

从工具到伙伴：以教师为主体的教师表现性评价系统研究

From Tool to Partner: Research on a Teacher-Centered Human-Computer Collaborative Teacher Performance Assessment System

牛梦晓¹, 刘欢², 孙众^{*}

^{12*}首都师范大学

^{*}sunzhong@cnu.edu.cn

【摘要】 为破解传统教师表现性评价高度依赖专家、成本高昂且难以规模化的实践困境,本研究构建了一个以教师为反思主体的人机协同评价系统。系统遵循分析—协商—改进的人机协同范式,通过整合课前教学设计、课中教学实录视频及 AI 课堂分析报告等多源数据,构建结构化教学案例库。研究采用 BGE-M3 模型进行文本向量化处理以支撑案例检索,并训练了专用的教师表现性评价模型。该系统形成了从教师提问到人机协商,最终生成改进策略的闭环,旨在规模化降低评价成本的同时保障教师的主体性,为人工智能支持教师专业成长提供了可操作的实践路径。

【关键词】 教师表现性评价; 人机协同; 教师专业成长

Abstract: To address the practical challenges of traditional teacher performance assessment, which heavily relies on experts, incurs high costs, and is difficult to scale, this study constructed a human-computer collaborative assessment system with teachers as the reflective subject. The system follows the analysis–negotiation–improvement paradigm of human-computer collaboration. It integrates multi-source data including pre-class instructional designs, in-class teaching videos, and AI classroom analysis reports to build a structured teaching case library. The study employed the BGE-M3 model for text vectorization to support case retrieval, and trained a dedicated teacher performance assessment model. The system forms a closed-loop empowerment path from teacher questioning to human-computer negotiation, ultimately generating improvement strategies. It aims to reduce assessment costs at scale while safeguarding teacher agency, providing an actionable practical path for AI-supported teacher professional growth.

Keywords: Teacher Performance Assessment, Human-Computer Collaboration, Teacher Professional Growth

1. 研究背景

近年来,我国基础教育改革持续深化,教师专业成长与评价机制的完善已成为推动教育高质量发展的核心任务。中共中央国务院《关于深化教育教学改革全面提高义务教育质量的意见》提出要构建以过程性、发展性为导向的教师评价体系,促进教师在评价中实现专业成长。尽管政策导向已明确指向过程性、发展性评价,但教师在实际评价过程中仍面临实践困境。传统评价体系主要依赖于专家听评课,因资源局限,无法常态化,难以满足教师评价的实时性需求(Borko H et al., 2008)。教师表现性评价为困境解决提供理论支持,其强调将评价融入日常教学中,可以满足教师对于评价常态化与实时性的需求。同时可以为教师提供系统化的评价支持,促进教师专业成长。然而,教师表现性评价依赖于专家评价,导致成本高、反

项目基金: 本文系国家自然科学基金人机协同的教师表现性评价关键智能技术研究(项目编号: 62477033)阶段性研究成果。通讯作者: 孙众 sunzhong@cnu.edu.cn

馈慢、难规模化等问题（周文叶，2022），同时教师的主体性在评价过程中难以保证。人工智能技术的发展与应用，为破解这一难题提供思路。通过人工智能赋能课堂行为智能分析、学情数据精准研判、人机多轮对话与意义理解等，为教师专业发展提供了全新路径（王雯等人，2024）。人工智能不再是单纯的工具载体，而是深度融入教学全流程的协同伙伴。因此，本研究设计了人机协同的教师表现性评价系统，助推高素质专业化教师队伍建设迈向高质量新阶段。

2. 理论基础

技术应用须回归评价的核心目标，正如教育评价学者斯塔弗尔比姆所言（Daniel L. Stufflebeam et al., 1966），评价最重要的意图不是为了证明，而是为了改进。唯有教师成为评价主体之一，将作为评估与管理工具的课堂教学智能分析，转变为助力个人与群体教学改进的技术和方法，才能切实改进课堂教学质量，促进教师专业成长。正是基于这一主体性认识，教师表现性评价成为契合上述理念的应然选择。教师表现性评价是一种旨在评价真实教学实践的专业评估，它强调在真实教学情境中捕获教师的专业实践，通过反思性对话充分表达自身观点，促进教师主体性生成。然而，传统教师表现性评价面临三重困境。其一，评价过程高度依赖专家视域，易受主观性干扰且成本高昂，难以实现规模化覆盖；其二，碎片化观察记录难以形成系统性证据链，无法支撑教师持续的自我监控与循证改进；其三，评价反馈往往滞后且缺乏情境化解读，未能将理论框架有效转化为教师可理解、可操作的实践智慧。

随着人工智能技术的飞速发展，技术赋能为破解上述困境提供了新的可能。生成式人工智能可以自动识别教学行为并生成分析报告，结合大语言模型模拟专家提供即时反馈，理论上可大幅降低实施成本、提升反馈时效。孙众等人基于第四代评价理论，提出的以分析-改进-协商为特征的人机协同教师表现性评价范式，突出了教师主体地位，促进教师反思，提升教师教学能力（孙众等人，2025）。如何将AI技术优势与教师表现性评价深度融合，进而构建即时、有效、可操作的评价系统，成为实践性探索的迫切需求。为此，本研究旨在构建基于人机协同的教师表现性评价系统，在实现技术赋能的同时确保教师主体地位，进而促进教师专业成长，推动高素质专业化教师队伍建设。

为进一步为技术实现提供理论指导，本文对以下概念做出界定。

（1）教师表现性评价

教师表现性评价在教师评价领域的运用，是一种旨在评价真实教学实践的专业评估，它关注教师在促进学生学习中的实际表现。教师表现性评价有着明确的目标指向性——于教师评价而言即教师真实有效地促进学生学习（周文叶等人，2020）。具体而言，它基于真实课堂教学情境，通过教学计划、教学实施、评价学习三大任务收集教师的实践证据，评估教师促进学生学习的专业实践能力（周文叶等人，2021）。同时，它突出教师的主体地位，鼓励教师高度主动参与。

（2）人机协同

在本研究中，人机协同指教育评价场景下人类教师与人工智能系统构成的协作关系（武松等人，2025），其核心在于构建分析—协商—改进的智能支持体系。旨在通过技术赋能实现AI分析结果向个性化、可操作反馈的转化，同时恪守教师专业判断与反思主体性的教育本位。该界定并非简单的技术辅助逻辑，而是强调教师从被动评价对象向基于AI证据的实践性反思者的范式转型，这一转型直接回应了第四代教学评价理论所倡导的沟通与协商价值取向（孙众等人，2025）。实现从信息接收者→证据提供者→反思主导者的角色跃迁。

3. 技术实现

本研究旨在训练教师表现性评价模型，设计教师与模型的交互系统，最终构建一个以教师为反思主体、人机协同的教师表现性评价系统。通过本研究开发的人机协同的教师评价系统，帮助教师提升教学能力，促进教师专业成长。系统技术流程遵循多源资源库输入-教师表现性评价模型分析-人机协商、教学改进的闭环，如图 1 所示。

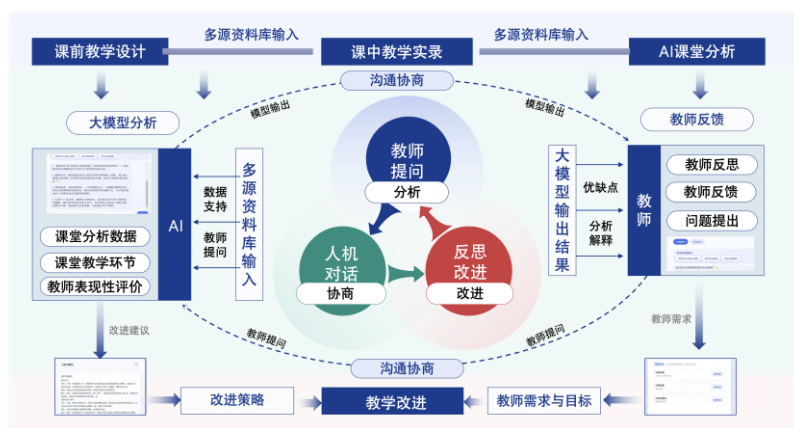


图 1 人机协同的教师评价系统

3.1 构建多源数据支持的教学案例库

系统首先要求教师上传涵盖教学实施前、中、后完整过程的资料，包括课前教学设计、课中教学实录视频及课后的 AI 课堂分析报告。在具体实现中，系统首先将课中教学实录视频转写为带时间戳的师生真实语录，根据时间特点将课堂实录内容与教学评价中的教学活动设计以及 AI 课堂分析报告中的分析数据一一对应。构建包含分析数据与课堂真实流程的映射关系。将其进一步进行文本向量化处理构建资料库，本研究选取三大文本向量化模型进行对比实验，最终选用 BGE-M3 模型作为文本向量化模型。通过实验验证 BGE-M3 模型在测试集上实现 96.7% 的召回率，确保模型反馈精准。通过对三大维度映射内容进行向量化处理，形成一个结构化、体系化的课堂教学案例库，为后续模型处理提供数据支持。

3.2 教师表现性评价模型

该模型的建构以教师表现性评价理论与团队自主研发的《教师表现性评价量表》为训练数据，量表已经过实验验证其有效性。通过采集并结构化处理教师表现性评价理论与相关文献构建专项数据集。并采用微调技术对基座大模型进行训练，训练模型的通用信息处理功能转化为教师表现性评价功能。在应用过程中，教师提出具体问题，模型对接教学案例库，融合教师问题与资源库依据作为模型输入，使用提示词调用模型进行处理。

RAG 层对接系统构建的多源数据支持的教学案例库。主要负责从海量教学资源中精准定位与当前评价情境高度相关的证据片段，为后续生成环节提供可靠的上下文支撑。该层以教师上传的课程资料及权威文献库为检索对象，构建行为数据—教学环节—教师表现性评价的三维语义索引。提示词工程是本系统实现从 AI 课堂分析数据到教师表现性评价转化的关键。通过结构化的提示词设计，将检索增强生成（RAG）所获取的多源证据转化为符合教师表现性评价标准的、可操作的反馈内容。

综合以上模型实现，系统接收教师查询意图后，首先通过 BGE-M3 嵌入模型将查询语义向量化，继而激活混合检索机制。检索结果经重排序算法筛选后，提取 Top-K 相关片段，涵盖：与查询契合的 AI 分析数据段落、对应教学环节的真实课堂实录、教师表现性评价规则及 edTPA 参考案例。

3.3 以教师为主体的人机协商交互系统

人机协商交互系统承接前述教师表现性评价模型,将其从静态分析框架转化为动态专业发展系统。该架构以教师表现性评价结果为协商基准,构建问题提出—证据协商—策略生成的协商路径。

在交互机制层面,本研究构建了以课程实例为单元的情境化对话系统。当教师选定特定课次发起提问时,后端自动聚合该课次的完整教学证据链。通过多模态数据与《教师表现性评价量表》的八条规则及教育理论文献共同构成大语言模型的上下文窗口。同时,系统支持教师对评价结果进行存疑、反驳与证据补充的多轮对话。当教师上传教案修订稿或学生作业样本时,触发二次检索与证据重评,生成协商轨迹的可视化输出。本研究已成功开发并部署了人机协同的教师表现性评价系统,如图2所示。



图2 人机协同的教师表现性评价系统

系统完整实现了分析—协商—改进的机制。通过构建轻量化 Web 交互平台,支持教师用户上传教学设计文档、课堂实录视频及 AI 分析报告等多模态数据。技术验证表明,系统在测试环境中稳定支持并发请求,初步验证了架构设计的可行性与技术路径的有效性,目前该系统开展小规模试点应用。系统实现不仅将理论原理转化为可运行的技术实体,更通过真实教学场景的测试,暴露了当前协商深度不足、改进策略可利用性较差等实践困境。这为后续研究指明了优化方向——需进一步探索人机协同中教师的改进机制,以及基于持续交互数据的模型建议生成路径,从而切实弥合技术支持与教师专业成长之间的转化鸿沟。

参考文献

- 中共中央,国务院.(2019).关于深化教育教学改革全面提高义务教育质量的意见.中华人民共和国中央人民政府. https://www.gov.cn/zhengce/2019-07/08/content_5407361.htm
- 王雯和李永智.(2024). 国际生成式人工智能教育应用与省思.开放教育研究, 30(3), 37–44.
- 孙众和韩硕.(2025).从人工智能课堂教学分析到人机协同的教师专业成长.教师教育论坛, 38(04), 4–12.
- 周文叶.(2020).教师评价:评什么和怎么评——访斯坦福大学李·舒尔曼教授.全球教育展望, 49(12), 3–12.
- 周文叶和董泽华.(2021).教师表现性评价系统的研发与实施逻辑:以 edTPA 为例.教育发展研究, 41(12), 20–27.
- 周文叶.(2022).指向立德树人的教师表现性评价.全球教育展望, 51(04), 129.
- 武松、安富海和张仁杰.(2025).功能论视角下人工智能赋能教育的边界与协同机制研究.教育理论与实践, 45(34), 25–30.
- Borko,H.,Jacobs,J.,Eiteljorg,E.,& Pittman,M.E.(2008).Video as a tool for fostering productive discussions in mathematics professional development. Teaching and Teacher Education,24(2),417–436.