

融入 ChatGPT 的 OBE 目标达成度模型可提升评价效率与个性化反馈研究

师春梅 张锦盛 杨清

云南工商学院, 云南昆明, 650000;

摘要: 成果导向教育 (OBE) 要求精准量化学生学习成果, 但传统的目标达成度评价方法存在耗时长、难以提供个性化反馈的局限性。本研究旨在构建一个融入 ChatGPT (C-OBE) 的教育目标达成度模型, 以显著提升评价效率和反馈的质量。该模型通过设计精细化的提示工程 (Prompt Engineering), 利用大型语言模型 (LLM) 强大的语义理解和逻辑分析能力, 实现对学生非结构化作业 (如项目报告、开放性问答) 的自动化、高一致性评分。实证研究将采用实验组 (C-OBE 模型) 与对照组 (传统人工评价) 对比的方式, 对评分耗时、评分一致性 (与专家评分相关系数) 和学生满意度进行量化评估。结果预期表明, C-OBE 模型能够大幅减少教师的评价负担, 显著提高评分效率, 并能依据评价标准 (Rubrics) 实时生成针对性强、指向明确的个性化反馈, 有效引导学生进行深度学习和持续改进。本研究为推动 OBE 评价体系的智能化升级提供了理论支持和可操作的技术路径。

关键词: 成果导向教育 (OBE); 目标达成度评价; ChatGPT; 大型语言模型 (LLM); 评价效率; 个性化反馈; 提示工程

DOI: 10.64216/3080-1494.26.03.023

1 引言 (Introduction)

1.1 研究背景与问题提出

在高等教育领域, 成果导向教育 (Outcome-Based Education, OBE) 已成为提升人才培养质量的核心理念。OBE 强调以学生实际达成的学习成果为导向, 要求对每一个课程目标 (CO) 的达成度进行精准、量化的评估。然而, 在实际教学过程中, 尤其是在针对报告、设计方案、论文等开放性、非结构化作业的评价环节, 传统的人工评估模式面临着严峻的挑战:

首先, 评价耗时长且效率低下。教师需要投入大量精力, 依据复杂的评分标准 (Rubrics) 对每一份作业进行细致的拆解和匹配, 极大地占用了教学和科研时间。其次, 个性化反馈难以实现。面对动辄数十甚至上百份的作业, 教师往往只能提供通用性或简单的评语, 难以针对性地指出每位学生在知识应用、逻辑推理、技能掌握上的具体弱点和改进方向, 这严重制约了反馈对学生学习的有效引导作用。

近年来, 以 ChatGPT 为代表的大型语言模型 (Large Language Models, LLMs) 在自然语言理解、文本生成和逻辑推理方面展现出颠覆性能力。这为解决 OBE 评价中的效率与质量难题提供了新的技术契机。如何将 LLM 的技术优势与 OBE 的评价逻辑深度融合, 构建一套高效、准确且能够提供深度个性化反馈的

智能化达成度模型, 成为当前教育技术研究的关键前沿问题。

1.2 理论基础与概念界定

本研究主要基于以下理论和概念展开:

成果导向教育 (OBE): OBE 的核心是学习成果, 强调通过系统的评价和反馈机制, 确保学生达到预定的专业和课程目标。目标达成度 (Achievement of Learning Outcomes) 是衡量教学质量的关键指标。

ChatGPT/LLM: 作为强大的自然语言处理工具, ChatGPT 能够进行复杂的语义分析和文本生成。在本研究中, 其主要功能是作为智能评分辅助器和个性化反馈生成器, 基于预设的 Rubrics 进行高效的匹配和输出。

评价效率与个性化反馈: 评价效率主要指单位时间内完成评分的准确性与速度; 个性化反馈则要求反馈信息具有针对性、建设性, 能够精准定位学生的学习盲点, 并提供具体的改进策略。

1.3 研究意义与创新点

本研究旨在提出并验证一种融入 ChatGPT 的 C-OBE 目标达成度模型, 其意义和创新点体现在:

提升评价效率与公平性: 通过 LLM 的自动化处理, 大幅缩短评分周期, 解放教师生产力, 同时以统一的、基于 Rubrics 的算法进行评估, 提高评分的一致性和客观性。

深化个性化反馈质量：模型能够基于学生的实际得分和 Rubrics 的要求，自动生成具体到知识点、逻辑步骤和技能应用层面的深度反馈，实现反馈的智能化和精细化，有效促进学生的持续性改进（Continuous Improvement）。

拓展智能教育评价的边界：将前沿的 LLM 技术引入 OBE 评价的复杂环节，为构建下一代智能化教育评价体系提供新的思路和可操作的模型框架。

本研究期望通过实证分析，验证 C-OBE 模型在提高评价效率和反馈质量方面的显著效果，从而为高等院校推进 OBE 教学改革提供坚实的理论支持和实践指导

2 理论模型构建：ChatGPT 驱动的 OBE 目标达成度模型

2.1 传统 OBE 目标达成度模型的局限性分析

传统的 OBE 目标达成度评价方法，虽然理念先进，但在大规模应用中，由于其高度依赖人工操作，面临着效率与质量的双重挑战，限制了教学的持续改进：

评价的主观性与一致性难题：OBE 评价核心在于将复杂的学习成果（如批判性思维、创新能力）与清晰的评分标准（Rubrics）进行匹配。然而，针对开放性、非结构化作业，不同教师在理解和应用 Rubrics 时容易产生主观差异，导致评分的一致性不高，影响了评价结果的客观性和可靠性。

非结构化数据处理的巨大耗时：学生的项目报告、代码、设计文档等产出属于非结构化数据。教师需要手动逐字逐句分析，将文本内容转化为量化的达成度指标，这一过程消耗大量时间。评估的滞后性使得反馈失去时效性，削弱了对学生学习行为的即时引导作用。

反馈机制的深度不足：传统模式下，教师难以对每位学生提供深度、具体的个性化反馈。反馈往往是概括性的，缺乏与 Rubrics 指标的精准关联，无法明确指出学生的具体失分点和改进方向，使得 OBE 要求的持续改进回路难以有效闭合。

2.2 ChatGPT 在目标达成度评价中的功能定位

以 ChatGPT 为代表的大型语言模型（LLMs）具备强大的自然语言理解、逻辑推理及文本生成能力，为克服传统 OBE 模型的局限性提供了技术突破口。在本研究构建的 C-OBE（ChatGPT-OBE）模型中，LLM 被定位为高度智能化的评价辅助核心，承担以下三重功能：

客观评分辅助器（Objective Scoring Assistant）：LLM 能够基于明确的 Rubrics，对大规模的开放性作业进行高效的语义匹配和自动化评估。它能将学生产出的复杂语义与 Rubrics 中的等级描述进行逻辑比对，输出

客观、量化的初评得分，有效提升评分效率和一致性。

非结构化数据结构化处理器：模型能够自动提取学生作业中的关键论点、技术细节和逻辑结构，并将其转化为易于计算和统计的结构化评价数据。这一功能极大地简化了数据分析流程，支撑了后续达成度指标的精确计算。

深度个性化反馈生成器（In-depth Personalized Feedback Generator）：这是模型的核心创新。ChatGPT 能同时输入学生的作业、评分标准和得分结果，基于此三者，实时生成针对性极强、具备建设性语言风格的反馈报告。该报告不仅指出得分，更重要的是能够精准定位学生的薄弱环节，并提供基于 Rubrics 更高要求的改进策略和学习资源推荐。

2.3 C-OBE 模型框架设计（ChatGPT-OBE Achievement Model）

C-OBE 达成度模型是一个以 LLM 为核心处理单元的四层架构，确保评价流程的智能化、自动化和反馈的定制化。

2.3.1 输入层（Input Layer）

作业数据：学生的项目报告、代码、案例分析等非结构化文本。

评价标准（Rubrics）：包含清晰的评价指标、权重和等级描述，作为 LLM 评分的唯一依据。

目标矩阵：明确该作业对应的课程目标（COs）及权重。

2.3.2 处理层（Processing Layer - LLM Core）

提示工程与规范化：这是实现精准评估的关键。通过设计精细化的指令集（Prompt），要求 LLM 严格遵守 Rubrics 评分，并以预定的 JSON 格式输出结果。

语义与逻辑分析：LLM 对学生产出进行深度文本解析，完成内容与 Rubrics 的多维匹配，评估其论证逻辑、技术应用和高阶思维的达成情况。

量化计算：依据预设的权重对 LLM 输出的各项指标得分进行加权平均，计算出最终的 CO 达成度得分。

2.3.3 输出层（Output Layer）

量化得分报告：提供学生在各个指标和总体 CO 上的精确得分。

个性化反馈报告：提供对得分的详细解释、具体的失分分析、以及基于 Rubrics 的下一步改进建议和资源指引。

2.3.4 反馈回路（Feedback Loop）

人机协同审核：教师对 LLM 的评分结果和反馈进行抽样审核，对提示工程进行持续优化和迭代，确保模型的准确性与权威性。

持续改进：学生利用反馈进行学习和作业修改，数

据回流系统,形成评价、改进、再评价的闭环,实现 OBE 的核心目标。

3 模型实现与实证研究设计

3.1 模型实现的技术路线

C-OBE 模型的实现依赖于对大型语言模型(LLM)的精细化控制和数据流程优化,核心在于提示工程(Prompt Engineering)。

提示工程的核心设计:必须设计精细化的指令集,指导 ChatGPT 严格依照 OBE 评价逻辑进行操作。这包括在提示词中明确模型扮演的“评分专家”角色、完整嵌入课程目标和 Rubrics(评价标准)、并要求模型以 JSON 或表格等结构化格式输出评分和反馈结果。同时,设定反馈的建设性语气和内容规范,确保反馈的质量。

API 调用与数据集成:利用 ChatGPT 的 API 接口,搭建数据处理平台,负责接收学生作业、调用 LLM 进行处理、解析结构化评分结果,并最终生成报告。

数据安全与隐私:确保学生作业和评分数据在处理过程中进行脱敏,并符合相关数据隐私保护规范。

3.2 实验研究设计与数据采集

为验证 C-OBE 模型的有效性,本研究将采用准实验设计。

实验对象与分组:选取某高校开设的具有明确 OBE 目标和开放性作业的课程(如项目报告)。将学生作业随机划分为:

对照组:教师按传统模式人工评分和反馈。

实验组:采用 C-OBE 模型进行自动化评分和反馈。

专家评分组(黄金标准):邀请资深学科专家对部分样本进行独立评分,用于验证模型的评分一致性。

数据采集:采集以下关键数据:

效率数据:记录两组完成评分和反馈的总耗时。

有效性数据:采集对照组、实验组和专家组的量化得分。

反馈质量数据:通过学生满意度问卷(关于及时性、针对性等维度)和后续作业成绩增量,评估反馈的实际驱动作用。

本研究将通过对比评分耗时比、评分一致性(皮尔逊相关系数 r_s)以及学生对个性化反馈的满意度与学习提升度,全面量化评估 C-OBE 模型在提高 OBE 评价效率和深化反馈质量方面的可行性与优越性。

4 结果与讨论

本研究构建并提出了融入 ChatGPT 的 C-OBE 目

标达成度模型,为解决传统成果导向教育(OBE)评价中效率低下和个性化反馈不足的难题提供了创新方案。通过利用大型语言模型(LLM)强大的语义理解和逻辑分析能力,C-OBE 模型能够实现对大规模非结构化作业的自动化、高一致性评分,预期将显著提升评价效率并解放教师生产力。更重要的是,模型能够基于评价标准(Rubrics)实时生成针对性强、具体到知识点的深度个性化反馈,有效闭合 OBE 要求的持续改进回路,对学生的后续学习具有强大的驱动作用。本研究成功将 LLM 技术与 OBE 评价逻辑结合,为高等院校构建智能化、可操作的下一代评价系统奠定了理论基础,其核心价值在于推动 OBE 评价体系的智能化升级和人机协同新范式的建立。

尽管 C-OBE 模型展现出巨大潜力,但在推广中仍需关注挑战。实践中,高校应引导教师从“评分执行者”转变为“评价标准设计者”和“智能系统管理者”,同时必须建立严格的数据安全和隐私保护机制。模型在处理涉及高阶复杂目标(如创新性、伦理道德)的评价准确性仍有待提升。未来研究应致力于多模态评价探索(例如结合视觉和语音技术评价现场展示)、深入研究模型与人类的信任机制,并开发反馈的自适应迭代系统,以期最终实现技术赋能教育,推动高等教育评价体系迈向高质量发展的新阶段。

参考文献

- [1]唐春娜.OBE 与 ChatGPT 的双元融合教育实践[J].创新创业理论与实践,2024,7(21):58-60.
- [2]郭帅,陈有.OBE 理念下外语人才培养目标达成度评价体系研究[J].佳木斯大学社会科学学报,2025,43(11):194-196+200.
- [3]殷凤梅,王筱薇.面向工程教育认证的操作系统课程目标达成度评价分析与持续改进方法研究[J].长春大学学报,2025,35(08):104-108.
- [4]王航,黄小平,甘国操,等.基于大型语言模型的异构多智能体系统[J].自动化与信息工程,2025,46(04):22-27.
- [5]黄涛.基于大型语言模型的智能知识服务平台[D].哈尔滨工业大学,2025.

作者简介:师春梅(1992.10-),女,汉族,云南省玉溪市,云南工商学院,本科,讲师,软件开发;
张锦盛(1977.01-),男,汉族,云南省昆明市,云南工商学院,硕士,教授,软件开发;
杨清(1993.09-),女,汉族,云南省昆明市,云南工商学院,学士,讲师,计算机网络。