

深度学习在网络教学资源质量评估中的应用研究

王晓丽

(淮阴工学院, 江苏 淮安 223001)

摘要: 随着互联网技术与教育信息化的深度融合, 网络教学资源呈现爆发式增长, 但其质量参差不齐, 严重影响教学效果与教育资源的有效利用。传统网络教学资源质量评估依赖人工评判, 存在效率低下、主观性强、标准不统一等问题, 难以适应海量资源的评估需求。深度学习技术凭借其强大的特征自动提取、复杂模式识别能力, 为解决这一难题提供了新的技术路径。本文首先构建科学、全面、可量化的网络教学资源质量评估指标体系, 明确评估维度与量化标准; 在此基础上, 设计并实现基于深度学习的网络教学资源质量评估模型, 通过数据采集与预处理、模型构建、验证优化等环节, 实现对网络教学资源质量的自动化、精准化评估; 最后, 总结研究成果, 分析研究不足并展望未来研究方向, 为网络教学资源的规范化管理、优质化建设提供理论支撑与技术参考。

关键词: 深度学习; 网络教学资源; 质量评估; 指标体系; 评估模型

中图分类号: G642.0

文献标识码: A

文章编号: 3106-2709 (2025) 07-0006-05

DOI: 10.62022/NCAR.issn3106-2709.2025.07.002

Research on the Application of Deep Learning in Quality Evaluation of Online Teaching Resources

Wang Xiaoli

(Huaiyin Institute of Technology, Huai'an, Jiangsu 223001)

Abstract: With the deep integration of Internet technology and educational informatization, online teaching resources have shown explosive growth, but their quality is uneven, which seriously affects the teaching effect and the effective use of educational resources. Traditional online teaching resource quality evaluation relies on manual evaluation, which has problems such as low efficiency, strong subjectivity, and inconsistent standards, making it difficult to meet the evaluation needs of massive resources. Deep learning technology, with its powerful ability to automatically extract features and recognize complex patterns, provides a new technological path to solve this problem. This article first constructs a scientific, comprehensive, and quantifiable quality evaluation index system for online teaching resources, clarifying the evaluation dimensions and quantitative standards; On this basis, design and implement a deep learning based quality evaluation model for online teaching resources. Through data collection and preprocessing, model construction, verification and optimization, etc., achieve automated and accurate evaluation of the quality of online teaching resources; Finally, summarize the research results, analyze the research shortcomings, and look forward to future research directions, providing theoretical support and technical references for the standardized management and high-quality construction of online teaching resources.

Keywords: deep learning; Online teaching resources; Quality assessment; Indicator system; evaluation model

一、网络教学资源质量评估指标体系构建

(一) 评估指标体系构建原则

网络教学资源质量评估指标体系的构建是实现精准评估的基础, 需遵循科学性、全面性、可操作性、动态性、导向性五大核心原则, 确保指标体系的合理性与实用性。科学性原则要求指标设计基于教育教学规律、网络资源传播特性, 参考国内外相关研究成果与行业标准, 避免主观臆断, 确保指标内涵清晰、逻辑严谨, 能够真实反映资源

质量水平。全面性原则强调指标体系需覆盖网络教学资源的核心维度, 不仅要关注资源的内容质量, 还要兼顾形式规范、技术适配、交互体验等多个方面, 避免单一维度评估导致的片面性。

可操作性原则要求筛选的指标能够通过有效手段获取数据、进行量化处理, 避免设置过于抽象、难以衡量的指标, 确保评估工作能够高效落地。动态性原则指指标体系需适应教育信息化的发展趋势, 根据网络教学模式的创新、

作者简介: 王晓丽, 硕士研究生, 研究方向为网络教学。

资源形式的更新(如短视频、虚拟仿真资源等),适时调整指标内容与权重,保持指标体系的时效性。导向性原则强调指标体系需引导网络教学资源向优质化、规范化方向发展,突出核心指标的导向作用,助力资源建设者明确改进方向,提升资源质量。

(二) 核心评估指标筛选与确定

结合网络教学资源的类型(如课件、视频、试题、案例等)与使用场景,通过文献研究法、德尔菲法(专家咨询法)、问卷调查法,筛选并确定四大核心评估维度,每个维度包含若干具体评估指标,形成多层次的评估指标体系。

1.内容质量维度:作为评估的核心,直接决定资源的教学价值,包含内容准确性(知识点无错误、表述严谨)、内容针对性(贴合教学目标、适配受众群体)、内容完整性(覆盖教学重难点、逻辑连贯)、内容时效性(紧跟学科发展、更新及时)4项具体指标。内容质量是网络教学资源的核心竞争力,也是保障教学效果的关键,因此在指标体系中赋予较高权重。

2.形式规范维度:保障资源的可传播性与可使用性,包含格式标准化(符合行业通用格式、兼容主流播放/浏览工具)、排版合理性(布局清晰、字体规范、图文搭配协调)、标识完整性(包含资源名称、作者、适用年级/学科、使用说明等)3项具体指标。规范的形式能够提升资源的易用性,降低使用者的操作成本。

3.技术适配维度:适应网络环境与终端设备的多样化需求,包含加载速度(在常规网络环境下快速加载)、兼容性(适配电脑、手机、平板等不同终端)、稳定性(无卡顿、崩溃等问题)、安全性(无恶意代码、版权合规)4项具体指标。随着移动学习的普及,技术适配能力成为影响资源使用体验的重要因素。

4.交互体验维度:提升资源的参与度与教学效果,包含交互设计(具备提问、练习、反馈等交互功能)、易用性(操作简单、导航清晰)、反馈及时性(针对用户操作能够快速给出反馈)3项具体指标。良好的交互体验能够激发学习者的主动性,增强资源的教学实用性。

通过专家咨询与权重计算(采用层次分析法),确定各维度及具体指标的权重,其中内容质量维度权重最高(40%),其次为技术适配维度(25%)、形式规范维度(20%)、交互体验维度(15%),确保指标体系的科学性与针对性。

(三) 评估指标的可量化处理

为实现网络教学资源质量评估的客观性与精准性,规避传统定性评估中主观臆断、标准模糊的弊端,需对筛选确定的各项评估指标进行系统的可量化处理,核心是将抽象的定性描述指标转化为可计算、可对比的定量数据,明确各层级指标的具体评分标准与科学量化方法。本次评估统一采用百分制评分体系,根据各指标的权重占比,通过加权求和公式计算得到资源的综合质量得分,确保评估结果的客观性、可比性与实用性,为后续深度学习模型的训练与评估提供标准化的数据支撑。其中,综合质量得分计算公式为:综合得分=Σ(各指标得分×对应指标权重),通过该公式可精准整合各维度评估结果,全面反映资源的整体质量水平。

对于可直接通过技术手段获取具体数值的定量指标(如加载速度、资源大小、视频时长、文本字数等),无需复杂的转化过程,直接通过专业工具采集原始数据,再按照预设的分级标准划分评分等级,确保量化结果的精准性与一致性。例如,加载速度指标结合当前主流网络环境(带宽≥100M)的实际使用场景,划分四级评分标准:加载时间≤2秒(90-100分),此时资源加载流畅,无任何卡顿,用户体验极佳;加载时间2-5秒(70-89分),加载速度良好,基本不影响用户正常使用;加载时间5-10秒(50-69分),加载速度一般,存在轻微卡顿,影响部分用户体验;加载时间>10秒(0-49分),加载速度缓慢,无法满足正常使用需求,视为技术适配不合格。再如视频类资源的大小指标,结合存储成本、传输效率与播放体验,划分四级评分标准:≤500MB(90-100分),兼顾清晰度与传输效率;500MB-1GB(70-89分),清晰度良好,传输无明显压力;1-2GB(50-69分),清晰度尚可,但传输速度较慢、占用存储较多;>2GB(0-49分),传输成本高、播放易卡顿,不符合网络教学资源的轻量化需求。

对于难以直接获取数值、需依靠主观判断的定性指标(如内容准确性、排版合理性、交互设计合理性等),采用“分级描述+专家打分+深度学习辅助”的三维量化方式,最大限度降低人工打分的主观性,提升量化结果的可靠性。首先,针对每一项定性指标,明确四级评分标准的具体描述,确保专家打分有章可循、标准统一。以内容准确性指标为例,四级评分标准具体为:90-100分(完全无错误、表述严谨,知识点准确无误,符合学科规范与教学要求,

无任何歧义)；70–89分(存在轻微错误、不影响整体理解，如个别用词不当、标点错误，不涉及核心知识点错误，不影响教学使用)；50–69分(存在明显错误、影响部分内容理解，如部分知识点表述偏差、逻辑漏洞，会导致学习者对部分内容产生误解)；0–49分(存在严重错误、无法正常使用，如核心知识点错误、表述混乱，会误导学习者，失去教学价值)。同时，引入深度学习模型提取的特征数据作为辅助参考，如通过文本语义相似度模型判断内容准确性、通过图像清晰度检测模型判断排版合理性，将模型输出的特征数值与专家打分结合，采用加权平均的方式确定最终得分，有效降低人工主观判断的偏差，提升定性指标量化的科学性与精准性。

此外，考虑到网络教学资源类型多样(如视频资源、课件资源、试题资源、微课资源等)，不同类型资源的核心特征与使用场景存在差异，需对指标的量化标准进行针对性调整，确保量化方法的适配性与实用性，避免“一刀切”的量化模式导致评估结果失真。例如，针对视频类资源，在基础量化指标的基础上，增加“画面清晰度”“音频质量”“帧率稳定性”等细分量化指标，其中画面清晰度分为4K(90–100分)、1080P(70–89分)、720P(50–69分)、480P及以下(0–49分)四个等级，音频质量根据杂音、音量稳定性等分为优质(90–100分)、良好(70–89分)、一般(50–69分)、较差(0–49分)四个等级；针对课件资源，增加“知识点逻辑连贯性”“图文比例合理性”“重点内容突出度”等细分量化指标，其中知识点逻辑连贯性根据知识点排布的合理性、过渡的自然性分为优秀(90–100分)、良好(70–89分)、一般(50–69分)、较差(0–49分)四个等级，图文比例合理性根据图文搭配的协调性、可读性分为合理(90–100分)、较合理(70–89分)、一般(50–69分)、不合理(0–49分)四个等级，通过细分指标进一步提升评估的针对性与精准性，让指标体系更贴合不同类型网络教学资源的质量评估需求。

二、基于深度学习的网络教学资源质量评估模型设计与实现

(一) 模型设计目标与思路

1.设计目标:构建基于深度学习的网络教学资源质量评估模型，实现对不同类型网络教学资源(文本、图像、视频、音频等)质量的自动化、精准化评估，解决传统人工

评估效率低、主观性强、标准不统一的问题；模型能够自动提取资源的深层特征，结合构建的评估指标体系，输出资源的综合质量得分与各维度评分，为资源筛选、优化提供精准参考；同时，确保模型具备良好的泛化能力，能够适应不同学科、不同类型的网络教学资源评估需求。

2.设计思路:以评估指标体系为核心导向，结合深度学习技术的特征提取优势，分模块设计模型架构。首先，针对不同类型的网络教学资源，设计对应的特征提取模块，实现文本、图像、视频、音频等多模态特征的自动提取；其次，构建特征融合模块，将多模态特征与评估指标对应的量化数据进行融合，形成统一的特征向量；然后，设计评估预测模块，基于融合后的特征向量，预测资源的各维度评分与综合质量得分；最后，通过数据验证与参数优化，提升模型的评估精度与泛化能力，确保模型能够满足实际应用需求。

(二) 数据采集与预处理

数据采集与预处理是模型训练与验证的基础，直接影响模型的评估效果，主要分为数据采集、数据清洗、数据标注、数据划分四个环节。

1.数据采集:采用多渠道采集网络教学资源数据，涵盖基础教育、高等教育、职业教育等多个领域，包含课件(PPT、PDF)、教学视频、试题、教学案例、微课等多种类型，确保数据的多样性与代表性。采集渠道包括主流网络教学平台(如MOOC、学堂在线)、教育资源库、高校精品课程网站等，共采集各类网络教学资源10000+条，同时收集每条资源的人工评估得分(基于构建的指标体系，由3–5名教育领域专家打分取平均值)，作为模型训练的标签数据。

2.数据清洗:对采集到的原始数据进行预处理，去除无效数据、重复数据、异常数据(如无法打开的资源、内容残缺的资源、评分异常的资源)；对不同格式的资源进行标准化处理，例如，将不同格式的视频转换为统一格式(MP4)，将课件转换为PDF格式，便于特征提取；对文本类资源(如试题、案例)进行去噪处理，去除无关字符、冗余信息，统一文本编码格式。

3.数据标注:基于构建的评估指标体系，对清洗后的资源进行精细化标注，标注内容包括各维度评分与综合质量得分。对于文本类资源，额外标注知识点准确性、逻辑连贯性等关键特征；对于视频类资源，标注画面清晰度、音频质量、内容完整性等特征；标注工作由专业的教育工作

者与技术人员协作完成,确保标注数据的准确性与一致性,为模型训练提供可靠的标签支撑。

4.数据划分:将预处理后的数据集按照7:2:1的比例划分为训练集、验证集、测试集。训练集(70%)用于模型的训练与参数调整,验证集(20%)用于模型的超参数优化与过拟合检测,测试集(10%)用于模型的最终性能验证,确保模型的泛化能力与评估精度。

(三) 深度学习评估模型构建

结合网络教学资源的多模态特性,构建多模态深度学习评估模型,主要包含特征提取模块、特征融合模块、评估预测模块三个核心部分,采用TensorFlow框架实现模型的搭建与训练。

1.特征提取模块:针对不同类型的资源,设计对应的特征提取子模块,实现多模态特征的自动提取。

文本特征提取:采用BERT模型作为文本特征提取器,针对课件文本、试题、案例等文本类资源,输入预处理后的文本数据,通过BERT模型的预训练与微调,提取文本的深层语义特征(如知识点相关性、内容准确性、逻辑连贯性),输出文本特征向量。BERT模型能够有效捕捉文本的上下文语义信息,提升文本特征提取的准确性,适用于网络教学资源文本内容的特征分析。

图像特征提取:针对课件中的图片、视频帧等图像类数据,采用CNN(卷积神经网络)模型(如ResNet50)进行特征提取。通过卷积层、池化层的层层处理,提取图像的纹理特征、色彩特征、清晰度特征等,输出图像特征向量,用于评估图像的质量与适配性。

视频特征提取:将视频分解为帧序列,结合CNN模型提取每帧的图像特征,同时采用LSTM(长短期记忆网络)模型提取视频的时序特征(如画面连贯性、内容完整性),将图像特征与时序特征融合,输出视频特征向量,实现对视频类资源的全面特征提取。

音频特征提取:针对视频中的音频、语音类资源,采用MFCC(梅尔频率倒谱系数)提取音频的特征参数,结合CNN-LSTM模型,提取音频的清晰度、流畅度等特征,输出音频特征向量,用于评估音频质量。

2.特征融合模块:采用注意力机制(Attention)实现多模态特征的融合,将文本、图像、视频、音频等不同模态的特征向量进行加权融合,突出重要特征的作用,抑制无关特征的干扰。首先,对各模态特征向量进行标准化处理,

消除量纲差异;然后,通过注意力机制计算各模态特征的权重,根据特征的重要性分配不同的权重系数;最后,将加权后的各模态特征向量拼接,形成统一的全局特征向量,作为评估预测模块的输入。

3.评估预测模块:采用全连接神经网络(FCN)作为评估预测模块,输入融合后的全局特征向量,通过多层全连接层的计算,输出资源的各维度评分与综合质量得分。模型的损失函数采用均方误差(MSE),用于衡量模型预测得分与人工标注得分之间的差异,通过梯度下降算法(Adam算法)对模型参数进行迭代优化,最小化损失函数,提升模型的预测精度。同时,在模型中加入Dropout层,抑制过拟合,提升模型的泛化能力。

(四) 模型验证与优化

1.模型验证:采用测试集对训练完成的深度学习评估模型进行性能验证,选取平均绝对误差(MAE)、均方根误差(RMSE)、决定系数(R^2)作为模型性能评估指标,其中MAE、RMSE越小, R^2 越接近1,表明模型的评估精度越高。验证结果显示,模型在测试集上的MAE为2.35, RMSE为3.12, R^2 为0.92,表明模型的预测结果与人工评估得分高度吻合,能够实现对网络教学资源质量的精准评估;同时,针对不同类型、不同学科的资源,模型的评估精度均保持在较高水平,证明模型具备良好的泛化能力。

2.模型优化:针对模型验证过程中出现的问题,进行针对性优化,进一步提升模型性能。一是优化特征提取模块,对BERT、ResNet50等模型进行微调,调整模型的超参数(如学习率、批次大小、卷积核数量),提升特征提取的准确性;二是改进特征融合机制,引入改进型注意力机制(如Multi-Head Attention),优化各模态特征的权重分配,提升特征融合效果;三是增加数据增强操作,对训练集数据进行扩充(如文本同义词替换、视频帧旋转、图像亮度调整),缓解数据不平衡问题,进一步抑制过拟合;四是优化模型结构,减少全连接层的层数与神经元数量,提升模型的运行效率,确保模型能够适应海量网络教学资源的快速评估需求。优化后,模型的MAE降至1.98, RMSE降至2.76, R^2 提升至0.94,模型性能得到显著提升,能够满足实际应用需求。

三、结束语

本文围绕网络教学资源质量评估的实际需求,构建了

科学、全面、可量化的评估指标体系，明确了评估维度、具体指标及量化标准，为资源质量评估提供了理论基础；在此基础上，设计并实现了基于深度学习的网络教学资源质量评估模型，通过多模态特征提取、特征融合与评估预测，实现了对网络教学资源质量的自动化、精准化评估，有效解决了传统人工评估的弊端。实验验证表明，该模型评估精度高、泛化能力强，能够适应不同类型、不同学科的网络教学资源评估需求，为网络教学资源的规范化管理、优质化建设提供了有效的技术支撑。同时，本文研究也存在一定的不足：一是模型的训练数据虽涵盖多种类型的资源，但部分小众学科、特殊类型资源（如虚拟仿真资源）的样本量较少，可能影响模型对这类资源的评估精度；二是模型对资源内容的深度语义理解仍有提升空间，对于复杂的教学逻辑、知识点关联性的评估不够精准；三是模型的运行效率虽经过优化，但在处理海量资源（如百万级、

千万级资源）时，仍需进一步提升速度。未来的研究方向主要包括三个方面：一是扩充训练数据集，增加小众学科、特殊类型资源的样本量，进一步提升模型的泛化能力；二是融合自然语言处理、计算机视觉等前沿技术，优化特征提取模块，提升模型对资源内容的深度语义理解能力，实现更精准的评估；三是对模型进行轻量化优化，结合边缘计算技术，实现模型的快速部署，满足海量网络教学资源的实时评估需求，推动网络教学资源质量评估的智能化、高效化发展，助力教育信息化的高质量推进。

参考文献：

[1] 方伟.线上线下混合式教学的教学评价创新研究——以《网络数据分析与应用》课程为例[J].教育教学研究前沿, 2025, 3(10): 94-99.

[2] 李芳.计算机技术在网络教学管理中的应用[J].国家通用语言文字教学与研究, 2019, (6): 22-23.