



生成式AI赋能“文检课”作业批改实践探索

范凤英 简云泓 曾晓牧 韩丽凤

清华大学图书馆



引言

文献检索课程是信息素养教育的重要内容，旨在提升学生适应数智时代与数字化社会的基础能力，强化其在职业发展与学术研究中的核心素养。文献调研报告作为文献检索课程考核的重要环节，既是学生检索能力和学术表达能力的综合体现，也是教师评估教学效果的重要依据。

然而，传统的人工批改模式在实际教学过程中面临诸多困境，生成式AI为此提供了解决问题、提高效率思路，借助生成式AI构建智能批改模块，可显著提升“文检课”文献调研报告的批改效率，并兼顾准确性。

研究方法

本研究综合运用实验对比法与问卷调查法，结合定量分析与定性分析设计评价体系，对人工批改与智能批改的效果进行系统评估。评价体系主要围绕评分一致性、批改质量和效率、学生反馈与接受度三个核心维度展开。

实验对比法

- 30份文献调研报告
- 人工批改与智能批改结果对比

问卷调查法

- 对智能批改方式的看法与态度
- 50名不同的学生

评分标准设计由五大模块构成：

- 选题简介
- 检索范围与策略
- 检索过程介绍
- 文献调研总结
- 经验总结与体会

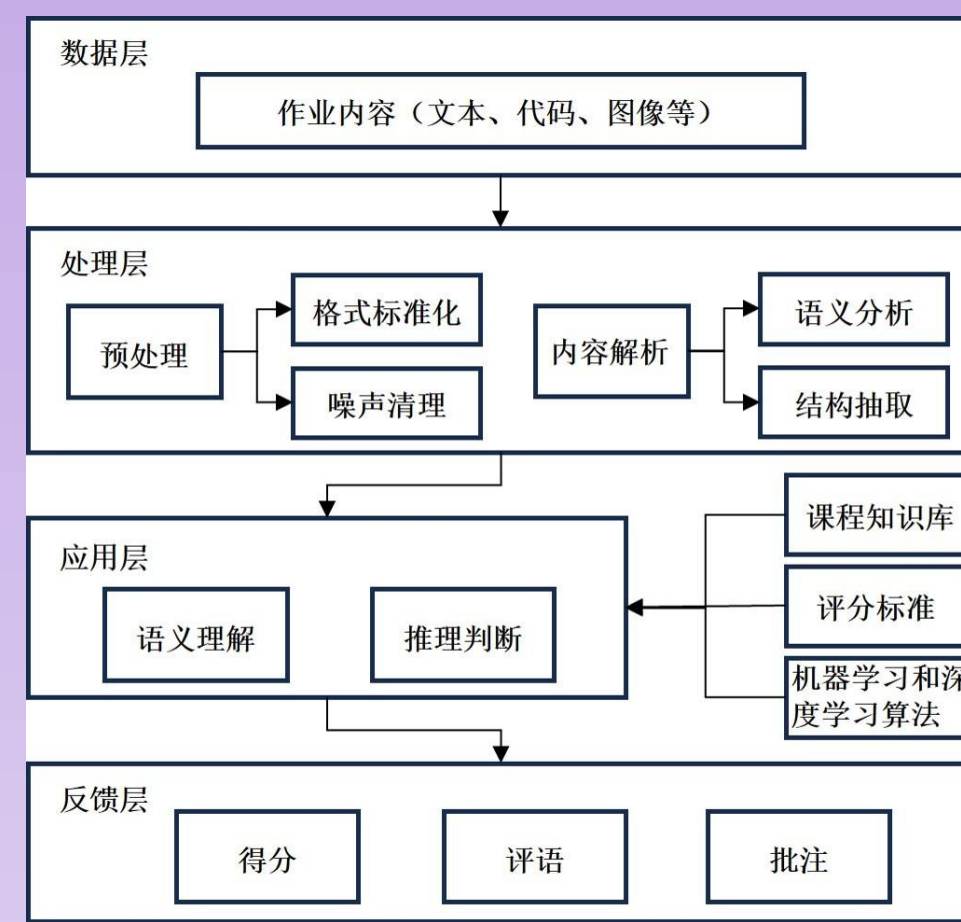


图1 智能批改模块工作流程

结果

- 两种批改方式在五个部分评分及总评分上均表现出较好的一致性。
- 其中，在“选题简介”和“检索范围及策略”两部分上，人工批改和智能批改评分一致性表现为中等。

表1 人工批改和智能批改评分一致性分析

报告内容	人工批改	智能批改	ICC	P
选题简介	8.94±0.62	7.41±0.91	0.469	0.003
检索范围及策略	19.78±1.07	21.81±0.68	0.409	0.009
检索过程介绍	26.97±2.88	35.34±3.63	0.603	0.000
文献调研总结	10.22±3.84	10.58±2.93	0.850	0.000
经验总结及体会	8.84±0.92	7.50±0.72	0.616	0.000
内总体内容	74.75±5.62	82.64±5.48	0.841	0.000

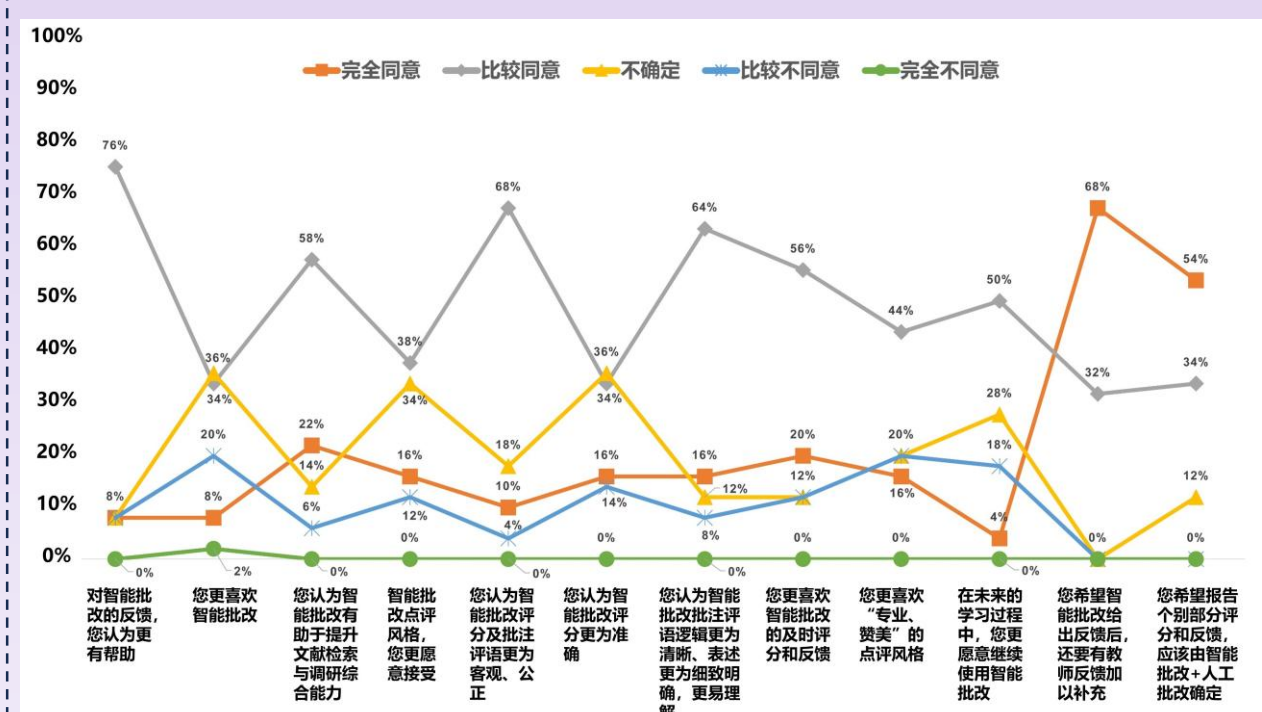


图3 问卷调查结果

- 54%的学生愿意在未来的学习过程中继续使用智能批改。
- 所有学生(100%)均希望在智能批改提供反馈的基础上，教师能够进一步补充反馈。

结论

- 智能批改模块极大提升批改效率，并生成更为全面、细致的反馈信息。
- 学生对智能批改的接受度较高，并认可其在客观性、清晰度和时效性方面的优势。
- 学生更倾向于采用智能批改与人工批改相结合的混合模式

参考文献

- 王媛, 韩丽凤, 金康斌. 清华大学信息素质教育的历史回顾与未来展望[J]. 大学图书馆学报, 2024, 42(3): 23-28. (WANG Y, HAN L F, JIN J B. Historical review and future prospects of information literacy instruction at Tsinghua University[J]. Journal of academic libraries, 2024, 42(3): 23-28.)
- 韩丽凤, 王媛, 曾晓牧, 等. 面向高质量本科人才培养的信息素养教育创新探索[J]. 大学图书馆学报, 2023, 41(5): 62-68. (WANG Y, HAN L F, ZENG X M. Exploring innovative information literacy education for high-quality undergraduate talent cultivation[J]. Journal of academic libraries, 2023, 41(5): 62-68.)
- 刘影婧, 韩丽凤. AIGC背景下高校信息素养教育的发展[J]. 大学图书馆学报, 2024, 42(2): 46-51. (LIU X T. The development of information literacy education in colleges in the context of AIGC[J]. Journal of academic libraries, 2024, 42(2): 46-51.)
- 张贵香, 贾君枝. 生成式AI时代下的提示语培育研究[J]. 大学图书馆学报, 2024, 42(6): 63-71. (ZHANG G X, JIA J Z. Research on prompt literacy cultivation in the era of generative AI[J]. Journal of academic libraries, 2024, 42(6): 63-71.)
- 郭亚军, 徐亮典, 梁艳丽, 等. 从ChatGPT到DeepSeek: 生成式人工智能迭代对图书馆的影响[J]. 图书馆论坛, 2024, 44(1306): 20250226.1616.005.html. (GUO Y J, XU Y X, LIANG Y L, et al. From ChatGPT to DeepSeek: the impact of generative artificial intelligence iterations on libraries[J]. Library tribune, 2024, 44(1306): 20250226.1616.005.html.)
- 政府工作报告——2024年3月5日第十四届全国人民代表大会第二次会议[EB/OL]. [2024-12-29].https://www.gov.cn/gongbao/2024/issue_11246/202403/content_6941846.html.
- Page E B. Grading essays by computer: Progress report [A]. Proceedings of the Invitational Conference on Testing Problems [C]. New York: Princeton, NJ: Educational Testing Service, 1967:87-100.
- Ramesh, D., Sanampudi, S. K. An automated essay scoring system: a systematic literature review[J]. Artificial Intelligence Review, 2022, 55(3):2495-2527.
- Jiang X J. Automatic Correction Method for English Composition Errors Assisted by Artificial Intelligence[C]/2024 International Conference on Interactive Intelligent Systems and Techniques (IIST). IEEE, 2024: 264-268.
- Shao H, Liu Z. Mobile Smart App and its Application in Improving the Efficiency of English Homework Correction[J]. Scalable Computing: Practice and Experience, 2024, 25(3): 1631-1646.
- 于瑞利, 韩宗斌. 智能批改赋能作文教学减负增效的验证实践——以中文作文智能批改系统应用为例[J]. 中小学信息技术教育, 2025, (01):86-87. (YU R L, HAN Z P. Evidence-Based Practice of Intelligent Grading Empowering Composition Teaching: Reducing Burden and Increasing Efficiency—A Case Study of a Chinese Composition Intelligent Grading System[J]. Primary and secondary school information technology education, 2025, (01):86-87.)
- 王家豪. 基于深度学习的数学口算试题智能识别与批改算法研究[D]. 南宁: 广西民族大学, 2023. DOI:10.27035/d.cnki.ggxmc.2023.000581. (WANG J H. Research on intelligent recognition and correction algorithm of mathematical oral test questions based on deep learning[D]. Nanning: Minzu University, 2023. DOI:10.27035/d.cnki.ggxmc.2023.000581.)
- Ganguli, I., Bhowmik, R. S., Biswas, S., et al. Intelligent program correction and evaluation system[J]. INTERNATIONAL JOURNAL ON ARTIFICIAL INTELLIGENCE TOOLS, 2024, 33(02): 2350071.
- 刘小峰. 基于立体结构的工程制图作业智能批改技术的研究与实践[D]. 镇江: 江苏科技大学, 2016. (LIU X T. The research and practice on intelligent correcting technology of engineering drawing assignment based on three-dimensional reconstruction[D]. Zhenjiang: Jiangsu University of Science and Technology, 2016.)
- 花芳, 孙平. 利用“文献检索与利用”课培养大学生的信息素养[J]. 图书馆工作, 2006, 50(10): 128-130. (HUA F, SUN P. Cultivation of students' information literacy by the course of information retrieval and utilization[J]. Library and information service, 2006, 50(10): 128-130.)
- 杨光武. 项目课程法在文献检索课教学中的应用[J]. 大学图书馆学报, 2008, (04): 81-84. (YANG G W. Exploring the application of project-based curriculum method in literature retrieval course[J]. Journal of academic libraries, 2008, (04): 81-84.)
- 郭太敏, 曹志梅, 谭黎娟, 等. 大学生信息素养一体化教育体系及其构建对策[J]. 大学图书馆学报, 2012, 30(02): 102-104. (GUO T M, CAO Z M, TAN L J, et al. Countermeasure research on integrated education system of information literacy for undergraduate students[J]. Journal of academic libraries, 2012, 30(02): 102-104.)
- 花芳. 以馆员素质培养为目标的《文献检索与利用》课程教学内容的模块化[J]. 图书馆工作, 2016, 60(16): 86-92. (HUA F. Design and modularization of teaching contents of "literature retrieval and utilization" aiming at cultivating intelligence literacy[J]. Library and information service, 2016, 60(16): 86-92.)
- 花芳, 程爱平. 曾晓牧. 以“学”为中心: 文献检索与利用课程的教学设计理念[J]. 教育探索, 2008, 11: 63-64. (HUA F, CHEN A P, ZENG X M. Student-Centered Approach: Teaching Design Philosophy for Information Retrieval and Utilization Courses[J]. Education exploration, 2008, 11: 63-64.)
- 程安林, 范亮. 高校学生评教结果信度实证研究[J]. 黑龙江高教研究, 2016, (07): 21-25. (CHEN A L, ZHANG L. The reliability test on the college student evaluations of teaching[J]. Heilongjiang research on higher education, 2016, (07): 21-25.)
- Jon Aik K. Artificial Intelligence in Academic Grading: A Mixed-Methods Study [D/OL]. Göteborg: University of Gothenburg/Department of Education, Communication and Learning, 2024-10-02 [2025-02-27]. https://hdl.handle.net/2077/83561.
- 周同强, 罗晋尧, 张和祥, 等. IgA肾病患者随机尿中ACR与24hUTP定量检测的相关性及其对临床诊断的一致性分析[J]. 现代检验医学杂志, 2024, 39(06): 162-166. (FU Y Y, LUO H L, ZHANG H J, et al. Correlation between random urinary ACR and 24-hour UTP quantification in patients with IgA nephropathy and consistency analysis of clinical diagnosis[J]. The journal of modern laboratory medicine, 2024, 39(06): 162-166.)
- Hackl V, Mueller A E, Granitzer M, et al. Is GPT-4 a reliable rater? Evaluating consistency in GPT-4's text ratings[J]. FRONTIERS IN EDUCATION, 2023, 8: 1272229.