

人工智能赋能研究生教育： 多维智能评价体系设计与实践^{*}

朱承璋 谢泓玥 肖亚龙 冯杰

郭克华^{**}

中南大学人文学院，长沙 410083

中南大学计算机学院，长沙 410083

摘 要 针对研究生教育评价中评价维度单一、过程反馈滞后及个性化支持不足等问题，本文以中南大学研究生课程为实践场域，构建了一种基于人工智能的多维智能评价体系。该体系以学术成果、创新能力、实践能力、团队协作与综合素质五个维度为核心，融合结构化数据、行为数据与文本数据，采用自然语言处理、机器学习与层次分析法等方法实现多源数据建模与综合评分。在实际应用中，系统对研究生学习过程进行持续跟踪与动态分析，生成能力雷达图与成长轨迹曲线，为教师提供精准教学干预依据。系统应用以来，支持 500 余项学生课题实践，覆盖 59 门交叉课程，学生在学习投入度、协作参与度及能力发展均表现出明显提升，能够有效推动研究生教育由结果导向向过程导向转变，实现评价的个性化与数据驱动，为高层次人才培养提供可推广的技术路径与实践范式。

关键字 人工智能，智能评价体系，研究生教育，教学评价改革，跨学科教育

Artificial Intelligence – Empowered Graduate Education: Design and Practice of a Multi-dimensional Intelligent Evaluation System^{*}

Zhu Chengzhang Xie Hongyue Xiao Yalong Feng Jie

Guo Kehua^{**}School of Humanities
Central South University
Changsha 410083, China;School of Computer Science and Engineering
Central South University
Changsha 410083, China

Abstract—To address key limitations in graduate education evaluation — namely the narrow scope of evaluation dimensions, delayed process feedback, and insufficient personalized support — this study develops a multi-dimensional intelligent evaluation system based on artificial intelligence, using graduate courses at Central South University as the application context. The system is structured around five core dimensions: academic achievement, innovation capability, practical ability, teamwork, and overall competence. It integrates structured data, behavioral data, and textual data, and employs methods such as natural language processing, machine learning, and the Analytic Hierarchy Process (AHP) to enable multi-source data modeling and comprehensive evaluation. In practical implementation, the system continuously tracks and dynamically analyzes students' learning processes, generating competency radar charts and developmental trajectory curves to support precise instructional interventions. Since its deployment, the system has supported over 500 student projects and been applied across 59 interdisciplinary courses. Results indicate notable improvements in students' learning engagement, collaborative participation, and competency development. The findings demonstrate that the proposed system effectively facilitates the transition of graduate education evaluation from outcome-oriented to process-oriented approaches, enabling personalized and data-driven assessment and providing a scalable technical pathway and practical paradigm for advanced talent cultivation.

Keywords—Artificial intelligence, Intelligent Evaluation System, Graduate Education, Teaching Evaluation Reform

^{*} **基金资助：** 本文得到中南大学研究生教育教学改革研究项目资助项目“人工智能赋能研究生教育改革与实践”（2025JGA004），湖南省教育科学规划 2024 年度课题“世界一流大学与一流学科建设的基本模式、经验与成效评价”（ND246250），2024 年湖南省学位与研究生教学改革研究项目“课程思政纳入“双一流”建设成效评价体系路径研究”（2024JGYB051），2025 年湖南省普通本科高校教学改革研究项目“数智赋能的中华文化国际传播课程体系建设与改革研究”（202502000114）资助。

^{**} 通讯作者：郭克华 guokehua@csu.edu.cn。

1 引 言

人工智能（Artificial Intelligence, AI）技术的迅猛发展正在深刻重塑教育领域的格局，推动教育信息化向智能化阶段加速演进^[1]。习近平总书记也在全国教育大会上强调“完善高校学科设置调整机制和人才培养模式，加强基础学科、新兴学科、交叉学科建设和拔尖人才培养。”在此背景下，研究生教育作为培养高层次创新人才和驱动科技创新的关键环节，

其评价机制的现代化改革尤为迫切。然而，当前研究生教育评价体系普遍存在显著不足：评价指标侧重单一学术成果，忽视综合素质；反馈机制滞后且缺乏个性化，难以有效捕捉学生多维能力的动态发展等，这些问题制约了研究生培养质量的持续提升。

基于此，本研究以中南大学人文学院为研究场域，主动调整高等教育结构^[2]，借助自然语言处理、机器学习、情感分析等 AI 方法，探索构建融合人工智能技术的研究生多元化评价体系，实现对学习行为、学术成果、实践能力等多维数据的自动识别与深度分析，推动教育教学、评价重心从结果导向往过程引导、能力发展转变，提升人才培养的精准性与适应性。

2 研究生教育评价面临的问题

2.1 评价维度单一，系统性不足

目前多数高校的研究生评价体系过度聚焦于可量化的学术成果，如论文发表数量、科研项目参与情况等，忽视了学生的跨学科能力、创新意识、实践能力、团队协作能力等综合素质的考查。这种以“成果挂帅”为核心的单向度评价模式不仅难以全面反映学生能力结构，也无法为其未来的职业发展提供有效指导。而数智化评价方法能够通过构建多维指标体系、整合多源教育数据，实现对人才培养质量的动态监测与持续改进^[3]。

2.2 理论实践脱节，权重失衡突出

研究生教育强调知识应用与理论创新并重，但现实中实践性成果往往难以获得等同于论文的评价权重^[4]。例如，实验报告、产品设计、调研报告等实践产出常因缺乏清晰的评价标准与主观性过强而被边缘化，导致学生在项目执行过程中缺乏动力，教师也缺乏评价抓手，弱化了“学以致用”的教育目标。

2.3 动态反馈缺位，个性化指导弱

当前的研究生评价模式大多为阶段性终结性评价（如学期成绩、结业考核），难以反映学生在学习过程中的成长轨迹与动态变化。且缺乏基于个体差异的能力诊断与发展建议，导致教学反馈滞后、培养路径同质化^[5]，不能满足“精准教育”与“因材施教”的目标。而基于行为检测、深度学习和层次分析法的量化评估模型，能够将学生课堂行为转化为可计算的过程性指标，为教师实时掌握学情、优化教学策略提供数据支撑^[6]。

2.4 跨学科适配性弱，数据整合受阻

随着交叉学科趋势日益显著，研究生教育中出现了大量跨领域的项目与课题。但目前主流评价体系在指标设计与模型适配上依据单一学科逻辑构建，缺乏

弹性调整机制^[7]。同时，不同平台和部门间数据分散、格式不统一，难以支持全面的综合评价，形成“数据孤岛”现象。

3 智能化教育评价体系的研究理念与总体设计

人工智能技术在高等教育领域的应用正从“工具辅助”走向“结构重构”。基于前述理论基础与问题分析，本研究提出一种面向研究生教育的多维度智能评价体系构建路径。该体系以“能力导向、多模态数据融合、人工智能建模与个性化反馈”为核心逻辑，通过结构化设计实现评价内容、方法与技术的深度融合，推动教育评价从“静态描述”向“动态洞察”转变。图 1 展示了整体研究架构。

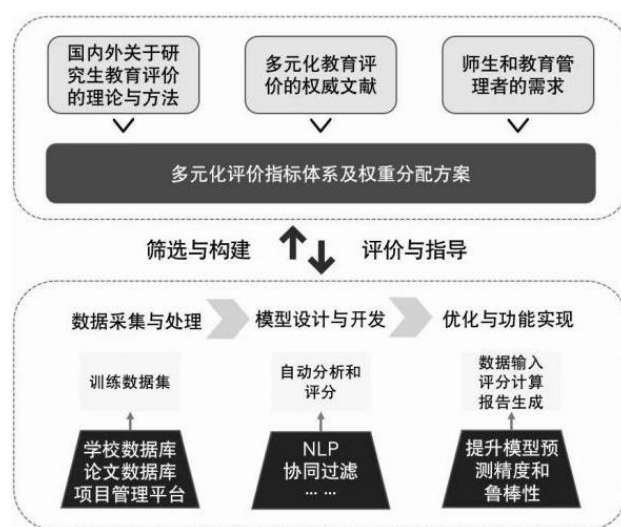


图 1 研究生教育智能评价体系总体框架

面对研究生教育日益复杂的培养要求与评价诉求，本研究以“能力导向、多维融合、技术驱动、过程赋能”为理念核心，结合系统论与教育测量理论，提出一种适配人工智能时代的智能化研究生教育评价体系。该系统分为四大核心层级：一是评价维度设计层，明确系统需达成的能力目标与指标结构；二是数据融合处理层，采集多源异构教育数据并进行标准化处理；三是 AI 建模分析层，匹配不同维度的模型算法实现自动评分与特征提取；四是教学反馈应用层，实现结果可视化展示与教学行为干预支持。这一架构不仅覆盖了从数据输入到教育应用的全过程，也为系统后续模块拓展与任务迭代提供了可持续性基础。

系统设计理念中最具特色的一点在于引入了“多主体参与评价机制”，即在系统功能模块中融合了学生自评、教师协评与 AI 自动评分三类评价来源，通过加权整合实现个体视角、教学视角、数据视角的有机结合，提升系统应用的可接受性与教育适配性。

为保障系统的跨学科适应性与教育公平性，平台架构特别设计了“可配置化能力模板”机制。教师可根据课程特点自定义评价维度与评分模型，并选择系统内置的通用模板作为参考，系统自动匹配推荐相应数据指标与分析模型，有效避免“一刀切”式的评价失真问题。

4 智能评价体系的构建与应用实践

针对当前研究生教育评价中多维指标量化建模范式缺失、AI 模型可解释性与伦理机制尚不完善等关键挑战，本研究以中南大学研究生教育实践为依托，融合多维评价理论、教育大数据技术与可解释人工智能（XAI）方法，构建了一套结构完整、操作规范且具备适应性的研究生多元化智能评价体系，具体构建方法如下。

4.1 多维度能力指标体系与标准制定

高质量研究生教育的核心目标在于全面提升学生的学术素养、创新意识、实践能力与综合素质。传统评价体系多以学术成果（如论文发表数量）作为主要依据，忽视了能力发展的结构性和过程性。本研究在能力导向教育理念指导下，构建了多维能力指标体系：学术成果、创新能力、实践能力、团队协作和综合素质，旨在实现对研究生培养全过程、全要素的系统评价。

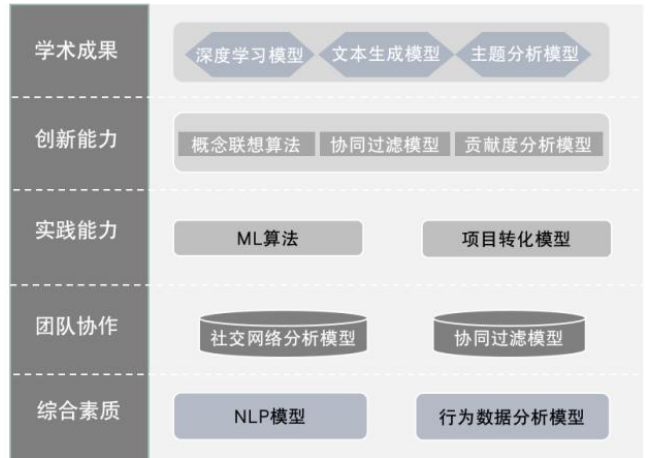


图 2 研究生教育多维能力指标体系构建图

在指标标准制定方面，本研究采用混合方法路径：一方面通过德尔菲法邀请 15 位专家反复打分确定一级维度与关键要素；另一方面运用层次分析法（AHP）计算各指标的权重分布，以构建科学合理的评分模型。例如，学术成果维度的权重在不同专业中略有差异，理工类学生中占比达 35%，而在人文学科中则为 25%，更多权重分配给综合表达与研究设计部分。系统允许教师在评分前进行参数微调，实现跨学科定制。

此外，为保障评分标准的可操作性与可解释性，

每项指标均配备具体的“数据映射规则”，如“研究主题聚焦度”由论文摘要与关键词相似度、引用焦点分析共同生成得分；“协作贡献度”基于任务提交日志与小组互评数据计算加权参与率。通过这些精细化设计，系统可实现指标与模型的直接匹配，减少中间人工干预，提高效率与一致性。

4.2 模型构建与维度映射机制

实现多维度智能评价的前提在于对研究生学习全过程数据的系统化采集与高质量整合。与传统以“单一成绩输入”为主的方式不同，本研究构建的智能评价系统强调对教务、科研、教学、平台行为、社会交互等多源异构数据的全面感知与融合处理，从而为 AI 模型的训练与评价推理提供高质量、结构化的数据基础。

首先，在数据采集方面，系统通过标准化接口与中南大学现有的研究生管理系统、教务系统、学术科研管理平台、在线课程平台（如雨课堂、学堂在线）及校内 GitLab/项目系统对接，自动调取学生课程成绩、项目参与记录、科研成果提交、实验室活动日志等结构化数据。同时，为补充非结构化信息，平台支持学生上传文本报告、研究摘要、反思日志等文档，并引入自然语言处理技术进行语义标签提取。此外，系统还通过 Web 插件采集在线学习行为数据，如学习视频完成率、互动频次、答题轨迹等，形成过程性学习画像。

表 1 多源教育数据采集机制表

数据类型	来源渠道	样例数据	技术处理方法
结构化数据	教务系统、科研管理平台	成绩、论文、项目记录	SQL提取、格式化
非结构化数据	汇报材料、课程作业	文本、PPT、PDF	OCR、文本清洗、NLP编码
行为数据	学习平台、协作系统	登录频次、参与时长	行为特征提取
社交评价数据	同伴互评、教师点评	自然语言文本	情感分析、主题挖掘

其次，在数据处理流程上，系统部署数据清洗与预处理模块，包括重复值剔除、数据脱敏、格式转换、时序对齐、异常值识别等关键步骤。平台融合架构方面，采用“数据湖+动态中台”的设计理念，所有原始数据先进入教育数据湖，分类存储后进入数据中台，由中台完成各类数据的标准建模、维度转换与特征抽取。通过这一机制，系统可灵活支持不同维度、不同粒度的特征融合，如将学生的“课程表现+项目评分+科研产出+社交行为”联动构建个体学习画像，形成高

维度能力表征向量。

针对教育数据的敏感性与伦理风险，系统在采集与处理环节引入隐私保护机制，对教师和研究团队开放的数据均已过脱敏处理。此外，平台遵循学校伦理审批与知情同意原则，在学生入学阶段获得数据授权，并在数据可视化呈现中设置访问权限分级，确保评价结果。

4.3 智能模型设计与评价映射路径构建

在完成评价指标体系构建与数据融合处理之后，系统需通过合理的模型设计，将输入的数据特征精准映射到具体的评价维度得分上。为此，本研究采用“任务导向+模型组合”的建模策略。针对不同维度特征的异构性，系统分别采用适配性最强的 AI 模型：在文本类数据（如论文摘要、项目报告）中，系统使用 BERT 模型提取语义特征，并结合 TF-IDF 计算论文关键词的创新性与相关度，输出学术成果的语义匹配度得分；在结构化行为数据中（如学习频率、项目打卡等），采用 XGBoost 回归模型预测学生行为与表现之间的非线性关系；针对团队协作维度，引入社交网络分析（SNA）构建小组互动图谱，并基于 PageRank 等算法量化学生在项目中的角色与贡献度。

创新能力维度的建模策略是通过学生过往科研数据、课程选修记录与报告关键词等信息，构建其个体知识图谱。在该图谱中，系统检测其概念节点之间的跨领域连接程度与路径密度，量化学生的跨学科思维能力。这一模型不仅体现了学生的“已知”知识结构，还可预测其未来研究方向的多样性与潜力。

为了保障模型在教育场景中的可解释性，系统统一部署了 SHAP（SHapley Additive exPlanations）解释机制，对每项模型输出结果进行特征重要度可视化展示。例如，某学生在“项目执行能力”维度得分偏低，系统可明确指出是由于“任务延迟率高”“协作频率低”等特征导致评分下降，教师据此可进行有针对性的干预指导，增强评价过程的透明度与信任感。此外，系统为教师设置了“评分边界校正”功能，即教师在最终评分阶段可查看 AI 建议得分并进行人工修正，系统会将教师的操作记录为反馈数据，用于后续模型微调与自我优化，从而实现人机协同的闭环建模。

5 应用成效评估与持续优化路径

智能评价系统的核心目标之一在于通过数据驱动的分析方式，促进学生多维能力的全面发展。在研究生课程试点中，系统通过课程成绩、协作行为、作业反馈等数据生成个性化“能力雷达图”与“成长轨迹曲线”，帮助学生识别短板、调整策略，实现个性化

发展。部分跨学科背景的学生在知识整合与迁移、促进个性化学习、自主反思与能力提升方面表现出明显提升。问卷与访谈反馈表明，大多数学生认为系统能帮助发现不足、提供有价值的学习建议，部分学生指出系统反馈更全面、连贯，也有人提出系统仍有待在个性化和语言表达上优化。个体案例也显示，在系统持续反馈与教师辅导下，学生在协作、贡献等维度有明显转变。整体来看，系统有效支持了从“完成任务”向“主动发展”的学生角色转变，体现其在研究生教育中的应用价值。

从更大范围的改革实践看，该评价体系嵌入学校人工智能交叉人才培养的整体机制之中，已形成较为清晰的实施支撑条件与反馈闭环。如图 3 所示，学校已协同 15 个学院建立跨学科协同机制，每年组建 70 余个导师混编团队和 40 余个联合导学中心，构建 54 个跨领域知识图谱与 187 条三阶知识链，支撑 59 门交叉课程和 21 个跨域案例库建设；同时依托 8 个国家级教研平台，与 20 余家行业领军企业共建 35 个实体性交叉实验室，8 年来支持 500 余项学生交叉课题，孵化 258 项创新成果，并推动 170 余项课题成果转化。上述数据表明，多维智能评价体系已经与课程建设、导师协同、实践训练和资源配置形成联动关系，在促进培养过程可追踪、评价结果可校准、教学改进可反馈方面展现出较强的应用潜力，也为研究生教育评价的智能化转型提供了可复制的实践样本。

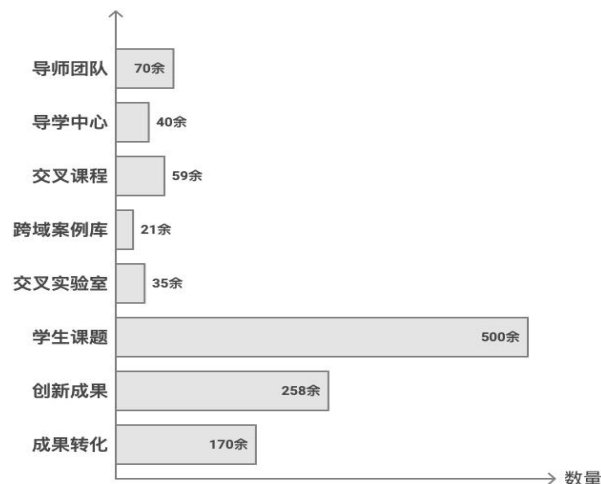


图 3 多维智能评价体系实施成效核心指标分布

6 结束语

在教育数字化与人工智能快速发展的背景下，本研究构建了一种适用于研究生教育场景的多维度智能评价体系，通过系统性的理论设计、技术实施和教学应用，验证了人工智能技术在研究生教育评价领域的有效性与适配性。研究成果的核心贡献体现在三个方面：明确提出以能力导向、多维融合为核心理念的评

价指标体系,解决了传统评价体系单维性、片面性问题;构建了数据驱动的智能评价体系与系统平台,实现从“静态终结性评价”向“动态形成性反馈”的转型,推动评价模式的精准化和个性化;创新性地设计了“人机协同评价机制”,在保障AI评分客观性的同时,保留教师主观评价的灵活性与可控性,增强了系统的可解释性与师生认可度。

本研究在中南大学人文学院的研究生课程中开展为期一学年的应用试点后,学生综合能力显著提升,教师教学反馈更加精准与及时,教学负担明显减轻,教学质量与管理效率均有提升。同时,系统应用也引发了教师角色的积极转型,推动教师从经验型管理向数据支持型教学转变。尽管系统推广过程中遇到了一些技术适配、数据整合与教师接受度方面的挑战,但通过实施灵活的指标自定义策略、技术培训与人机协同机制,相关困难已获得有效缓解,逐步为系统的规模化推广提供了可靠保障。

未来,随着教育数字化的不断深入,本研究提出的智能评价体系在跨学科、跨院系甚至跨校区的应用中具备较强的适配潜力与推广价值。建议高校管理部

门在政策层面给予更多支持,将平台使用情况纳入教学考核指标体系,并配套相应激励机制与技术保障措施。此外,系统未来的发展还需重点关注数据伦理与隐私保护问题,持续加强模型性能优化与用户体验改进,探索更丰富的教育数据来源与更广泛的应用场景。

参 考 文 献

- [1] 姚琳,宋晏,石志国.基于新工科的大学生计算机基础课程体系思考与探索[J].计算机教育,2019,(03):112-116.
- [2] 王浩,吴共庆,胡学钢,等.新工科背景下人工智能通识系列课程建设与实践[J].计算机教育,2019,(02):112-114.
- [3] 冯艳娜,张峰,吴敏宁.应用型本科高校人才培养成效数智化评价方法研究[J].计算机技术与教育学报,2026,14(1):12-17.
- [4] 陈龙,黄鑫,张喆,等.大学计算机课程引入人工智能内容的教学探索[J].计算机教育,2023,(03):203-207.
- [5] 刘浩,刘笑笑,辛涛.人工智能赋能基础教育监测的应用与挑战[J].北京师范大学学报(社会科学版),2022,(02):136-142.
- [6] 傅翠娇,杨一泓.基于深度学习和层次分析的学生专注度量化评估[J].计算机技术与教育学报,2026,14(1):101-106.
- [7] 肖睿,肖海明,尚俊杰.人工智能与教育变革:前景、困难和策略[J].中国电化教育,2020,(04):75-86.