于运用元认知策略的学生，往往在学业表现上更具优势，能够更高效地理解、记忆和提取知识。例如，王晓平^[2]等认为元认知知识是个体关于认知主体、认知任务和认知策略等方面的知识，学生若了解自身擅长形象思维，在学习科学、社会科学数据时，便会倾向于借助图表、思维导图等工具辅助理解。张羽^[3]等强调元认知监控是元认知的核心，学生通过监控学习进度、评估学习效果，及时调整学习策略，如在记忆困难时运用思维导图梳理知识体系，从而加深理解与记忆。

教师的元认知能力同样对教学质量起着关键作用。具备较强元认知能力的教师能够在教学过程中不断反思教学行为，通过观察学生反应、课堂提问、作业批改等方式监控教学效果，及时发现问题并调整教学方法，以提升教学质量。

（二）LLM 在教育领域的应用现状

LLM 基于深度学习技术，其核心架构 Transformer 摒

基金项目：浙江省高职教育“十四五”第二批教学改革项目：专业群深度对接省域数字化数字精密制造产业链人才培养模式的研究与实践（jg20240427）；2023年度浙江省产学研合作协同育人项目“机电专业人才“双元制”培养模式的探索与实施”（浙教办函〔2023〕241号，序号 433）；2025年校级教育科学研究项目：基于 OBE 理论与能力图谱的智能电气专业群课程体系构建研究。

作者简介：蒋帆（1984—），女，硕士，副教授，研究方向为高等教育，机电一体化技术。

弃了传统循环神经网络 (Recurrent Neural Network, RNN) 和卷积神经网络 (Convolutional Neural Network, CNN) 的序列处理方式, 引入自注意力机制 (Self-Attention) 和位置编码 (Positional Encoding), 显著提升了模型对长序列数据的处理能力与效率^[4]。自注意力机制使模型在处理词汇时, 能够关注输入序列中所有词汇的重要性, 从而精准捕捉语义关系。基于 Transformer 架构, 研究人员通过不断扩充模型参数规模与训练数据量, 开发出一系列功能强大的 LLM。以 GPT 系列为例, 2023 年推出的 GPT-4 在多模态能力、逻辑推理能力等方面实现重大突破, 不仅能够处理文本信息, 还能理解图像等多模态内容, 在复杂逻辑推理任务和专业领域测试中表现出色, 为其在教育领域的广泛应用奠定了坚实基础。

在教育实践中, LLM 已在多个领域发挥重要作用。在教学辅助方面, LLM 能够为教师提供丰富的教学资源与多样化的教学工具, 助力教师快速生成教案、课件及教学案例, 提升备课效率与质量, 同时协助设计教学活动, 激发学生学习兴趣, 促进课堂互动。在智能评测领域, LLM 可实现学生作业与考试的自动批改与评估, 依据预设评分标准给出分数, 并提供详细反馈与改进建议。在个性化学习方面, LLM 能够根据学生学习情况与特点, 提供个性化学习建议与指导, 充当学生的专属学习伙伴, 满足个性化学习需求, 提升学习效果。

(三) RAG 在教育领域的潜在价值

RAG 技术是信息检索与生成式技术的创新性融合, 旨在提升人工智能系统处理自然语言任务时的回答准确性与信息可靠性。该技术具有显著优势, 能够有效缓解 LLM 在生成内容时的“幻觉”问题^[5]。由于 LLM 的知识来源于预训练数据, 在处理复杂、新颖或需要实时信息的问题时, 可能生成与事实不符的内容。而 RAG 技术通过引入实时检索机制, 从可靠知识库中获取准确信息, 为生成模型提供事实依据, 降低“幻觉”发生概率。例如, 在回答历史事件相关问题时, RAG 技术可从权威历史文献数据库中检索资料, 确保答案真实可靠。

在教育资源获取与整合方面, RAG 技术具有重要价值。教育领域拥有海量学习资源, RAG 技术能够帮助师生快速、准确地检索所需信息, 并整合为有用的学习资料。对于教师而言, RAG 技术可协助收集教学素材, 丰富教学内容; 对于学生, RAG 技术与智能教学系统结合, 能够实现个性化学习, 在学生遇到问题时迅速检索知识, 提供解答与学习建议。

二、元认知视域下 LLM 与 RAG 技术赋能教学模式创新

(一) 促进学生元认知能力发展

1. 提供个性化学习支持

元认知视域下, LLM 与 RAG 技术通过深度分析学生学习数据, 为学生提供精准的个性化学习支持, 助力学生明晰学习状况, 明确学习目标, 加强自我监控与调节。以超星学习通在线学习平台为例, 该平台运用 LLM 与 RAG 技术, 实时采集和分析学生学习时间、答题情况、课程完成进度、提问内容等多源数据, 精准把握学生学习状态与需求, 据此生成个性化学习建议, 如推荐特定知识点讲解视频、针对性练习题, 并在答题过程中实时提供提示与指导。

2. 引导学生自我反思与监控

借助 LLM 与 RAG 技术构建的智能学习环境, 能够有效引导学生反思学习过程, 培养自我监控能力。学生完成学习任务或解答问题后, LLM 生成引导性问题, 帮助学生反思学习过程, RAG 技术检索类似题目不同解法与相关知识点拓展资料, 拓宽学生思维视野。在线课程平台利用 LLM 与 RAG 技术生成学习过程可视化分析报告, 展示学生学习进度、参与讨论情况、作业完成质量及与同学对比分析等信息, 使学生直观了解学习表现, 发现优势与不足, 主动反思并调整学习策略, 提升元认知能力。

(二) 助力教师教学策略优化

1. 教学内容设计与组织

LLM 与 RAG 技术为教师设计教学内容提供强大支持, 打破传统教材局限, 使教学内容更具深度与广度。基于 LLM 与 RAG 技术生成的教学内容, 融入前沿学术成果与丰富细节, 教师还可利用 LLM 生成多样化教学素材, 如知识点动画脚本、教学案例, 通过角色扮演等活动帮助学生深入理解知识。

2. 教学过程的动态调整

LLM 与 RAG 技术能够实时收集和分析学生学习数据, 为教师提供全面教学反馈, 助力教师动态调整教学策略。在课堂教学中, 教师借助智能教学平台, 利用 LLM 与 RAG 技术监测学生课堂参与度、注意力集中程度、知识点理解程度等, 根据反馈信息生成个性化辅导方案, 检索强化练习材料与拓展学习资源。在小组合作学习中, LLM 与 RAG 技术能够帮助教师监测小组讨论情况, 及时干预指导, 促进小组合作学习顺利进行, 提升学生团队协作能力。

（三）推动教育资源的整合与利用

1. 构建丰富的教育知识库

RAG 技术凭借强大的信息检索与整合能力，汇聚各类教育资源，构建全面丰富的教育知识库。这些资源涵盖教材、学术论文、教学视频等多种类型，经过系统整理、分类与标注索引，为精准检索与利用提供便利。现有的教学资源平台利用 RAG 技术整合多方资源，构建专业知识图谱，为师生提供丰富学习资料，满足个性化学习需求。

2. 实现资源的精准推送

基于师生行为数据、学习偏好与教学需求等多源信息，运用机器学习算法与数据分析技术，实现教育资源的精准推送。通过分析学生学习历史数据，了解学习进度、知识掌握情况与学习难点，推送适配学习资源；根据教师教学计划、课程内容与教学风格，推送匹配教学资源与工具，提高资源利用率与教学效率，促进教育教学个性化发展。

三、元认知视域下 LLM 与 RAG 技术赋能教学模式创新面临的挑战与应对策略

（一）面临的挑战

1. 技术层面的问题

尽管 LLM 在自然语言处理领域取得显著成就，但其生成内容时的“幻觉”现象仍亟待解决。如斯坦福研究人员，用形成性试卷来验证 ChatGPT 的呈现效果时，发现其准确率只有 58.4%^[6]。又比如历史教学中，LLM 可能生成错误的历史时间或虚构情节，误导学生学习^[7]。RAG 技术在检索效率与准确性方面面临挑战，随着教育资源增长，庞大的知识库对 RAG 技术检索能力提出更高要求，可能出现检索缓慢、结果相关性不强等问题。此外，LLM 与 RAG 技术应用成本较高，研发与维护需要大量计算资源和专业人才，技术更新换代快，增加了教育机构的应用负担。

2. 教育教学层面的问题

LLM 与 RAG 技术的应用对教师角色与能力提出新挑战。传统教学中教师是知识传授者，而在新技术教学模式下，教师需转变为学习引导者与促进者，要求教师具备良好的信息技术素养。然而，目前许多教师对新兴技术掌握不足，缺乏培训与实践经验，影响技术优势发挥。学生长期形成的被动学习习惯也阻碍了新技术教学模式的创新，部分学生过度依赖技术答案，缺乏独立思考与批判性思维。现有的教育评价体系以考试成绩为主，难以评估学生在新技术教学模式下的学习过程、元认知能力与综合素养，导致教学与评价脱节。

3. 伦理与安全层面的问题

LLM 与 RAG 技术应用涉及大量师生数据收集、存储与分析，存在数据隐私泄漏风险^[8]。教育平台若技术漏洞或管理不善，可能导致学生个人信息被非法获取利用。学术诚信问题也不容忽视，LLM 强大的文本生成能力增加了学生作弊风险，教师难以判断作业与论文原创性，破坏教育公平与学术环境。此外，技术滥用可能传播虚假信息、不良价值观，误导学生思想与行为，影响社会稳定。

（二）应对策略

1. 技术优化与创新

为解决 LLM “幻觉”问题与 RAG 技术检索瓶颈，降低技术应用成本，可从多维度进行优化。针对 LLM “幻觉”，在数据层面强化筛选清洗，增加数据多样性；在模型层面改进架构与训练算法，引入事实性验证机制。对于 RAG 技术检索问题，优化检索算法，采用结合语义理解与深度学习的改进版 DPR 算法，构建分布式索引结构，探索量子计算应用。在成本控制方面，教育机构与云服务商合作采用弹性云计算，运用模型压缩技术，加强内部技术团队建设，参与开源社区，降低研发成本。

2. 教育教学改革与教师培训

教育机构应以技术为驱动，推动教学改革。创新教学模式，推广项目式学习、问题导向学习，设计开放性任务，培养学生自主学习与创新思维。构建基于技术的学习社区，为学生提供交流平台，教师发挥引导监督作用，邀请行业专家分享经验。加强教师培训，系统开展 LLM 与 RAG 技术原理、应用场景与操作方法培训，采用线上线下融合模式，邀请专家与一线教师分享案例，鼓励教师实践应用，提升教学能力。

3. 伦理规范与安全保障措施

制定完善的伦理规范与安全机制是保障技术合理应用的关键。教育机构与相关部门联合制定伦理准则，明确技术应用边界与责任，规范数据隐私保护、学术诚信管理与技术使用透明度。健全数据安全保障机制，采用加密技术、权限管理、数据备份与安全监测系统，确保数据安全。防范学术诚信问题，开发检测工具，加强学术道德教育，构建监督机制，严肃处理违规行为，维护学术环境公正。

四、结论与展望

本研究基于元认知理论，系统探讨了 LLM 与 RAG 技术赋能教学模式创新的作用机制、路径与成效。研究表明，LLM 与 RAG 技术从促进学生元认知能力发展、助力教师教学策略优化、推动教育资源整合利

用等多个维度,为教学模式创新提供了强大动力。

展望未来,LLM与RAG技术在教育领域前景广阔。在技术层面,有望通过算法优化进一步降低LLM“幻觉”发生率,大幅提升RAG检索与生成效率,并结合多模态技术,实现图文音视频融合,为学生带来更加丰富、多元的学习体验。在教学模式创新方面,个性化、自适应与混合式教学将得到进一步深化,推动因材施教的教育理想向更高水平发展。教育评价体系也将顺应技术发展趋势,向注重学习过程、强调元认知能力评估的方向不断完善。随着技术与教育的深度融合,LLM与RAG技术将持续驱动教学变革,为创新型人才培养提供坚实支撑,开创教育教学的崭新局面。

参考文献:

- [1] FLAVELL J H. Cognitive development[M]. Englewood Cliffs, NJ: Prentice Hall, 1977: 115.
- [2] 王晓平, 胡娇. 图表理解的认知机制及其教学策略[J]. 江西教育科研, 2007(8): 103-105.
- [3] 李曼丽, 丁若曦, 张羽, 等. 从认知科学到学习科学: 过去、现状与未来[J]. 清华大学教育研究, 2018, 39(4): 29-39.
- [4] Shazeer N, Cheng Y, Parmar N, et al. Mesh-TensorFlow: Deep Learning for Supercomputers[J]. CoRR, 2018, abs/1811.0208.
- [5] Lewis P, Perez E, Piktus A, et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks[J]. Advances in Neural Information Processing Systems, 2020, 33: 9459-9474.
- [6] Margaret Ryznar, Exams in the Time of ChatGPT[J]. Wash. & Lee L. Rev. Online, 2022(80): 305-322.
- [7] Zhao, H., & Wang, Z. (2023). Siren's Song in the AI Ocean: A Survey on Hallucination in Large Language Models[J]. arXiv preprint arXiv:2309.15666.
- [8] Shao, R., He, J., Asai, A., et al. (2024). Scaling Retrieval-Based Language Models with a Trillion-Token Datastore[J]. arXiv preprint arXiv:2407.12854.