

基于大语言模型的在线学习者群体情感演化研究——以视频弹幕分析为视角

Research on the Evolution of Group Emotions in Online Learners Based on Large Language

Models from the Perspective of Bullet Screen Analysis

李菁菁¹, 邵鑫源¹, 王玥², 梁媛媛¹, 王达¹, 黄鑫^{1*}

¹南阳师范学院人工智能与软件工程学院

²河南开放大学教师网络学院

* huangxin@nynu.edu.cn

【摘要】 在线学习已经成为数字时代的主要学习模式。在线学习者情感实时分析与反馈可以改善学习体验与效果,对提升在线学习效果至关重要。然而,目前在线学习情感分析仍以问卷、评论等事后文本为素材,存在时效性差,情感分类细粒度不足、忽视群体情感变化视角等问题。本文基于大语言模型对在线学习视频实时弹幕进行分析,从群体角度出发对不同类型课程中弹幕反映出学习者的细粒度情感进行深入探讨,精准捕捉群体情感动态演化过程,总结不同类型课程对学习者的差异化影响,为提升在线学习者学习体验与优化在线学习环境提供了有益的参考。

【关键词】 大语言模型;弹幕;情感细粒度分析;群体情感

Abstract: Online learning has become a major mode of education in the digital age. Real-time emotional analysis and feedback of online learners can enhance learning experiences and outcomes, playing a critical role in improving learning effectiveness. However, current emotional analysis in online learning relies on post-event data like surveys and reviews, which suffer from issues such as poor timeliness, coarse emotional classification, and lack of attention to group emotional changes. This paper analyzes real-time bullet screen comments in online learning videos using large language models, exploring the fine-grained emotions expressed by learners from a group perspective. It examines how different course types impact learners' emotions, offering valuable insights for improving learning experiences and optimizing the online learning environment.

Keywords: Large language models, Bullet screen, Fine-grained analysis of sentiment, Group emotion

1. 引言

数字时代的在线学习已逐渐成为一种主要的学习形式,但由于在线学习环境下师生处于分离状态,教师难以实时捕捉到学习者情感动态变化,无法及时调整教学策略导致在线学习效果不佳。已有研究表明学习者情感变化会影响在线学习效果(Pan,X. et al.,2023),因此研究学习者在线学习过程中的情感变化有助于构建和优化在线学习环境。

近年来,研究者围绕在线学习情感分析展开了研究。Guo 等人通过问卷调查,分析学生对在线学习的态度(Guo, C. et al.,2020)。Anwar 等人利用自然语言处理技术分析学生评论、问卷及讨论内容,研究情感类别(Anwar, A. et al.,2023)。Feng 等人提出基于方面的卷积神

本研究受下列项目资助:教育部人文社会科学研究青年基金项目(项目号:24YJCZH135)、河南省科技攻关项目(项目号:242102211019)、河南省本科高校智慧教学专项研究项目“南阳师范学院建设智能化教学场景的探索与研究”、南阳师范学院校级自然科学研究项目(项目号:2025PY003)、南阳师范学院博士专项(项目号:2023ZX007、2024ZX010)、南阳师范学院校级教改项目(项目号:2024-JXYJYB-4)。

神经网络(CNN)和注意力机制长短期记忆模型,用于自动化识别学生情感(Feng, X. et al.,2020)。高彦太通过改进的 Text CNN 模型,对中国大学慕课平台的学习评论进行积极、消极及混合情感分类,扩展了情感分析的应用(高彦太,2022)。这些基于自然语言处理和深度学习的方法显著提升了在线学习中情感识别能力(Vistorte, A. O. R. et al., 2024)并推动了技术发展。尽管研究成果显著,但仍有局限性:一是现有方法多分析课后评论等滞后性数据,难以实时捕捉学习者当下情感状态,影响反馈时效性;二是情感分类维度多为正面(积极)、中性和负面(消极)的粗粒度维度,缺乏在线学习情境下更细粒度的情感类别,难以全面理解学习者复杂情感;三是研究多聚焦个体情感,较少关注群体情感动态及其对学习互动的影响,限制了学习体验的提升与环境优化。

视频学习已成为在线学习的主要形式(Noetel et al.,2021),学习者通过弹幕实时互动。弹幕以文本形式沿时间轴滑动,具有即时性和互动性,分析弹幕可实时捕捉学习者细粒度情感表达。相比传统 RNN 和 CNN 方法,大语言模型(LLMs)凭借强大的语言理解与情感生成能力,在细粒度情感分类及动态变化捕捉方面更具优势。本研究通过分析视频弹幕,采用 LLMs 对新授课和习题课中的学习者情感进行细粒度分析,为构建适配在线学习的视频情感反馈机制,优化教学设计和提升在线学习效果提供科学依据。

1.1. 在线学习情感分析

在线学习平台中,学习者常通过视频评论和论坛问答表达情感,但因师生分离、关注不足和反馈不及时等问题,学习效果仍有很大提升空间。Wang 和 Zhang 对在线学习社区评论进行情感分析(Wang & Zhang,2020),研究帮助深入了解学习者情感,并支持教师或平台调整教学策略。已有研究指出,细粒度情感数据支持优化个性化学习,提升效率并增强积极性(许桂芳 & 穆肃,2023)。当前情感分析多基于评论数据,因缺乏实时性,难以及时调整反馈,影响效果。已有研究多关注正面、中性、负面情感的粗粒度分析,而情感细粒度分析能更精准地识别情感细节及强度,如正面情感的“兴奋”“认同”,负面情感的“愤怒”“伤心”,为动态调整学习策略提供支持。

当前在线学习研究主要关注个体情感反应,但是在视频学习中,学习者的情感状态不仅受个人与视频内容的影响,还可以通过弹幕实时感知群体情感状态,群体情感可通过感染、认同等因素深刻影响个体情感(Chmiel et al., 2011)。在线学习群体情感分析能够研究情感的动态演化,揭示学习者群体情感的变化趋势。本研究基于在线视频弹幕数据,聚焦群体情感的实时细粒度变化,为在线学习平台提供有效的情感分析方法,优化学习环境。

1.2. 弹幕情感分析

弹幕具备即时反馈、动态互动和情感表达的特性,能实时记录学习者感受,有效解决实时反馈不足的问题。目前弹幕情感分析多采用情感词典方法。张腾等利用 Word2Vec 进行词嵌入实现分析(张腾等,2022)。在线教育中的弹幕情感分析面临非正式语言处理、多语种交叉和情感类别简化等挑战,特别是在处理中英日多语种融合的复杂弹幕时尤为明显。弹幕因字数简短、语气直接、语义碎片化,且夹杂学科术语和流行语(丁国栋 & 杜华,2021),在语序、内容和表达上与传统评论文本有显著差异,传统方法难以直接适用(李浩君等,2023)。本研究将结合实时数据分析与大语言模型,深入分析学习者情感,提升分析的准确性。

1.3. 大语言模型情感分析

GPT、讯飞星火、文心一言等大语言模型应用已经成为文本情感分析的重要工具。相较于传统的情感分析方法,大语言模型在情感分析中能理解深度语境情感,灵活处理多样化情感表达并且具备跨语种识别的能力,适应性强,能够处理较为复杂的情感。大语言模型通过预训练和微调过程,可以在没有大量标注数据的情况下自动化标注数据的情感以进行有效的情

感分析（张钦彤等，2024）。在调用大语言模型进行情感分析时，提示设计至关重要。提示应明确描述任务需求，避免使用模糊、不具体的表述。本研究将针对在线学习场景及情感分析需求设计并训练提示词，为进一步提升大语言模型在情感分析任务中的性能提供支持。

1.4. 研究目的和问题

研究目的：本研究旨在基于大语言模型方法探究在线视频学习中弹幕所体现的学习者群体情感的动态演化过程。具体而言，研究的核心目标包括 1.探讨在线学习视频中弹幕反映的学习者情感的细粒度类别。2.分析不同类型的课程（新授课与习题课）对学习者情感产生的不同影响。3.观察在线学习视频中学习者群体情感的动态演化过程。为此，本研究提出了以下三个研究问题，并结合数据分析得出相应结论：

- 问题 1： 在线学习视频中学习者情感细粒度类别有哪些？
- 问题 2： 学习者对不同类别的在线学习视频情感细粒度是否存在差异？
- 问题 3： 不同类别的在线学习视频是否对学习者群体情感的演化倾向产生影响？

2.研究设计

2.1.数据选取

本研究以哔哩哔哩学习型社区的考研高等数学教学视频为素材，此类视频学习者 为考研群体，有很强的学习动机，视频学习完成度高，弹幕能够反映学习者在视频学习中的真实情感。通过爬虫技术获取了 131 条视频数据（45 条新授课视频和 86 条习题课视频），总时长约 5240 分钟。提取视频中的弹幕数据，共计约 48.1 万条，其中新授课视频约 17.5 万条，习题课视频约 30.6 万条。经过人工清洗，剔除无关内容后，保留约 11.8 万条有效弹幕，其中新授课视频约 3.1 万条，习题课视频约 8.7 万条（包含中性弹幕）。

2.2.研究方法

本实验对两类视频的弹幕进行了情感细粒度分析。首先，使用讯飞星火大模型应用程序编程接口（API）对清洗后的有效弹幕进行情感词提取。输入提示词时，确保简明清晰并包含关键字，以提高查询的准确性和效率。随后，将相同情感词的弹幕进行聚类，并为每个情感词分配特定颜色。为确保相同情感词在多次运行中保持固定颜色，采用字典映射方法，通过字典的唯一性和高效查找机制，快速匹配情感与颜色，降低计算复杂度。最后，基于弹幕的情感、颜色和数量生成可视化图表。具体流程见图 1。

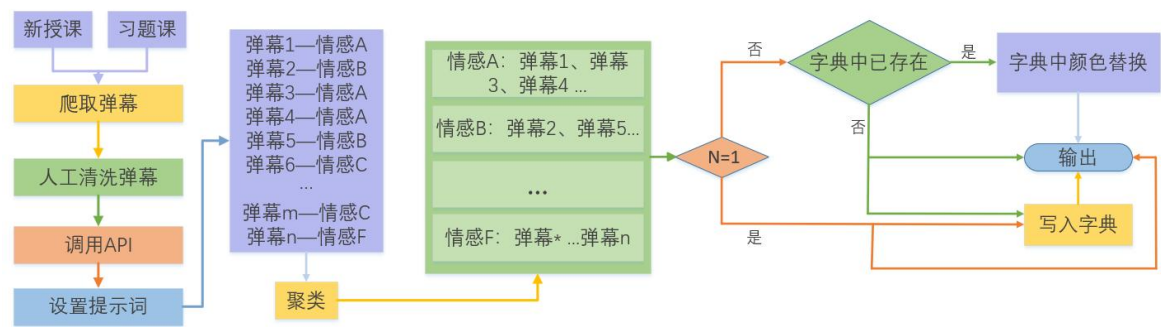


图 1：弹幕情感细粒度分析流程图

3.数据可视化

3.1.词云图可视化



图 2：词云图

本文以一节新授课和一节习题课弹幕为例，使用 jieba 中文分词工具进行分词，筛选高频名词，并通过词云图可视化弹幕中的核心内容和关键词频率。“中性”情感未展开讨论。新授课词云图（图 2a）显示，“函数”、“数值”、“方程”等关键词体现了学习者对核心概念的关注，弹幕主要围绕数学考研的核心知识点和理论阐释展开。“部分”、“建议”、“研究生”等词则反映了学习者对复杂概念理解和解决考研问题的需求。习题课词云图（图 2b）中，“题”、“导数”、“问”、“证明”等高频词突显其理论性和复杂性，“证明”和“导数”为教学重点，也易引发困惑。同时，“牛”、“思路”、“答案”等词与兴奋情感相关，表明学习者在理解难题或接触新思路时的激动与对教师讲解的赞扬。总体来看，新授课弹幕聚焦概念知识传授，习题课弹幕关注解题步骤和考试情境，反映了不同课堂模式下学习者的关注重点与情感反馈差异。

3.2. 聚类图可视化

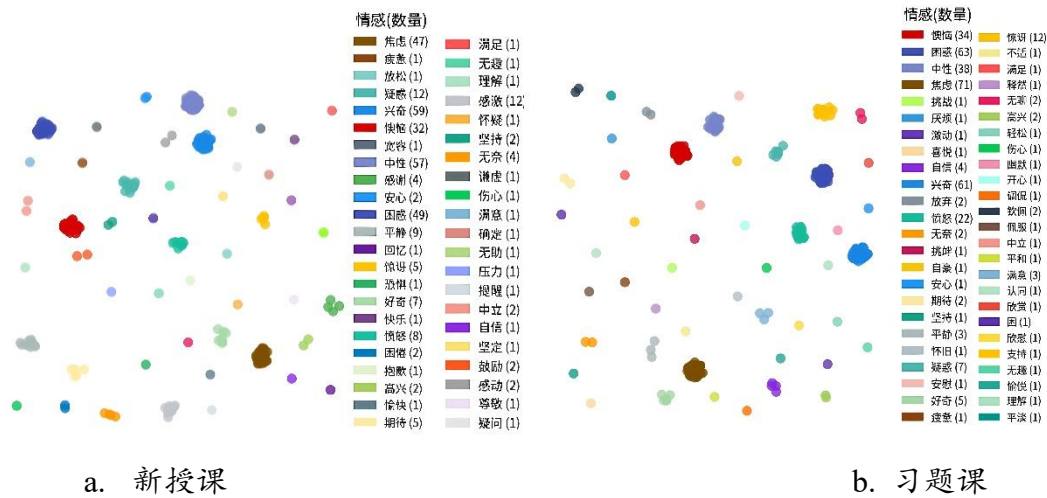


图 3：情感聚类图

图 3a 中的弹幕内容选自新授课某节视频完整有效的弹幕内容，图 3b 中的弹幕内容来自习题课真题某题的完整有效的弹幕内容。图中左侧的每个点代表一条有效弹幕，不同颜色对应不同的情感类别，相同情感的弹幕聚集在一起。为确保每种颜色与特定情感一一对应，采用了字典映射的方法。

4. 结果分析

4.1. 在线学习视频中学习者的细粒度情感类别分析

表 1：情感示例表

粗粒度情感	细粒度情感	示例
-------	-------	----

正面	兴奋、自豪、轻松、满意 自信、开心、佩服、认同	“芳芳真牛”（佩服）、“太神奇了”（兴奋） “一气呵成，令人拍案”（兴奋）...
负面	困惑、焦虑、愤怒、畏惧 疑惑、无奈、沮丧、疲惫	“我大抵是病了 极坐标现在就学吗”（无奈） “哭了，看 0 和 1 于是下意识在 1/2 展开了”（沮丧）...
中性	\ “积分不等式，积分的绝对值小于等于绝对值的积分” “无定义点和分段函数分段点”...	

表 1 显示新授课视频与习题课视频的正面情感集中于知识点掌握、教师讲解认可和能力提升，激发学习积极性。例如，“兴奋”表现对新知识的激动，“佩服”体现对教师的认可，“认同”反映对解题思路的接受，“轻松”和“自豪”则源于问题解决后的满足感。负面情感主要来自学习中的困惑和焦虑，如“疑惑”表达对知识点的障碍，“无奈”表现对复杂内容的无助，“沮丧”反映尝试失败的挫败，“疲惫”则来自学习压力。正负情感在两类视频中既有共性，也因教学内容不同而有所差异。不同类型视频中的学习者情感既存在共性，也表现出一些差异。

4.2.不同类别的在线学习视频中学习者情感细粒度的差异

根据图 3a 情感聚类图分析，弹幕中情感强度分布存在显著差异。高频情感如“焦虑”、“兴奋”、“懊恼”、“困惑”强度较高，反映学生对新授课程内容的高度关注与强烈表达；低频情感如“好奇”、“感动”强度较弱，多为短暂反应。高频情感的波动对课堂氛围和学生体验影响显著。积极情感如“兴奋”与较好的学习效果相关，表明对课程内容与讲解方式的认可；负面情感如“焦虑”和“困惑”可能影响学习效果，适度焦虑可促进学习，但过度焦虑需引导转化为学习动力。

根据图 3b 情感聚类图分析，弹幕中情感强度差异显著。高频情感如“焦虑”、“困惑”、“兴奋”强度较高，对学生心理状态和学习参与度影响显著；低频情感如“激动”、“喜悦”、“期待”强度较弱，多为短暂反应，未主导学习过程。积极情感如“兴奋”、“满意”、“自信”等与较高学习效果相关，反映学生对课程或教学方式的认可；负面情感如“焦虑”、“困惑”、“懊恼”可能产生负面影响，尤其高频“焦虑”和“困惑”可能表明课程难度或节奏过高。教师需平衡课堂情感氛围，激发积极情感并引导负面情感，优化学习效果。

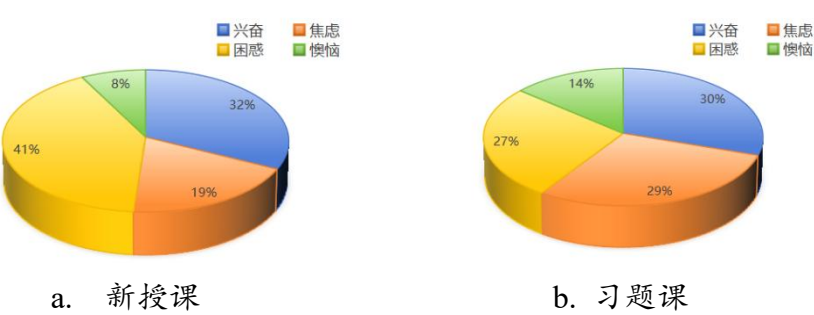


图 4：高频情感饼状图

在两类视频弹幕中，学习者的情感主要表现为“困惑”、“兴奋”、“焦虑”和“懊恼”。分别以新授课某节视频及习题课真题某题的完整有效的弹幕内容为例进行分析。通过对比两种不同课型，发现新授课的新知识点更能激发学生兴趣，这可能源于学生对未知知识的探索欲和对新内容的期待心理。而在习题课中，虽然学生仍表现出一定的兴趣，但程度略低，可能是因为习题课的内容更偏向知识应用，缺乏新鲜感。此外，与新授课相比，习题课中的“困惑”情感显著减少，这表明习题课能够有效解决部分学生的知识盲点，提高学生对知识的掌握程度。然而，习题课中的“焦虑”情感显著上升，这可能与学生对习题结果的高关注度有关。学生在习题课中更容易因为对解题正确性的压力产生心理负担，同时习题的难度和失败

体验可能导致部分学生产生懊恼心理。综上所述，新授课和习题课在激发学生兴趣和缓解困惑方面存在显著差异，而习题课在提高知识应用能力的同时，也可能增加学生的心理压力。

4.3. 不同类别的在线学习视频中学习者群体情感演化

图 5 展示了新授课一节视频完整有效的弹幕的类别，并将其划分为四个阶段：第一阶段代表课堂引入，第二阶段至第四阶段分别对应知识点 1 至知识点 3。针对这四个阶段，分别绘制了四张情感柱状图。新授课中，第一阶段引发较多“兴奋”和“焦虑”，学习者对新课程既期待又紧张。第二阶段“兴奋”显著上升，表明内容吸引力增强，但“困惑”增加，部分学习者遇到理解障碍，“焦虑”则因适应课程节奏而下降。第三阶段“兴奋”继续上升，但“懊恼”和“焦虑”增加，可能因复杂内容引发挫败感和紧张。第四阶段“兴奋”略有回落，但“焦虑”、“困惑”、“懊恼”显著减少，表明学习者逐渐专注于知识吸收，负面情绪减弱。

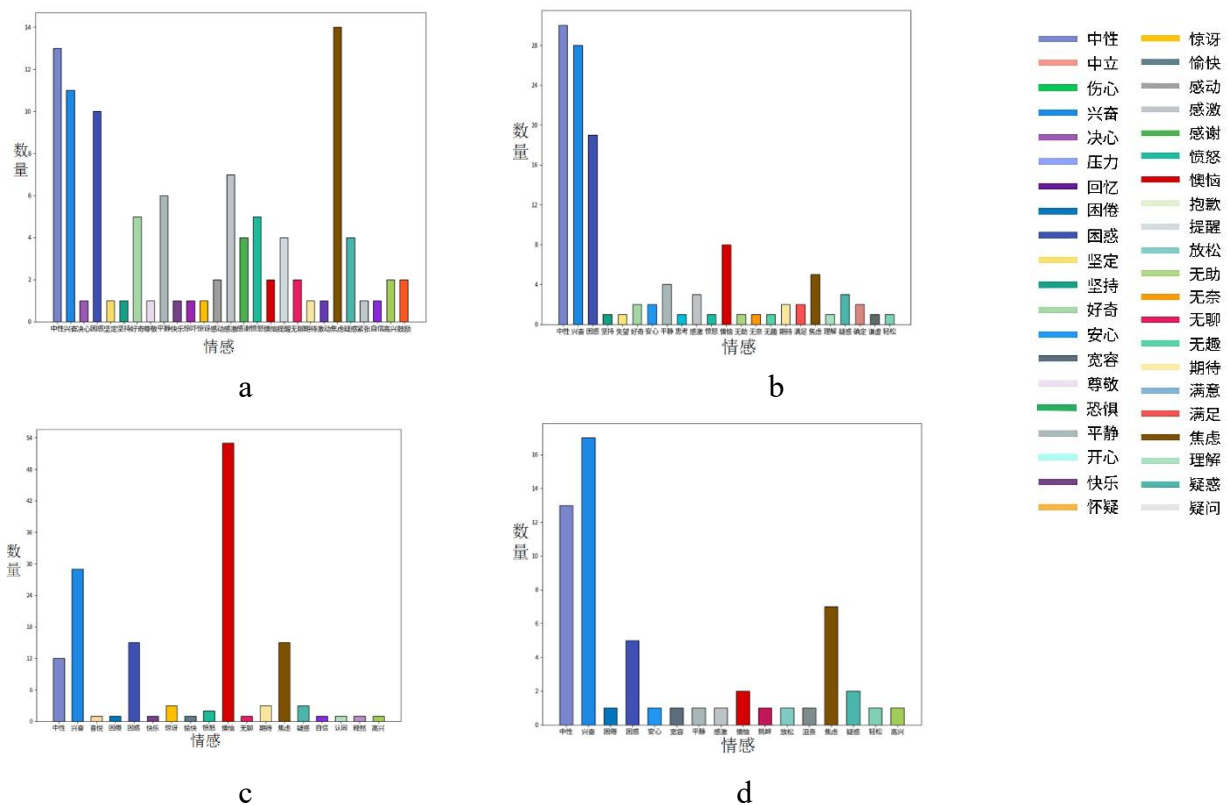
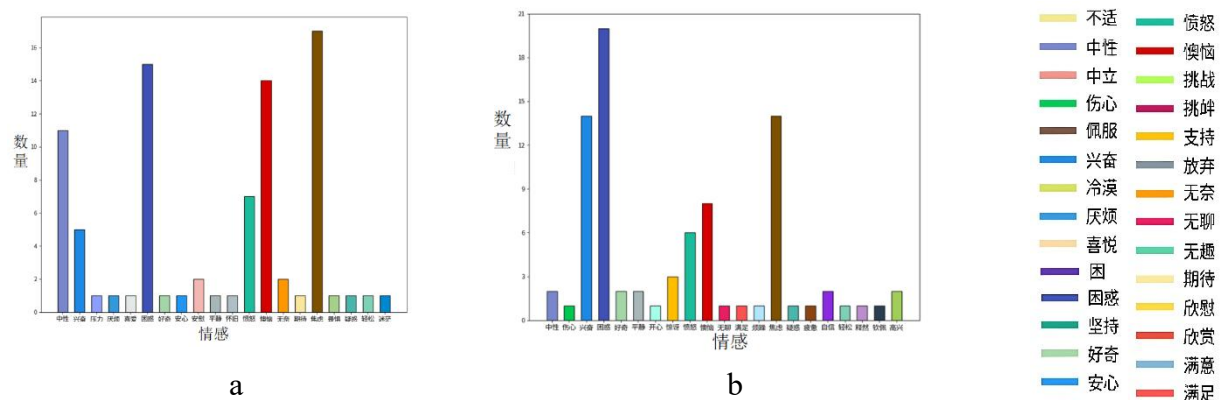


图 5：新授课情感演化柱状图



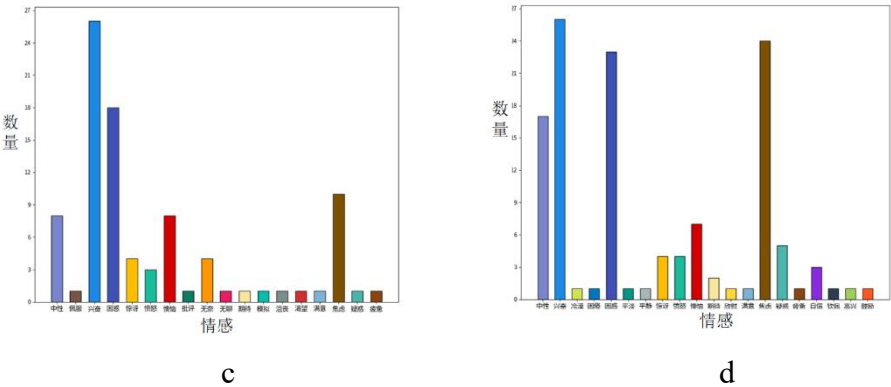


图 6：习题课情感演化柱状图

图 6 展示了习题课中一道真题的弹幕情感变化情况，共分为四个阶段：图 6a 为题目表述与引导阶段，图 6b 为第一小问讲解，图 6c 为第一小问第二种方法讲解，图 6d 为第二小问讲解。柱状图显示新授课与习题课中学习者情感的演化趋势。在习题课初期，弹幕以“困惑”、“焦虑”和“懊恼”为主，反映学习者对题目内容的不解和挫败感，“兴奋”较少，积极情绪低迷。随着讲解深入，“兴奋”逐渐增多，显示对题目理解的认可；“困惑”和“懊恼”减少，表明疑虑和沮丧情绪缓解。“焦虑”随讲解减少，但在后期，“兴奋”超越“困惑”成为主导情感，体现对问题解决的满意度。“焦虑”在最后略有回升，可能反映对题目难度的深层担忧。

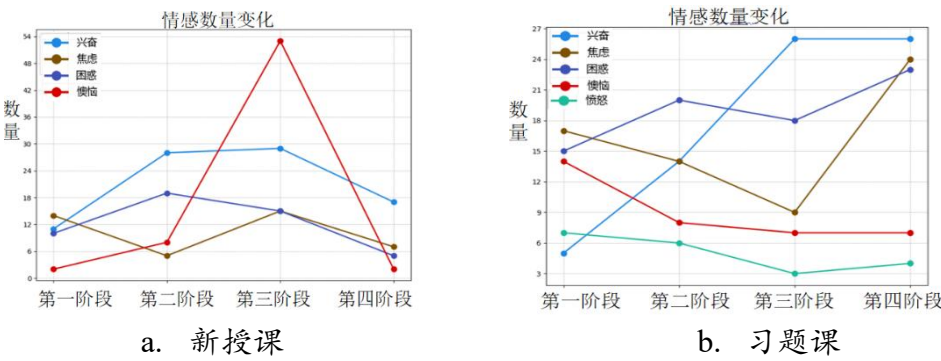


图 7：情感折线图

图 7a 展示新授课视频中高频情感词“兴奋”、“焦虑”、“困惑”、“懊恼”的曲线图。“懊恼”在第三阶段达到峰值后迅速下降，显示挫败感减弱；“兴奋”在第二阶段达到高点后逐渐减弱；“困惑”在第三阶段达峰值，第四阶段回落；“焦虑”整体较低，仅在第三阶段稍有上升。图 7b 为习题课情感曲线，“困惑”和“焦虑”主导初期情感，逐渐过渡到“困惑”和“兴奋”，尾声情感交织。总体而言，教师的讲解在两类视频中有效缓解负面情绪，并激发学习者积极性与参与热情。

5. 结论

本研究以在线学习视频弹幕为素材，分析学习者细粒度情感表达，克服了问卷和课堂评论等传统方法的滞后性，实现情感状态的实时捕捉。借助大语言模型的语言理解与情感处理能力，不仅支持细粒度情感分类，还能精准捕捉情感动态变化，为在线学习情感分析提供高效支持。通过对不同类型视频中学习者群体情感的分析，探索情感特征，为构建情感反馈机制提供理论依据，并为优化教学设计和提升学习效果提供科学指导。

针对在线视频学习中弹幕所体现的学习者群体情感的动态演化过程，本文首先对在线学习视频中学习者情感细粒度类别从正面、负面、中性的粗粒度情感进一步划分为“兴奋”、“自

豪”、“轻松”、“满意”等正面的细粒度情感；“困惑”、“焦虑”、“愤怒”、“畏惧”等负面的细粒度情感以及中性情感。其次，分析了新授课和习题课在激发学生兴趣和缓解困惑方面的显著差异，最后，通过分析相应数据直观说明了不同类别的在线学习视频对学习群体情感演化倾向的影响。

本研究在情感分析方面取得了一定进展，但仍存在局限性。首先，现有模型对讽刺、幽默等复杂情感的识别能力不足，难以准确把握语境和语气的细微变化。其次，学习者的情感反应受多种因素影响，如学习前的情绪、心理状态和外部环境等，但本研究尚未充分考虑这些因素。后续研究将加强模型复杂情感的识别能力，结合情境感知与个性化因素，提高模型的精准性和灵活性，并采用多种情感分析方法进行对比验证，以提升模型的有效性和适用性。本研究成果为提升在线学习环境的情感交互和学习效果提供了参考，并为在线教育平台的发展提供了支持。

参考文献

- 丁国栋 & 杜华.(2021).教学类视频弹幕文本的批判话语分析——以哔哩哔哩网站的某套高中物理教学视频为例.现代教育技术, (07), 72-79.
<https://doi.org/10.3969/j.issn.1009-8097.2021.07.009>
- 许桂芳 & 穆肃.(2023).个性化学习资源推荐中文本情感识别的作用及关键技术.中国电化教育, (05), 105-112.
- 张钦彤, 王昱超, 王鹤羲, 王俊鑫 & 陈海.(2024). 大语言模型微调技术的研究综述. 计算机工程与应用, 2024,60(17):17-33. <https://link.cnki.net/urlid/11.2127.tp.20240611.0941.002>
- 李浩君, 汪旭辉 & 廖伟霞.(2023).在线教育弹幕情感信息智能识别模型研究——融合变式情感词典与深度学习技术.现代远距离教育(01), 2023.0003. 19-31.
<https://doi.org/10.13927/j.cnki.yuan.2023.0003>
- 张腾, 倪渊, 莫同 & 吕学强. (2022). 弹幕视频的情感时间曲线聚类与传播效果. 数据分析与知识发现, 6(6), 32-45. <https://link.cnki.net/urlid/10.1478.g2.20220304.1544.004>
- 高彦太.(2022).基于深度学习的在线学习智能评价方法研究(硕士学位论文,西北大学). 硕士
<https://link.cnki.net/doi/10.27405/d.cnki.gxbdu.2022.000248>
- Anwar, A., Rehman, I. U., Nasralla, M. M., Khattak, S. B. A., & Khilji, N. (2023). Emotions matter: A systematic review and meta-analysis of the detection and classification of students' emotions in stem during online learning. *Education Sciences*, 13(9), 914-952.
<https://doi.org/10.3390/educsci13090914>
- Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A., & Hołyst, J. A. (2011). Collective emotions online and their influence on community life. *PloS one*, 6(7), e22207-22214.<https://doi.org/10.1371/journal.pone.0022207>
- Feng, X., Wei, Y., Pan, X., Qiu, L., & Ma, Y. (2020). Academic emotion classification and recognition method for large-scale online learning environment—Based on A-CNN and LSTM-ATT deep learning pipeline method. *International journal of environmental research and public health*, 17(6), 1941-1956. <https://doi.org/10.3390/ijerph17061941>
- Guo, C., Yan, X., & Li, Y. (2020). Prediction of student attitude towards blended learning based on sentiment analysis: Proceedings of the 2020 9th International Conference on Educational and Information Technology. [Conference]. IEEE, Qingdao, China,
<https://doi.org/10.1145/3383923.3383930>

- Noetel, M., Griffith, S., Delaney, O., Sanders, T., Parker, P., del Pozo Cruz, B., & Lonsdale, C. (2021). Video improves learning in higher education: A systematic review. *Review of educational research*, 91(2), 204-236. <https://doi.org/10.3102/0034654321990713>
- Pan, X., Hu, B., Zhou, Z., & Feng, X. (2023). Are students happier the more they learn?—research on the influence of course progress on academic emotion in online learning. *Interactive Learning Environments*, 31(10), 6869-6889. <https://doi.org/10.1080/10494820.2022.2052110>
- Vistorte, A. O. R., Deroncelle-Acosta, A., Ayala, J. L. M., Barrasa, A., López-Granero, C., & Martí-González, M. (2024). Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review. *Frontiers in Psychology*, 15, 1387089. <https://doi.org/10.3389/fpsyg.2024.1387089>
- Wang, K., & Zhang, Y. (2020). Topic sentiment analysis in online learning community from college students. *Journal of Data and Information Science*, 5(2), 33-61. <https://doi.org/10.2478/jdis-2020-0009>