

基于深度学习的学习情绪分析：方法与应用

朱颖雯¹, 张涛¹, 金鸽¹, 乔宇龙¹, 张迁^{1,*}

1. 江苏开放大学, 信息工程学院, 江苏 南京, 210036

摘要: 学习情绪作为调节学习过程与成效的核心非认知因素, 其精准识别与分析是落实《教育信息化 2.0 行动计划》中“发展智能教育”战略的关键突破口。本文基于深度学习技术构建面向教育场景的多维度情绪分析框架, 重点解决教育情境特异性(如课堂/MOOC 等)挑战。通过整合控制-价值情绪理论(Control-Value Theory)与多模态学习方法, 建立包含文本语义特征(讨论区发言)、视觉行为特征(眼动/表情)及交互行为特征(点击流)的三维情绪表征体系。基于此构建的“情绪-认知-行为”动态干预模型, 有效提升学习者课程完成率, 缓解在线学习的高辍学率问题。

关键词: 深度学习; 学习情绪; 多模态; 在线教育

Academic Emotions Analysis Based on Deep Learning: Methods and Applications

ZHU Ying-Wen¹, ZHANG Tao¹, JIN Ge¹, QIAO Yu-Long¹, ZHANG Qian^{1,*}

1. School of Information Technology, Jiangsu Open University, Nanjing, Jiangsu, China, 210036

Abstract: Academic emotions, as a core non-cognitive factor regulating learning processes and outcomes, requires precise identification and analysis to achieve the “developing intelligent education” strategy outlined in the “Education Informatization 2.0 Action Plan.” This paper constructs a multi-dimensional emotion analysis framework for educational scenarios based on deep learning technologies, with particular emphasis on addressing education-specific contextual challenges (such as classroom/MOOC environments). By integrating Control-Value Theory with multimodal learning methods, we establish a three-dimensional emotion representation system that incorporates textual semantic features (discussion forum posts), visual behavioral features (eye movements/facial expressions), and interactive behavioral features (clickstream data). The “emotion-cognition-behavior” dynamic intervention model built on this framework effectively improves learners' course completion rates and alleviates the high dropout rate problem in online learning.

Keywords: Deep learning; Academic emotion; Multimodal; Online education

伴随着教育信息化的深入发展, 学习情绪作为影响学习效果的关键非认知因素受到广泛关注。学习情绪涵盖学习过程中产生的享受、焦虑、疑惑、无聊等多类情绪状态, 不仅直接影响学习体验, 而且通过行为投入间接作用于学习成效^[1]。传统的情绪识别方法如问卷调查和人工观察存在主观性强、滞后性大和生态效度不足等固有缺陷, 特别在后疫情时代, 学习者因居家隔离和跨媒体不适等特殊状况, 更易出现困惑、焦虑甚至抑郁等负面情绪问题, 严重影响学习效率^[2]。近年来, 深度学习技术在情绪识别领域取得显著突破, 为解决教育场景的情绪分析难题提供新思路, 大语言模型通过检索增强架构^[3]和语音描述转化技术^[4], 将文本、语音等多模态情绪识别推向新高度。然而, 通用情绪识别模型直接迁移至教育场景, 识别精度往往大幅下降, 特别是区分“困惑”与“挫折”等学习特有情绪时, 凸显构建教育特异性情绪分析框架的必要性。

本文基于深度学习技术构建面向教育场景的多模态情绪分析框架,重点解决教育情境特异性(如课堂/MOOC等)与教育学有效性验证两大挑战。通过整合控制-价值情绪理论与多模态学习方法,建立包含文本语义特征(讨论区发言)、视觉行为特征(眼动/表情)及交互行为特征(点击流)的三维情绪表征体系,为智能教育系统提供理论指导与技术支持。

1 学习情绪分析理论基础

1.1 学习情绪的定义

学习情绪(Academic Emotions)是指与教学、学习和成就直接相关的情绪体验,是学习者在学习过程中产生与认知活动和学习成果密切相关的情感状态^[5-6]。学习情绪不仅影响学习者的认知加工过程,还通过调节学习动机、自我调节和学习策略的使用对学习效果产生深远影响。与一般情绪不同,学习情绪具有明显的情境特异性,更加关注与学习任务、学业成就和学习环境相关的情感体验,如课堂中的愉悦感、考试焦虑、学习兴趣和成就自豪等^[7]。

学习情绪可从多个维度进行分类。根据效价维度可分为积极情绪(如愉悦、希望、自豪和满足)和消极情绪(如愤怒、焦虑、羞耻和无聊);根据激活维度可划分为激活型情绪(如愉悦、希望、焦虑、兴奋和愤怒)和去激活型情绪(如放松、无聊、绝望和失落);根据情绪产生的核心对象可分为活动情绪(如学习过程中的享受、无聊、愤怒、沮丧)、结果情绪(如成功期望、失败焦虑、预期自豪、预期羞耻)。积极学习情绪通过认知资源分配机制显著影响学习投入度,且存在明显的阈值效应,表明学习情绪的精准分类和识别对制定有效的教学干预策略具有重要意义^[8]。随着人工智能技术在教育领域的应用,基于深度学习的多模态情绪识别方法为捕捉学习过程中的细微情绪变化提供新的技术路径,有望进一步丰富和完善学习情绪的理论分类体系^[9-10]。学习情绪分类如表1所示。

1.2 情绪认知理论

控制-价值理论由德国心理学家 Reinhard Pekrun 于 2006 年提出,用于解释学习情绪的产生机制及其对学习过程的影响^[11-12]。学习者对学习活动的控制感(对学习结果的可控性评估)和价值感(对学习任务重要性的主观评价)是诱发特定学习情绪的关键认知因素。近年来,控制-价值理论的研究范式已从传统课堂扩展到数字学习环境。Loderer 等人^[13]对 1965-2018 年间 186 项研究进行系统综述和元分析,探讨情绪(如愉悦、好奇/兴趣、焦虑、愤怒/挫折、困惑、无聊)的前因(控制-价值评估、先验知识、性别、技术学习环境特征)和结果(参与度、学习策略、学业成就)。Pekrun 等人^[14]提出三维分类法(效价×唤醒度×对象焦点),涵盖 12 类成就情绪(如愉悦、希望、焦虑、羞耻等),并验证其结构及与动机、学习策略、健康等的关系。此分类法为理解情绪在成就情境中的作用提供了全面框架,并揭示了情绪如何通过认知-动机机制影响学习、工作与健康。

1.3 教育场景中的情绪特殊性

教育场景中的情绪体验呈现明显的情境特异性和复杂的动态演变特征,源于教育环境特有的评价压力、社会互动和认知挑战。与一般情境中的情绪不同,学习情绪在微观表现和宏观发展轨迹上都具有显著的结构化特征。教育环境中的情绪通常与特定学科内容紧密关联,如数学焦虑和语言学习愉悦感等,这种学科特异性使得情绪识别需要结合具体的知识领域背景^[15]。同时教育场景中情绪的社会建构属性尤为突出,师生互动、同伴评价和集体课堂氛围共同形成了独特的“情绪生态系统”,系统中个体情绪不仅受自身认知评估影响,还受到群体情绪传染效应和社会比较机制的调节。情绪

的社会互构性使得单一维度的情绪识别方法难以准确捕捉教育场景中情绪的丰富内涵和动态变化。

此外，教育场景中的情绪呈现出明显的时序特征和认知关联性，尤其在数字化学习环境中表现更为复杂。学习过程中困惑-探索-理解的情绪转换过程与认知负荷理论密切相关，积极和中性情绪（如好奇、轻度困惑）促进认知资源有效分配，而强烈的负面情绪（如焦虑、无聊）则占用大量工作记忆容量，降低认知处理效率。特别是在线学习环境中，由于缺乏面对面互动的即时反馈，学习者更容易经历情绪波动，导致学习中断或放弃。这些复杂的情绪特征及其与认知过程的交互关系，为学习情绪分析方法提出了独特挑战。而深度学习技术凭借其强大的表征学习能力和时序数据处理优势，为捕捉教育场景中情绪的微妙变化和复杂模式提供了技术基础。

表 1 学习情绪分类

Table 1 Academic Emotions Taxonomy

分类维度	情绪类型	典型情绪	特征描述	认知过程	学习行为
效价维度	积极情绪	愉悦、希望、自豪、体验正向情感，增强满足	自我效能感和学习价值感	促进创造性思维，提高认知灵活性，优化资源分配	增强内在动机，提高任务坚持度，促进深度学习策略使用
	消极情绪	愤怒、焦虑、羞耻、体验负向情感，降低无聊	自我效能感和学习价值感	限制工作记忆容量，分散注意力，增加认知负荷	诱发回避行为，降低学习投入，增加表层学习策略使用
激活维度	激活情绪	愉悦、希望、焦虑、高唤醒水平，兴奋、愤怒	生理激活状态增强，精力集中	加快信息处理速度，提高警觉性，但可能降低精确性	增强任务参与度，提高学习强度，但可能降低学习持久性
	去激活情绪	放松、无聊、绝望、低唤醒水平，失落	生理激活状态降低，精力分散	降低信息处理速度，适合反思性加工，但可能导致注意力涣散	降低学习投入强度，可能增加反思性学习，但也可能导致学习惰性
对象维度	活动情绪	学习享受、无聊、愤怒、沮丧	与学习活动本身相关，受任务特性和学习环境直接影响	直接影响认知参与度和加工深度，调节注意力分配和资源使用	影响学习投入程度，调节任务坚持性，影响流体验形成
	结果情绪	成功期望、失败焦虑、预期自豪、预期羞耻	与学习结果预期和评价相关，受目标设定和成败归因影响	影响目标定向的认知过程，调节元认知监控和评估策略	影响学习目标选择，调节学习策略使用，影响面对挑战的行为模式
社会维度	自我情绪	自豪、羞愧、内疚、焦虑	聚焦于自我的评价和体验	影响个体信息加工策略选择，调节个人认知风格	影响自主学习行为，调节自我监控过程，形成独特学习习惯
	社会情绪	感激、同理心、嫉妒、羞辱	与他人互动相关的情绪体验	促进群体认知协调，影响社会元认知，调节协作信息加工	影响协作参与度，调节人际交往质量，塑造学习氛围

2 基于深度学习的学习情绪分析方法

2.1 基于 RNN 的情绪识别方法

基于循环神经网络（Recurrent Neural Network，RNN）的学习情绪识别方法通过序列建模有效捕捉情绪表达的时序特征与上下文依赖关系，较好地解决忽略上下文信息的传统情感识别网络的局限。如图 1 所示，该方法将文本数据（如学生反馈、课堂交互记录等）通过分词处理并转换为词向量表示，常用词嵌入方式包括 Word2Vec、GloVe 或字符级嵌入；向量序列被输入到长短期记忆网络

(Long Short-Term Memory, LSTM) 或门控循环单元 (Gated Recurrent Unit, GRU) 中进行处理; 通过注意力机制选择性聚焦于与情绪相关的关键词或句段, 并通过全连接层将高维特征映射到情绪标签空间。除单一 RNN 结构外, 研究者还探索了多层次上下文建模和多模态特征融合方法以增强 RNN 的表示能力。Hu 等^[16]提出 DialogueCRN 模拟人类大脑的信息处理过程, 将情绪分析分为感知和认知两个阶段, 感知阶段使用注意力机制匹配上下文线索, 认知阶段使用双向 LSTM 网络捕获上下文逻辑, 这种“感知-认知”过程使模型能够基于上下文线索推断情绪。Li 等^[17]设计双向情感循环单元, 建模对话中发言者之间的情感互动过程, 通过对局部上下文的有效利用获得优异的情绪识别性能。Wen 等^[18]提出 DIMMN 利用注意力机制分别捕获跨模态依赖关系, 同时引入时间卷积网络和 GRU 分别获取局部和全局上下文信息。

基于 RNN 的学习情绪识别方法序列处理能力天然适合捕捉情绪表达的时序演变模式, 能够有效建模情感状态的渐变过程; 门控 RNN 变体 (如 LSTM 和 GRU) 通过特殊的记忆机制缓解梯度消失问题, 能够捕捉长距离情感依赖关系; 模型结构相对简洁, 训练所需计算资源较少, 适合在资源有限的教育环境中部署。然而, 该方法序列处理的本质决定了其难以并行化计算, 处理长文本时效率低下; 在处理超长序列时面临长期依赖问题, 即使是 LSTM 和 GRU 也难以保持极长距离的情感上下文; 纯 RNN 模型在捕捉全局语义结构方面能力有限, 对于理解复杂学习情境中的隐含情感表达 (如讽刺、隐喻) 表现不佳; RNN 结构在建模多方对话中的情感交互时灵活性不足, 需要辅助机制来处理说话者转换和情绪流动。这也为后续基于图神经网络和 Transformer 的研究奠定了基础。

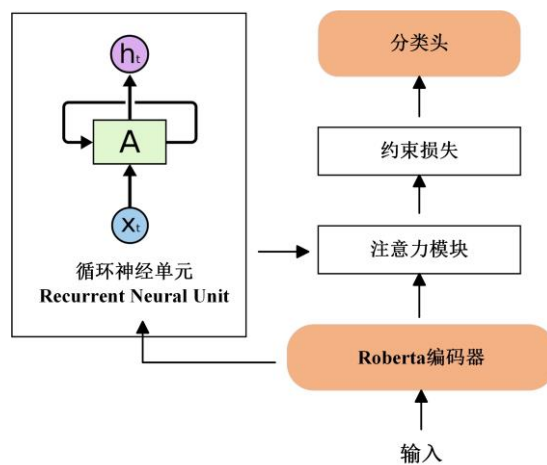


图1 基于 RNN 模型情绪分类任务
Fig.1 Emotion Recognition Method Based on RNN

2.2 基于 GNN 的情绪识别方法

基于图神经网络 (Graph Neural Network, GNN) 的学习情绪识别方法通过将学习互动编码为图结构, 有效建模多方交互中的情绪动态与依赖关系。与基于 RNN 的方法相比, GNN 能够明确确定不同上下文对当前话语的影响, 且获取的上下文不受距离限制, 从根本上解决长距离依赖问题。如图 2 所示, 其对学习情境中的对话或互动构建为图结构 $G=(V,E)$, 节点 V 表示话语、参与者或概念, 边 E 捕捉它们之间的关系, 边权重反映互动强度或情感依赖程度; 结合预训练语言模型 (如 BERT、

RoBERTa 等)提取的语义特征和交互元信息为每个节点初始化特征表示;图构建完成后通过多层图卷积操作递归更新节点表示;通过图池化操作或注意力机制聚合节点表示,并通过分类层映射到情绪标签空间,完成学习情境中的情绪识别任务。Ghosal 等人^[19]首次将图神经网络引入对话情绪分析,提出 DialogueGCN 模型,将对话视为话语节点和话语关系边的集合,通过独特的节点聚合方法连接话语的上下文信息以理解当前语境。图注意力网络(Graph Attention Network, GAT)将注意力机制引入图神经网络,不同于图卷积网络, GAT 将顶点之间的相关性纳入图结构,逐节点的计算方法使其适用于各种有向图的计算。Ishiwatari 等人^[20]在 DialogueGCN 和 GAT 基础上进行深入研究,增加了相对位置编码器,捕获话语在对话中的相对位置信息,反映话语的时间顺序及情绪的进展变化。然而,这种构建完全连接图的方法存在时空复杂度高的问题,且并非所有上下文都对话语情绪分析起积极作用。Shen 等人^[21]设计有向无环图(Directed Acyclic Graph, DAG)来构建对话结构,以更直观的方式建模话语附近的上下文信息,仅关注特定窗口内上方的话语,忽略下方话语。近年来,多模态信息与 GNN 的结合成为研究热点,如 M3GAT^[22]和 DCGCN^[23]。

基于 GNN 的学习情绪识别方法在处理复杂学习互动的情感分析方面表现出独特优势,其非欧几里得数据结构天然适合建模教育场景中的社交网络和多方交互,能够同时考虑学习群体的整体情感氛围与个体情绪状态;消息传递机制使模型能够捕捉情绪的传导效应,解决传统序列模型难以处理的长距离、非线性情感依赖关系;图结构的灵活性使其能够整合多模态信息(如文本、语音、面部表情等)和背景知识(如学科领域知识、社会关系等),提供更全面的情绪理解。然而,该方法也面临若干挑战,图构建过程通常依赖启发式规则,难以自动确定最优图结构;在大规模学习场景中,完整图的计算复杂度较高,可能导致效率问题;同时,图神经网络的表达能力受限于其聚合函数设计,在捕捉复杂情绪模式时可能不如特定设计的专用架构。

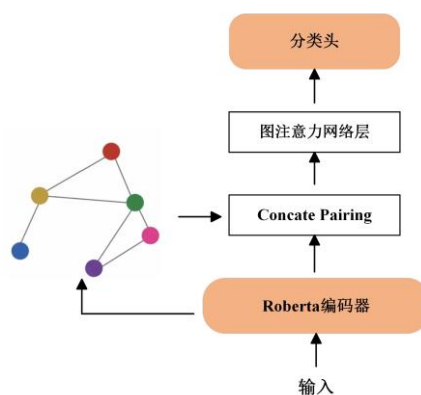


图2 基于 GNN 模型情绪分类任务
Fig. 2 Emotion Recognition Method Based on GNN

2.3 基于 Transformer 的情绪识别方法

基于 Transformer 的学习情绪分析方法采用 Transformer 架构,其核心是注意力机制,能够高效捕获序列内或序列间的依赖关系。如图 3 所示,其对学习场景中的文本输入(如课堂讨论、在线评论或学习反馈)进行预处理,包括分词、嵌入表示和特殊标记添加,在对话场景中,通过添加等特殊标记区分不同说话者;输入经过多层 Transformer 编码器处理,每层包含多头自注意力模块和前馈

神经网络，自注意力机制通过计算序列中每个位置与所有位置的相关性来捕捉上下文依赖关系；通过特定的分类层将 Transformer 提取的特征映射到情绪类别空间，完成学习情绪的识别。与 RNN 和 GNN 相比，Transformer 通过注意力机制可以有效捕获远距离上下文信息，从根本上解决了长期依赖问题；采用并行计算方式处理序列数据，大幅提高了计算效率；能够准确聚焦序列中的重要信息。Zhong 等^[24]首次提出知识增强型 Transformer 结构 KET，将外部知识库中的常识知识融入文本，通过设计的上下文感知情绪图注意力机制动态选择知识实体。Ghosal 等^[25]提出 COSMIC 框架使用 COMET 预训练编码器生成不同的常识向量，表示说话者意图、对说话者的影响、说话者的回应和对他人影响，能够捕获上下文状态、说话者内部状态、外部状态、行为意图状态与情绪之间的交互。BERT 及其衍生模型 RoBERTa 作为基于 Transformer 的预训练语言模型，适用于各种 NLP 任务，显著提升了对话情绪分析的性能。Transformer 的优势使其在多模态信息融合和知识集成方面表现出色。Li 等^[26]提出情感特征提取器 Emoformer 可有效提取三种模态的情感向量并融合为胶囊形式。Ma 等^[27]利用注意力机制和 Transformer 结构提出 SDT 模型，能自适应地整合多模态信息，在训练过程中引入硬标签和软标签知识，获得更准确的特征表示。Shen 等^[28]则结合图注意力机制捕获主题与对象之间的依赖关系，设计堆叠式跨模态 Transformer 层（CIT）捕获交互模态的特征关联。

基于 Transformer 的学习情绪分析方法中并行计算结构有效解决了传统 RNN 在处理长序列时的梯度消失问题，能够同时捕捉全局情感依赖关系；预训练-微调范式使模型能够迁移通用语言理解能力到特定情绪分析任务，极大减少了对标注数据的需求，且模型设计的灵活性使其易于整合多模态信息和领域知识，提升对复杂学习情境的理解能力。然而，该方法计算复杂度高，限制处理超长对话的能力；模型参数量大，资源受限场景下难以部署；解释性较差，难以向教育工作者提供直观可理解的情绪分析结果；此外，对于学科特定的情感表达（如数学焦虑、编程挫折）缺乏足够敏感性，在专业学习情境中表现有限。Transformer 的研究热度也推动了大型语言模型（Large Language Model, LLM）的兴起，如 GPT 系列、Gopher 和 Llama 2 等。这些拥有数十亿参数的大型语言模型无需针对特定下游任务进行微调和优化即可表现良好，但在情感分析等特定场景中仍有提升空间。以上三种方法的优缺点如表 2 所示。

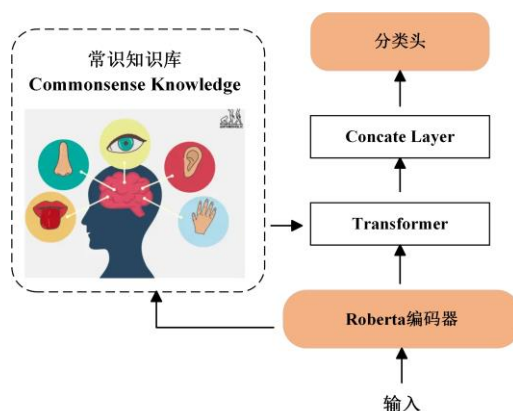


图3 基于 Transformer 模型情绪分类任务
Fig. 3 Emotion Recognition Method Based on Transformer

表 2 各种方法优缺点
Table 2 Comparison of advantages and weakness of each method

方法	基于 RNN	基于 GNN	基于 Transformer
优点	拥有最广泛的研究基础	清晰建模交互关系	丰富的变体
	具有历史记忆能力	考虑远程上下文	高计算效率
	包含时间信息	高可扩展性	捕获关键信息
缺点	长期依赖问题	高扩展成本	缺乏时间信息
	并行训练困难	动态建模能力差	模型复杂度高
	计算成本高		

2.4 基于大语言模型的情绪识别方法

基于大语言模型(LLM)的学习情绪识别方法充分利用预训练-微调范式与强大的上下文理解能力,实现更精确、更具语境感知的情绪分析。以 GPT 系列、LLaMA 或 BERT 等为代表的大规模语言模型通过自监督学习在海量文本上预训练,习得丰富的语言表示与世界知识;针对学习情绪识别任务,采取任务适应性微调策略,包括指令微调(Instruction Tuning)和上下文学习(In-context Learning)两种主要范式,通过精心设计的提示(Prompt)引导模型执行情绪分析;在特定教育场景中,研究者常采用领域适应技术,利用课堂对话、学习反馈等数据进一步优化模型,如使用参数高效微调方法中的 LoRA 或 P-tuning 等技术,仅更新部分参数以降低计算成本;为增强情绪识别的可靠性与稳定性,通常结合外部知识库(如情感词典、心理学知识图谱)或多模态信息(如学生表情、语音、生理信号等),通过知识增强型提示设计或检索增强生成(Retrieval-Augmented Generation, RAG)技术提升模型对学习环境特定情绪表达的理解能力,并通过思维链或自洽性等推理技术提供可解释的情绪分析结果。

基于大语言模型的学习情绪识别方法在教育场景中展现出独特优势与潜在挑战。预训练赋予的丰富世界知识使模型能够理解复杂、隐晦的情感表达,包括文化背景差异、特定领域术语和情感隐喻等,对把握多元化学习环境中的情感微妙变化至关重要;强大的上下文理解能力使模型能同时考虑历史对话、学习任务性质和社交动态等因素,提供更全面的情绪分析;模型的通用性与灵活性使其能通过简单提示适应不同教育场景和多样化情感分析需求;基于自然语言生成的情绪分析结果具备更好的可解释性和实用性,能直接作为教育干预的依据。

3 教育场景应用

3.1 多模态在线学习场景下的学习情绪识别

多模态在线学习场景下的学习情绪识别需同时应对学习环境复杂性与情绪动态变化的双重挑战。本文提出一种整合控制-价值情绪理论与多模态深度学习的教育情绪分析框架,旨在捕捉并解析在线学习过程中的复杂情绪动态,框架如图 4 所示。该框架通过构建三维情绪表征体系实现综合分析,在文本语义维度,利用 BERT、RoBERTa 等预训练语言模型分析学习者在讨论区与聊天室的发言内容,提取情感极性、认知投入度和社会临场感三类关键语义特征;在视觉行为维度,借助计算机视觉技术捕捉学习者的眼动模式(包括注视持续时间、扫视频率等)、面部微表情(通过 AU 单元激活序列表征)及姿态变化,构建视觉情绪特征空间;在交互行为维度,通过学习管理系统记录的点击流数据、交互时序与内容选择偏好等构建学习行为图,并应用图神经网络提取反映学习参与度与策略选择的高阶特征。这三维特征通过多模态注意力融合机制整合为统一情绪表征,并经由面向教育

情境的层次化分类器映射到学习情绪空间, 考虑个体差异与学习环境特性, 实现情境自适应的情绪识别。

面对在线教育场景的多样性与特异性挑战, 采用迁移学习与场景自适应机制, 确保情绪分析的教育情境针对性。针对三类典型在线教育场景设计特定的特征提取方案与分析模型。在 MOOC 学习环境中, 重点关注“视频-测验-讨论”学习周期内的情绪动态变化, 通过深度时序建模技术捕捉学习者从挫折到困惑再到恍然大悟等认知情绪转变序列, 并将这些情绪变化与课程设计参数(如视频长度、内容复杂度、测验难度)进行关联分析, 识别关键情绪触发因素, 为课程设计优化提供数据支持。在实时在线课堂环境中, 重点分析师生互动质量与情绪同步性, 引入跨模态情感传导模型, 量化教师情感表达对学生情绪状态的影响程度与传播模式, 为情感教学策略提供实证依据。在协作学习环境中, 通过整合社交网络分析(SNA)与协作对话行为分析, 构建群体情绪动态模型, 识别情绪领导者与情绪传播路径, 支持情感智能干预。

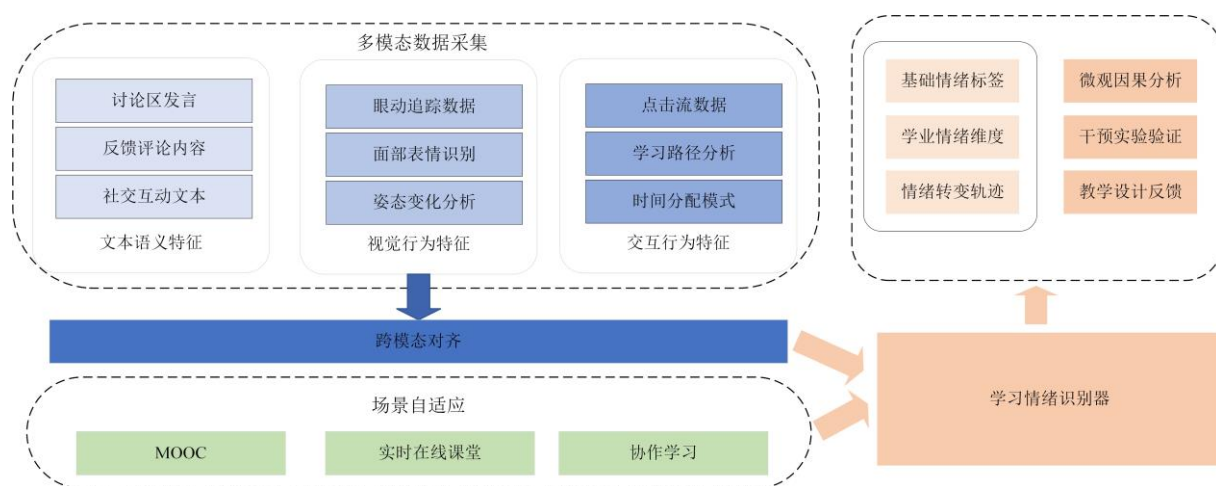


图 4 面向教育场景的多维度情绪分析框架

Fig. 4 Multi-dimensional emotion analysis framework for educational scenarios

3.2 学习情绪干预机制

基于 3.1 多维度学习情绪分析框架, 提出一种情绪驱动的自适应学习干预框架, 如图 5 所示, 该框架将干预系统设计为“感知-决策-执行”三层结构。在感知层, 系统持续监测并分析学习者情绪状态, 重点识别四类关键情绪事件, (1)持续性负面情绪(如连续 15 分钟以上的困惑或焦虑状态); (2)情绪急剧波动(如从专注到挫折的突变); (3)情绪与任务不匹配(如面对简单任务却表现高焦虑); 以及(4)群体情绪异常(如协作学习中的情绪不同步或冲突)。通过时序模式挖掘算法自动从情绪数据流中检测, 并以优先级队列方式传递至决策层。

决策层是框架的核心, 负责基于检测到的情绪事件、学习者特征及环境上下文, 生成最佳干预策略。采用基于知识的强化学习方法构建干预决策引擎, 将干预过程建模为马尔可夫决策过程。状态空间包含情绪向量、学习进度、任务难度及历史干预响应等维度; 动作空间则由四类系统化干预策略构成, (1)认知干预(Cognitive Interventions)针对困惑、焦虑等认知性负面情绪, 提供概念解析、思维导图或问题分解等支持, 降低认知负荷; (2)动机干预(Motivational Interventions), 针对厌倦、脱离等动机性问题, 实施目标提醒、成就可视化或个性化挑战等策略, 重燃学习兴趣; (3)情绪调节

干预（Emotional Regulation Interventions）提供冥想引导、认知重评提示或积极情绪诱导，帮助调节高强度负面情绪；(4)社会支持干预（Social Support Interventions）激活同伴协作、教师指导或虚拟伙伴互动，利用社会临场感缓解孤独或增强归属感。奖励函数通过情绪改善度、学习行为指标与任务完成效果的加权组合设计，实现干预效果的多维优化。

执行层，将抽象干预策略转化为具体可执行的多模态干预操作，并通过多通道协同方式呈现给学习者。(1)环境式干预（Ambient Interventions）通过学习界面背景色调变化、辅助视觉元素或环境音效等微妙方式传递情绪支持，最小程度打断学习流程；(2)集成式干预（Integrated Interventions）将干预内容自然嵌入学习材料中，如动态调整练习难度、插入自适应休息点或重构学习路径；(3)显式干预(Explicit Interventions)在情绪状态严重影响学习时,系统启动虚拟助教对话或提供直接指导，明确引导情绪调整。该框架为构建情感智能教育系统提供了实用且有效的解决方案，为下一代适应性学习环境设计提供了理论与技术基础。

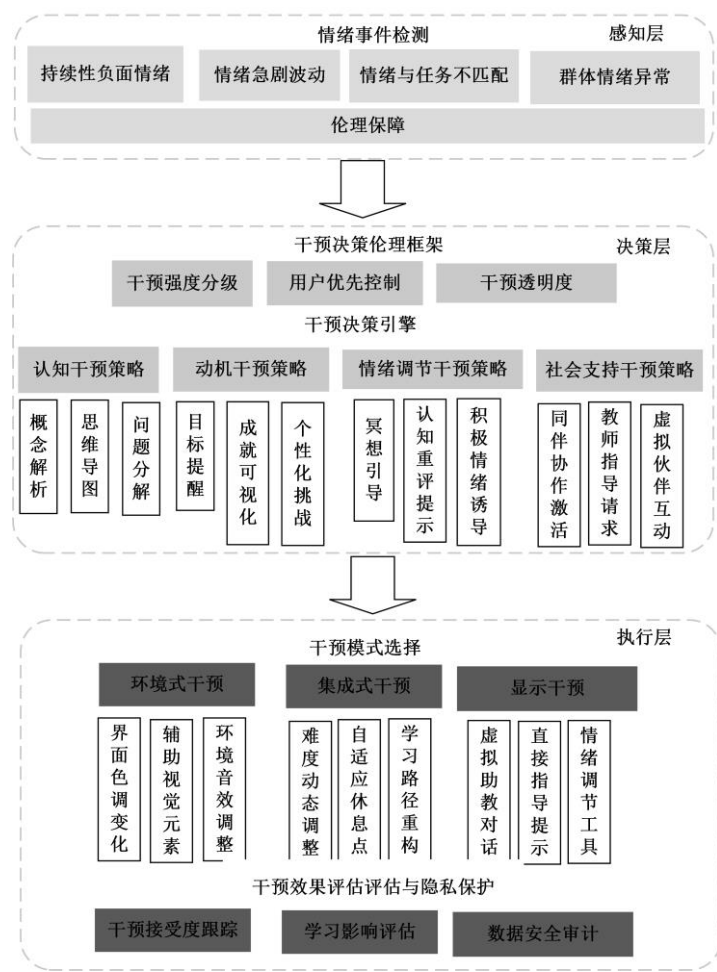


图 5 情绪驱动的自适应学习干预框架

Fig. 5 Emotion-driven adaptive learning intervention framework

4 结论

本研究构建基于深度学习的学习情绪分析框架，针对教育情境特异性的关键挑战提出系统化解

决方案。通过整合控制-价值情绪理论与多模态深度学习技术,成功建立包含文本语义、视觉行为及交互行为的三维情绪表征体系,实现对在线学习环境中复杂学习情绪的精准识别。创新点在于提出情绪驱动的自适应学习干预框架,该框架通过“感知-决策-执行”三层结构,将情绪分析与学习干预有机结合,形成闭环的情感智能教育系统,提升学习者课程完成率,缓解在线学习的高辍学率问题,为智能教育发展提供了有力支持。

未来研究将沿着三个方向深入拓展,探索大语言模型在教育情境下的情绪理解能力,通过领域适应性微调提升模型对教育特定情绪术语和表达的识别精度;加强纵向研究设计,通过长期追踪学习者情绪发展轨迹,揭示情绪-认知-行为的动态交互机制及其对学习成效的长期影响;推进情感计算与教育学的深度融合,建立更加完善的理论框架。本研究为“以学习者为中心”的智能教育生态建设提供技术路径,对推动教育信息化 2.0 发展具有重要意义。

参考文献

- [1] 左秀媛,谭翔宇. 愉悦情绪、认知评价及行为投入对学术英语学习成效的影响[J]. 外语研究,2025,42(01):30-37.
- [2] 俞建慧. 在线学习环境下基于交互视角的情绪分析及干预研究[D]. 浙江:浙江师范大学,2022.
- [3] Lei S, Dong G, Wang X, et al. InstructERC: Reforming Emotion Recognition in Conversation with Multi-task Retrieval-Augmented Large Language Models[J]. arXiv preprint arXiv:2309.11911, 2023.
- [4] Wu Z, Gong Z, Ai L, et al. Beyond silent letters: Amplifying llms in emotion recognition with vocal nuances[J]. arXiv preprint arXiv:2407.21315, 2024.
- [5] Pekrun R. Control-value theory: From achievement emotion to a general theory of human emotions[J]. Educational Psychology Review, 2024, 36(3): 83.
- [6] Lin H, Chen Q. Artificial intelligence (AI)-integrated educational applications and college students' creativity and academic emotions: students and teachers' perceptions and attitudes[J]. BMC psychology, 2024, 12(1): 487.
- [7] Linnenbrink-Garcia L, Pekrun R. Students' emotions and academic engagement: Introduction to the special issue[J]. Contemporary educational psychology, 2011, 36(1): 1-3.
- [8] Liu Y, Ma S, Chen Y. The impacts of learning motivation, emotional engagement and psychological capital on academic performance in a blended learning university course[J]. Frontiers in Psychology, 2024, 15: 1357936.
- [9] Hasnine M N, Nguyen H T, Tran T T T, et al. A real-time learning analytics dashboard for automatic detection of online learners' affective states[J]. Sensors, 2023, 23(9): 4243.
- [10] Liu S, Liu S, Liu Z, et al. Automated detection of emotional and cognitive engagement in MOOC discussions to predict learning achievement[J]. Computers & Education, 2022, 181: 104461.
- [11] Pekrun R. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice[J]. Educational psychology review, 2006, 18: 315-341.
- [12] Pekrun R, Frenzel A C, Goetz T, et al. The control-value theory of achievement emotions: An integrative approach to emotions in education[M]//Emotion in education. Academic Press, 2007: 13-36.
- [13] Loderer K, Pekrun R, Lester J C. Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments[J]. Learning and instruction, 2020, 70: 101162.
- [14] Pekrun R, Marsh H W, Elliot A J, et al. A three-dimensional taxonomy of achievement emotions[J]. Journal of Personality and Social Psychology, 2023, 124(1): 145.
- [15] Gan C, Zheng J, Zhu Q, et al. A survey of dialogic emotion analysis: Developments, approaches and perspectives[J]. Pattern Recognition, 2024: 110794.
- [16] Hu D, Wei L, Huai X. Dialoguecrn: Contextual reasoning networks for emotion recognition in conversations[J]. arXiv preprint arXiv:2106.01978, 2021.
- [17] Li W, Shao W, Ji S, et al. BiERU: Bidirectional emotional recurrent unit for conversational sentiment analysis[J]. Neurocomputing, 2022, 467: 73-82.
- [18] Wen J, Jiang D, Tu G, et al. Dynamic interactive multiview memory network for emotion recognition in conversation[J].

- Information Fusion, 2023, 91: 123-133.
- [19] Ghosal D, Majumder N, Poria S, et al. Dialoguegen: A graph convolutional neural network for emotion recognition in conversation[J]. arXiv preprint arXiv:1908.11540, 2019.
- [20] Ishiwatari T, Yasuda Y, Miyazaki T, et al. Relation-aware graph attention networks with relational position encodings for emotion recognition in conversations[C]. Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP). 2020: 7360-7370.
- [21] Shen W, Wu S, Yang Y, et al. Directed acyclic graph network for conversational emotion recognition[J]. arXiv preprint arXiv:2105.12907, 2021.
- [22] Zhang Y, Jia A, Wang B, et al. M3GAT: A multi-modal, multi-task interactive graph attention network for conversational sentiment analysis and emotion recognition[J]. ACM Transactions on Information Systems, 2023, 42(1): 1-32.
- [23] Yang Z, Li X, Cheng Y, et al. Emotion recognition in conversation based on a dynamic complementary graph convolutional network[J]. IEEE Transactions on Affective Computing, 2024, 15(3): 1567-1579.
- [24] Zhong P, Wang D, Miao C. Knowledge-enriched transformer for emotion detection in textual conversations[J]. arXiv preprint arXiv:1909.10681, 2019.
- [25] Ghosal D, Majumder N, Gelbukh A, et al. Cosmic: Commonsense knowledge for emotion identification in conversations[J]. arXiv preprint arXiv:2010.02795, 2020.
- [26] Li Z, Tang F, Zhao M, et al. EmoCaps: Emotion capsule based model for conversational emotion recognition[J]. arXiv preprint arXiv:2203.13504, 2022.
- [27] Ma H, Wang J, Lin H, et al. A transformer-based model with self-distillation for multimodal emotion recognition in conversations[J]. IEEE Transactions on Multimedia, 2023.
- [28] Shen X, Huang X, Zou S, et al. Multimodal knowledge-enhanced interactive network with mixed contrastive learning for emotion recognition in conversation[J]. Neurocomputing, 2024, 582: 127550.

基金项目: 江苏省教育科学规划重点课题《人工智能赋能学习情绪分析及干预研究》(B-b/2024/01/166), 江苏省高等学校基础科学(自然科学)研究面上项目《基于肠道超声图像分割的炎症性肠病辅助诊断方法研究》(24KJB520005), 江苏开放大学(江苏城市职业学院)“十四五”2023年度科研规划课题学科研究青年专项(2023XK003)

¹ **第1作者简介:** 朱颖雯(1982-), 女, 博士, 江苏开放大学, 研究方向: 数据挖掘、机器学习、人工智能。 E-mail: zhu_ying_wen@163.com

*** 通讯作者简介:** 张迁(1994-), 男, 博士, 讲师, 江苏开放大学, 研究方向: 信息化教育教学, 计算机视觉, 深度神经网络。 E-mail: zhangqian@jsou.edu.cn