

基于多模态数据分析的在线学习智能评估反馈

□文 / 李睿，刘子韬

(北京世纪好未来教育科技有限公司，北京 100080)

摘要：基于互联网新一代信息技术的飞速发展，在线学习在我国基础教育中蓬勃发展，已经成为推动教育创新的重要力量。在线学习为学生提供了时空灵活性，丰富的学习资源共享以及全场景的行为数据化记录。然而，由于在线课堂发生在虚拟的网络空间中，其教学质量参差不齐的问题也日渐暴露。因此，我们提出一套合理的在线课堂智能质量评估反馈体系。该评估体系基于真实的多模态在线课堂数据，覆盖不同的在线教学形态，能够从老师视角和学生视角，全面考虑和衡量质量评估核心要素，进而实现在线教学的精准评估反馈。

关键词：在线学习；智能评估；多模态数据分析

中图分类号：G43 **文献标志码：**A **文章编号：**2096-5036(2022)02-0068-07

DOI：10.16453/j.cnki.ISSN2096-5036.2022.02.008

0 引言

在线教育是 20 世纪 90 年代逐渐兴起并飞速发展的教育创新领域，是我国教育信息化促进教育现代化的重要发展阶段，是信息技术对教育产生变革影响的典型实践。在疫情防控期间，我国大部分基础教育中小学校，依托网络构建的学习环境，运用云计算、学习分析、物联网、人工智能、网络安全等新技术，跨越学校和班级的物理限制，面向学习者个体，提供优质、灵活、个性化的新型教育服务。在线学习在为学生提供了时空灵活性和丰富的学习资源的同时，其教学质量参差不齐的问题也日渐暴露。相比于传统的线下教学，一些老师由于缺乏在线授课经验，导致学生缺乏充分的在线形式的交互，学习单调枯燥。同时，整个在线课堂都被记录和复制，参差不齐的在线内容和教学质量，在传播广、覆盖面大、无地域限制的互联网环境中，一旦出现问题，会被无限放大，带来重大负面影响。因此，我们提出一套合理的在线课堂智能质量评估反馈体系。该评估体系基于真实的多模态在线课堂数据，覆盖不同的在线教学形态，能够从老师视角和学生视角，全面衡量质量评估核心要素，进而实现在线教学的精准评估反馈。

基金项目：科技创新 2030—“新一代人工智能”重大项目——“智慧教育人工智能开放创新平台”项目（2020AAA0104500）；
北京市科技新星项目（Z201100006820068）

1 中国 K12 在线课堂教学与评价的现状

在线课堂教学是指学习环境依托于新一代互联网构建，整个学习流程都在线上开展的教学模式。相比于传统的线下教学，在线课堂中的学习活动、教师指导、学习资源都依赖于互联网信息技术的支持。在互联网音视频技术日益成熟的推动下，在线课堂教学为学生提供了时空上的灵活性、学习资源普遍的共享化，以及学生在整个课堂中学习行为的数据化。同时，在疫情防控的助推下，学生的学习场景逐渐从线下向线上迁移，学生的学习使用习惯和家长对在线课堂的认知大幅提升，在线教学的行业渗透率迅速提升。

在线教育的快速发展是一把双刃剑，它在解决了教育资源普及问题的同时，其教学质量参差不齐的问题也日渐集中和暴露，具体体现在如下两个严峻的现实问题^[1,2]。第一，老师在线上授课的效果有待提升。在疫情防控的紧迫压力之下，很多授课老师紧急从自身熟悉多年的线下课堂，转为线上的教学授课，其中很多优秀的线下教学的经验，并无法直接复制到在线课堂，导致老师和学生的交互匮乏，学习体验形式单一，在线学习的效果大打折扣。第二，在线课堂里所有的教学过程均可被记录和复制，参差不齐的在线内容和教学质量，在传播广、覆盖大、无地域限制的互联网环境中，一旦出现问题，会被无限放大，带来重大负面影响。如何保障当前在线教学内容合规且高质，成为当前面临的棘手问题。

因此，在线教学能够长期存在和发展的前提是需要一套合理的在线课堂质量评估反馈体系。该评估体系需要能够基于真实客观的在线课堂数据，覆盖当前多种多样的在线教学形态（在线大班、在线小班、在线 1 对 1 等），同时从老师视角和学生视角，全面衡量和考虑质量评估核心要素，进而实现在线教学的精准评估反馈^[3,4]。本文将以好未来集团研发的基于多模态数据分析的在线学习智能评估反馈系统，围绕当前中国 K12 在线课堂教学评价的现状、智能评估反馈的基本框架原理和关键技术等问题展开分析，阐述多模态数据分析技术在在线教学评价中的应用，探索未来更多的可能性。

2 在线学习智能评估反馈国内外研究现状

按照在线学习的形式，国内外关于在线学习的智能评估反馈相关研究工作可以主要分为两大类——基于类 MOOC 的录播视频在线学习的评估反馈、基于实时音视频的在线学习的评估反馈。在基于类 MOOC 的在线学习平台上，平台会记录学生的行为和活动日志，进而用于分析学生的参与程度和课堂的质量。比如，Qiu 等通过对学生个体数据和其在课程论坛里的学习活动规律进行建模，提出用隐式动态因子图预测学生的学习行为和课堂质量^[5]。Ramesh 等研究学生在 MOOC 平台上的不同细粒度的行为并研发了一个隐式的表征模型，用于抽取学生不同的参与程度，进而预测学生的学习状态^[6]。在基于实时音视频的在线学习评估反馈工作中，借助音视频的实时交互能力，老师和学生在整个在线课堂的全场景数据能够被有效地记录，进而提升在线学习评估反馈的覆盖度和准确性。例如，Chen 等基于在线课堂的多模态场景数据，搭建了在线课堂违禁词

自动检测系统，以及整个课堂质量的自动预估系统^[7]。

尽管国内外的研究者对于在线学习的智能评估反馈，已经开展了诸多相关研究工作，但其涉及的评估反馈能力较为离散和单一，缺乏对整个课堂情况的整体把控，对于老师和学生的综合考量缺失，没有形成体系化的端到端、可复制的模型框架。

3 在线学习智能评估反馈框架

一套完备的在线课堂教学质量评估框架体系是在线课堂能够长期存在和发展的前提。该框架需要同时覆盖以教师表现为核心质量评估要素的千百人同时在线的大班课堂，以及以学生反馈为核心评估要素的十人以下的在线小班课堂。我们依托科技创新2030—“新一代人工智能”重大项目——“智慧教育人工智能开放创新平台”项目，利用业界领先的多模态学习技术，分析研究了数十万小时的在线课堂，尝试通过自动化和系统化的手段，对在线课程的质量做合理的评估，提出了基于多模态数据分析的在线学习智能评估反馈系统框架，如图1所示，相关的系统实现落地在好未来集团多条产品线，实现了在线课程的全面质量监控，并且匹配了后续的评估和处理机制。在学生满意度达到行业领先水平的同时，教师管理队伍的规模和工作效率也达到了领先水平。

如图1所示，本文提出的在线课堂教学评估反馈系统理论框架，从基础课质、教学行为、互动反馈三个层次，全面衡量在线课堂的教学质量。



图1 在线课堂教学评估反馈系统理论框架

在基础课质层，我们需要对老师的基本表现有一定的约束和要求，避免“负面行为”。具体而言，我们将基础课质的评估分为两个纬度——教学仪态和教学红线。在教学仪态方面，系统会评估授课老师的教学环境，比如嘈杂或者昏暗的教学环境等。在教学红线方面，系统会检测老师着装、行为、语言的规范性，包括但不限于，教师个人不规范的着装，对学生进行类似“笨死了”“这么简单都不会做”等负面的表达。类似上述这些不规范行为会对学生的心产生十分负面的影响，对其检测是智能监课系统的基本功能，也可以叫做保底功能。

在教学行为层，系统需要对教学效果有帮助的积极行为，进行精准的检测。展开而言，在线课堂的教学行为可以分为具体的教学动作以及教师多年沉淀的教学技巧。教

学动作层面，典型的积极行为包括对学生给予正面激励，提醒学生完成课堂行为或流程，以及对课堂活动的有效组织。比如，老师在课堂上表扬孩子，提醒学生记笔记或者使用草稿纸，都是积极的课堂行为。在教学技巧层面，老师在多年教学经验中沉淀的教学技巧，比如用连续激昂的情感、丰富的肢体动作等，都能够带来积极正向的教学效果。这些积极的教学动作和教学技巧在不同年级、学科的课堂上会被赋予不同的权重，通过不断学习找到老师积极行为的语言特征，最终对课堂质量这个指标数据化。

在互动反馈层，系统质量评估的核心需要从老师转向学生。课堂学习的核心是老师和学生之间能够产生充分的沟通。只有在师生充分互动交流的基础上，学生的学习效果才能得到充分的保证。我们根据大量的实践分析和经验沉淀，将在线课堂中的师生反馈归结为如下三个方面。

一是表达。表达是指学生于在线课堂中要与老师有充分的互动。简单而言，学生要讲话，而不仅是听讲。对于在线课堂，学生只听讲而不与老师互动，那么在线课堂与看录播视频并无二致。而学生在课堂上说话也是有不同层次的。首先，可以判断的是句子的长短，“嗯、啊、知道了”这种短句子的比例如果过高，说明学生并没有实质性地参与到课堂互动中。其次，学生的疑问句比例非常重要，学生在课堂上能够提出问题，哪怕是“老师这里我没听懂”，也体现了在线互动的核心价值。最后，学生和老师的对话频率，有来有回的对话比单方面的提问回答效果更好。当然，需要在这里对互动内容有最基础的判断，比如上课聊游戏，除了活跃气氛，对教学效果而言不一定是价值最大化的选择。

二是动笔。动笔指的是学生于在线课堂中要有体现教学目的的书写内容，主要包括做题和笔记两个维度。书写的数据收集既可以采用拍照的方式，也可以采用手写板的方式。我们在实践中选择的是真实纸张和手写板的组合。在手写板上放置真实的A4纸，书写的过程就会被系统记录，既可以实时展示给老师，又可以作为课程质量分析的依据。在做题和笔记的分析上，我们的核心探索方向是找到这个行为与教学效果之间的连接点。比如记笔记，记录本身就比不记录要好得多，而记录的方式多种多样，整洁和有逻辑性是重点。至于记录得是否准确，记录得是否详实，实际上与教学效果间的因果关系没有那么强烈。在动笔的维度上，我们更关注用技术的方式实现“做了”和“合理”，通过这种方式达到最基本的教学要求。

三是总结。总结指的是学生于在线课堂中要有总结的环节。除了在课堂环节中专门设置“总结”环节外，从经验上看，我们关注到两个在教学效果中有突出表现的行为——第一，当出错的时候，学生对错误的原因进行分析，既可以是口头的分析，也可以是笔头的记录，错题本其实是这个行为的自然延伸；第二，每节课学生能够根据自己的理解梳理出本节课的思维导图。对思维导图的判断可以类比对笔记的判断，有比没有好，逻辑清晰结构合理比混乱好。对于识别方式，我们要求学生统一用大括号表示思维导图的层级，实现系统对思维导图层级的判断，以及每个层级内部的详实度的判断。

上述阐述的三层评估反馈体系，能够涵盖现所有主流的在线课堂学习形式，同时对授课老师和学生进行全方位、细粒度的数据量化，进而实现在线学习智能评估反馈的整体目的。

4 在线学习智能评估反馈关键技术

为了让在线教育长期健康可持续地发展，上述在线学习智能评估反馈框架需要能够以一种高可用的方式具体实现，其中包括在线课堂多模态的音视频数据的流转以及人工智能算法的精准检测与分析。

基于真实在线课堂的音视频多模态使用场景，我们设计了如下的通用多模态数据技术框架，如图 2 所示。在该通用技术框架下，可根据不同在线学习平台，不同在线学习形式（在线大班、在线小班、在线 1 对 1）的具体场景和需求，进行人工智能算法模型的运行，进而实现在线课堂智能评估反馈。具体而言，通过移动终端、PC、手写板、摄像头等教育场景下常见的终端设备，采集教师行为、学生行为有关的多模态数据，如教师讲课的视频和音频、学生听课状态的视频、教师批改作业的行为日志（如批改时间、时长）等。采集到实时数据后，将这些实时数据通过数据回流方式存储在数据库中，与教师、学生等画像数据，共同作为离线模型实验的数据来源。同时，通过批量数据标注，构建训练、验证、测试数据集，结合研究目标，评估相关的模态数据，明确模型的输入、输出和评价指标。

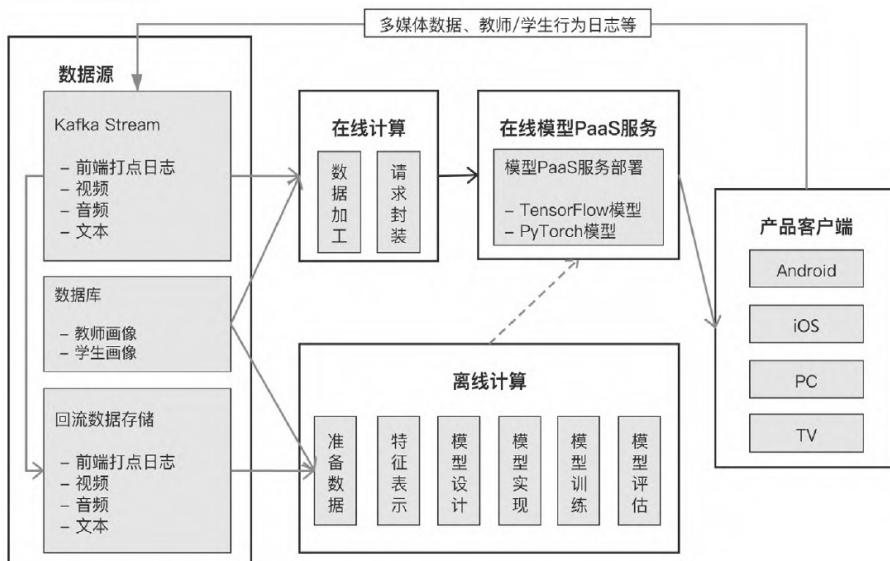


图 2 多模态在线课堂数据处理和人工智能算法部署通用框架

在人工智能的算法模型设计方面，不同维度不同层次的智能评估反馈，依赖的多模态数据和相关的算法模型也均有不同。整体而言，基础课质层，集中以计算机视觉和智能语音技术为主的感知层应用算法为主，而教学行为层和互动反馈层，则集中以自然语言理解技术为主的认知层应用算法为核心。同时，在算法分析的过程中，我们要求在线学习的智能评估反馈不需要有人工的参与，是一个完备的端到端解决方案。

具体而言，对于基础课质层，在自动检测教学仪态和教学红线的过程中，我们针对师生的在线直播的音视频流，进行实时的动态截图处理，抽取在线课堂的教师和学生侧的图像，其中包括师生的人脸图像以及当下的教学和学习环境的背景图像。然后，我们

利用基于注意力机制的卷积神经网络对相关的图像进行自动 AI 分类分析，进而达到相关维度的自动检测，比如人脸检测、着装检测等，相关精度高达 95% 以上。对于教学行为层和互动反馈层，我们分别利用前沿的智能语音和自然语言理解技术，对教师在课中讲述的每一句话或者师生之间的整段对话，进行实时多模态分析。具体而言，我们使用语音自动分割技术，将老师的授课音频进行实时的断句切分，然后再利用语音识别技术，同时保存老师说话的文本内容。最后，利用大规模预训练语言模型，我们对教师的内容进行深层次的语义编码转换，然后再通过若干神经网络的深度语义特征抽取，最终实现不同的教学动作和教学技巧的精准识别，精度达到 80% ~ 92%。值得一提的是，在诸如教师情感、流利度的智能评估中，我们同时结合了音频和文本中的有效信息，将两个模态进行联合学习并完成最终预测，相关精度达到 84% ~ 87%。

5 在线学习智能评估反馈实际应用效果

依托科技创新 2030—“新一代人工智能”重大项目——“智慧教育人工智能开放创新平台”项目，好未来集团集中研发了在线学习智能评估反馈整体解决方案，并将其落地在好未来学而思多条产品线中，不断提升、打磨、优化在线课堂的整体教学质量和体验。例如，在实际应用中，我们针对数十万节的 1 对 1 课堂引入了这一套智能评估反馈系统。如前文所述，我们在互动反馈这个方面，重点关注学生和老师之间的深层交互，基于表达、动笔、总结三个维度，提出“愿表达、勤动笔、善总结”的智能评估反馈驱动的教学理念，针对 1 对 1 在线学习制定了六个层面的质量评估标准。一是愿表达（问答）——学生上课能够积极、完整地回答老师的问题；二是愿表达（讲题）——学生能够完整讲述一道题的思路；三是勤动笔（笔记）——学生上课要进行笔记的记录；四是勤动笔（做题）——学生上课要完整地书写一道题的过程或结果；五是善总结（归因）——当发现错误时，学生要主动归纳错误的原因；六是善总结（导图）——在下课之前，学生要通过思维导图总结本节课的重点。

每一节课，针对上述的六个层面，在线学习智能评估反馈系统都会进行自动评判，在算法模型自动评判的基础上，我们随机抽取了 200 节课，对其判断结果进行了人工校验，其中算法判断和人工判断的一致性均在 90% 以上。基于在线学习智能评估反馈系统，我们可以第一时间发现个别没有达到优秀标准的课堂，可以针对问题有的放矢，在老师的选聘、培训方面提供更多的服务。

依托科技创新 2030—“新一代人工智能”重大项目——“智慧教育人工智能开放创新平台”项目，好未来集团已经将上述的基于多模态数据分析的在线学习智能评估反馈解决方案全面开放给全社会。

6 总结与展望

随着互联网技术与人工智能技术的不断发展，科技力量在基础教育领域的应用会

迎来更好的发展，也伴随着对学生的学习场景产生正面和深远的影响。对于在线学习的场景，如何进行高效的在线教学与课堂效果的评价，已经成为中国基础教育必不可少的一个环节。基于多模态数据分析的在线学习智能评估反馈技术，已经能够在在线课堂场景深入到教师教学和学生学习的部分场景和流程中，并起到相应的预警和提醒作用。在未来，这个场景会更加深入到学生学习的全流程。智能评估反馈技术，会逐步向学习者服务转化与过渡，尽可能多地为学生提供合适的教学服务。如何为学生提供全学习流程高质量的服务，是下一代教学产品和人工智能技术要解决的核心问题。少年强则国强，期待人工智能在推动中国基础教育进步上发挥进一步的作用。

参考文献

- [1] Li Hang,Yu Kang,Yang Hao,et al.A Multimodal Machine Learning Framework for Teacher Vocal Delivery Evaluation[C]// International Conference on Artificial Intelligence in Education.Cham:Springer,2021:251-255.
- [2] Huang,Gale Yan,Jiahao Chen,et al.Neural multi-task learning for teacher question detection in online classrooms[C]// International Conference on Artificial Intelligence in Education.Cham:Springer,2020:269-281.
- [3] Xu,Shiting,Wenbiao Ding,Zitao Liu.Automatic dialogic instruction detection for k-12 online one-on-one classes[C]// International Conference on Artificial Intelligence in Education.Cham:Springer,2020:340-345.
- [4] Liu Haochen,Zitao Liu,Zhongqin Wu,et al.Personalized Multimodal Feedback Generation in Education[C]//Proceedings of the 28th International Conference on Computational Linguistics.Enschede:ICCL,2020:1826-1840.
- [5] Qiu Jiezhong,Jie Tang,Tracy Xiao Liu,et al.Modeling and predicting learning behavior in MOOCs[C]//Proceedings of the ninth ACM International Conference on Web Search and Data Mining.California:ACM,2016:93-102.
- [6] Ramesh Arti,Dan Goldwasser,Bert Huang, et al.Learning latent engagement patterns of students in online courses[C]// Twenty-eighth AAAI Conference on Artificial Intelligence.Québec:AAAI,2014.
- [7] Chen Jiahao,Hang Li,Wenxin Wang,et al.A multimodal alerting system for online class quality assurance[C]//International Conference on Artificial Intelligence in Education.Cham:Springer,2019:381-385.



李 睿

任职于好未来集团，从事教学领域的产品设计工作。清华大学法学院学士、硕士。在培训行业深耕十余年，引领并推动在线教育行业的数次行业变革。参与设计在线录播课、在线直播课、双师课堂、在线1对1等多个行业主流产品。著有《产品之核》一书。



刘子韬

好未来集团 AI 技术总监，兼任国际人工智能教育协会执行委员。美国匹兹堡大学计算机专业博士。主要研究方向为多模态数据挖掘和机器学习，以及相关方法在教育场景的应用。在人工智能领域顶级会议和期刊上发表论文 70 余篇，国家授权发明专利 22 项。入选北京市青年人才项目、北京市科技新星项目。多次在国际顶级学术竞赛中获得冠亚军。



在线阅读(节选)