

学习者多模态情绪融合分析：动因、框架与路向

唐焯伟^{1,2} 卜凡丽¹ 赵一婷¹

(1. 东北师范大学 信息科学与技术学院, 吉林 长春 130117;
2. 教育部数字化学习支撑技术工程研究中心, 吉林 长春 130117)

[摘要] 如何精准识别学习者情绪, 深度融合各模态情绪数据, 剖析情绪发生的作用机制, 已成为智能教育发展的关键问题。本研究提出学习者多模态情绪融合分析概念, 从学习者多模态情绪融合分析的动因、框架与路向三个维度, 探索多模态情绪数据之间的有机融合, 精准感知学习者的情绪状态, 挖掘多模态情绪数据的潜在价值, 实现对学习发展规律的细粒度解析, 推进学习者情绪识别领域的实践探索。

[关键词] 情绪识别; 情绪融合分析; 多模态情绪融合分析

[中图分类号] G434 **[文献标识码]** A **[文章编号]** 1007-2179(2023)03-0096-08

学习者多模态情绪识别是近年来智能教育领域的重点研究问题, 关注学习者情绪状态感知、情绪发生机理与学习现象规律等(王一岩等, 2022)。研究表明, 积极情绪能激发学习者的学习兴趣, 消极情绪容易使学习者产生学习倦怠, 降低学习投入度(孙发勤等, 2020)。从情绪角度出发, 感知学习者表情、声音、神态、动作的情绪状态, 有利于针对性地调节学习者的心理状态, 提高学习效率(陈子健等, 2019)。

基于此, 本研究提出学习者多模态情绪融合分析(learner multimodal emotion fusion analytics, LMEFA)概念, 通过采集学习者不同模态数据, 融合各模态间的情绪数据, 建立模态数据之间的强关联性, 精准感知学习者的情绪状态, 探究学习者情绪与内在认知发展的关联关系, 为学习者高效学习提供数据支撑。

一、动因

(一) 研究现状

研究表明, 不同模态数据之间存在耦合关系。有效整合不同模态间的数据, 有利于发挥各个模态的作用, 从而精确地判断学习者的情绪状态。学习者多模态情绪融合分析研究主要集中于多模态情绪数据采集和多模态特征融合两方面。

1) 多模态情绪数据采集, 主要是对多种行为表现模态混合、多种神经生理模态混合和神经生理与行为表现模态混合三种模式的情绪数据进行采集(潘家辉等, 2020)。多种行为表现的数据采集主要是采集学习者的面部、语音、语调、手势、身体姿态等语言行为和非语言行为数据, 实现情绪识别; 多神经生理的信息数据采集主要通过智能传感设备, 采集学习者的眼动、心跳等生理数据, 进而实

[收稿日期] 2023-02-11 [修回日期] 2023-04-19 [DOI 编码] 10.13966/j.cnki.kfjy.2023.03.010

[作者简介] 唐焯伟, 副教授, 博士生导师, 东北师范大学信息科学与技术学院, 研究方向: 智慧教育、智慧学习环境等(tangyw100@nenu.edu.cn); 卜凡丽, 硕士研究生, 东北师范大学信息科学与技术学院教育技术系, 研究方向: 智慧教育; 赵一婷, 博士研究生, 东北师范大学信息科学与技术学院教育技术系, 研究方向: 智慧教育。

[引用信息] 唐焯伟, 卜凡丽, 赵一婷(2023). 学习者多模态情绪融合分析: 动因、框架与路向[J]. 开放教育研究, 29(3): 96-103.

现情绪识别;基于神经生理模态与行为表现模态的混合模式主要是综合行为表现和神经生理优点,精准识别学习者的情绪。不同模态情绪识别的效果见表1。相对于行为表现模态,神经生理模态的数据采集难度更高,识别率更精确。

表1 不同模态类型的情绪识别效果对比

模态属性	模态类型	采集难度	信噪比	识别率
神经生理	脑中枢变化	高	低	高
	外周生理变化	高	高	中
行为表现	脸部表情	低	高	中
	语音语调	低	中	中
	肢体动作	低	高	低

2)多模态特征融合,主要包括数据级融合、特征级融合、决策级融合与模型级融合。数据级融合直接将收集到的多模态数据进行组合,可以保持模态数据的完整性,但操作过程复杂;特征级融合通过提取不同模态的数据特征,形成多模态数据特征级,但未考虑模态之间的差异性;决策级融合先探索不同模态对情绪识别的重要程度,再进行融合,联合决策,操作相对容易,但探究各模态的重要程度非常重要;模型级融合需根据模态特性建立合适的模型,联合学习关联信息,优势在于自主定位融合路向。

本研究梳理的多模态情绪融合分析案例见表2。当前多模态情绪融合分析存在的棘手问题在于:一

是研究年限较短,识别准确度有待提升;二是多模态情绪融合分析技术复杂;三是多模态情绪数据和多融合算法的优化组合问题亟需解决。

(二)问题与解决思路

学习者情绪融合分析在智能教育领域存在的问题,主要体现在技术、理论和应用三个层面(见图1)。

1. 理论层面:情绪发生机理

教育学、心理学、脑科学、认知科学等学者从本领域的研究出发,识别学习者的情绪状态,探究个体情绪的发展规律,但没有统一的共识性情绪识别理论,支持情绪发生机制和演化机理的深层次挖掘分析,难以准确诠释情绪发生的内在机理,一定程度上阻碍了学习者情绪融合分析研究的发展。

这一问题的解决思路是结合脑科学、认知科学、心理学和人工智能领域开展跨界融合协同研究,丰富情绪融合分析的理论基础,并从行为表现和神经生理等角度多元化、精准识别学习者情绪,实现多维建模,深度探究学习者的情绪发展规律,诠释情绪的内在发生机理。

2. 技术层面:技术壁垒

多模态情绪数据库是情绪融合分析的基础。国外学者已基于不同需求建立情绪识别的开源数据库(李锦瑶等,2023),但样本量较小,且应用场景属非教育领域,一定程度上制约了情绪融合分析的教育应用发展。

表2 不同混合模式的多模态情绪融合分析案例

模态混合模式	作者和文献	模态与特征提取	多模态特征融合策略	分类算法
多种行为表现模态混合	Noroozi et al.(2017)	语音: Mel 频率倒谱系数、滤波器组能量和韵律特征 面部表情: 面部特征点的几何关系	模型级融合	支持向量机
	Castellano et al.(2008)	面部表情: 几何特征 手势: 手形、手掌方向和运动方向等 语音: 基于谱的特征	特征级融合	朴素贝叶斯
多种神经生理模态混合	Koelstra et al.(2012)	脑电(EEG)信号 δ 、 θ 、 α 、 β 节律功率谱特征, 时域特征	决策级融合	朴素贝叶斯
	Chanel et al.(2009)	脑电(EEG): 频域微分熵 血压、皮肤电、心率、呼吸	模型级融合	线性判别分析朴素贝叶斯
行为表现与神经生理模态混合	Soleymani et al.(2011)	脑电: 频域功率谱密度 眼动: 瞳孔功率谱眨眼、注视等时域特征	决策级融合	支持向量机
	Huang et al.(2019)	面部表情: 面部特征点的几何关系 脑电: 频域特征	决策级融合	SVM(脑电) CNN(表情)

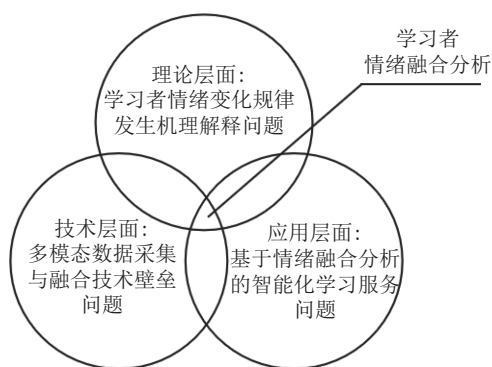


图1 学习者情绪融合分析现存问题

此外,如何精准采集学习者各模态情绪数据,理解并利用多种模态数据之间的关联性也亟待解决。

这一问题的解决思路是:开发融合算法,利用多模态融合技术采集各模态情绪特征数据,优化不同模态数据间的组合,尽可能不遗漏数据,最大程度地融合各模态信息,实现深层次的数据融合;建立情绪融合分析数据库,通过多学科联动,精细化学习情绪数据库,揭示学习情绪的内在发生机理。

3. 应用层面:基于情绪融合分析的智能化学习服务问题

学习者情绪融合分析的目的是精准识别学习者情绪状态,揭示情绪发生的内在机理,为学习者提供精准的个性化智能服务。已有研究证实学习者情绪变化在很大程度上反映了学习者的学习动机、学习投入度等个性化特征(王一岩等,2021a),但情绪融合分析的教育研究主要聚焦学习者情绪与学习情境、学习投入度和知识建构等单一特征的关联关系,不足以解释学习者情绪与认知过程间的联系,无法支持智能化学习服务。

这一问题的解决思路是:剖析学习者的认知发展规律,关联分析学习者的内部认知加工过程与情绪状态,探寻学习者情绪状态的内在发生机理,为学习者提供个性化支持服务,优化学习效果,实现科学化精准学习支持。

(三)学习者多模态情绪融合源息性分析

情感计算由美国学者皮卡德(Picard)提出,指对人的外在表现和内在生理信息进行测量、分析和计算,强调情感的发生、识别和表达。情绪识别作为情感计算的关键环节,通过测量个体的行为和生理特征指标预测其情绪状态,建立人体内外在表

现特征与情绪的联系。人体常通过多通道显露情绪,因此融合多通道的多模态情绪识别已成为智能教育的重要研究方向。

本研究提出的学习者多模态情绪融合分析主要包括多模态情绪数据采集和同步多模态融合表征两部分。学习者多模态情绪数据是多模态情绪融合分析的基础,源于学习者的外在行为特征和内在生理指标。多模态融合表征指从各模态数据提取特定特征,探寻多模态数据间的相关性。多模态情绪融合分析的核心是优化组合多模态数据和多融合算法,分析学习者多模态情绪数据,解释学习者行为、情绪和内在认知发展的关联机理,实现科学、精准教学,推动智慧教育创新发展。

二、教育应用框架

(一)学习者多模态情绪融合分析研究框架

多模态学习者情绪融合分析包括数据采集、数据特征提取与融合、情绪识别与分类、动态干预与调节四个步骤。基于此,本研究构建了学习者多模态情绪融合分析研究框架(见图2)。该框架以教育场景为导向,利用智能传感设备和智能头戴式设备,实时动态采集学习者多模态情绪信息,多维度、多方法融合分析情绪特征数据,形成学习者情绪特征数据集,并根据数据集精准识别学习者情绪状态,反馈、调节与干预学习者的情绪,激发积极的学习情绪,提升学习质量。

1. 以教育场景为导向,采集学习者多模态情绪数据

捕获学习者学习过程的情绪状态信息,有助于教师调整教学方式和教学策略,促进学习者智慧学习。目前,研究者采集情绪数据已从人工观察、自我报告等主观评判方法,逐步过渡到通过采集语音、面部表情、脑电和皮肤电等数据,识别学习者情绪的客观评判方法。

在不同教育场景中,学习者多模态情绪数据采集方式不同。本研究从学习者的语音、面部表情、身体姿态和生理信息等通道采集多模态情绪数据。语音数据指通过声学特征、语音语调、语言文字和上下文信息等特征分析学习者的情绪状态;面部表情指通过肌肉的动态变化反映学习者的情绪状态;身体姿态指通过身体语言表达情绪,如通过“低头

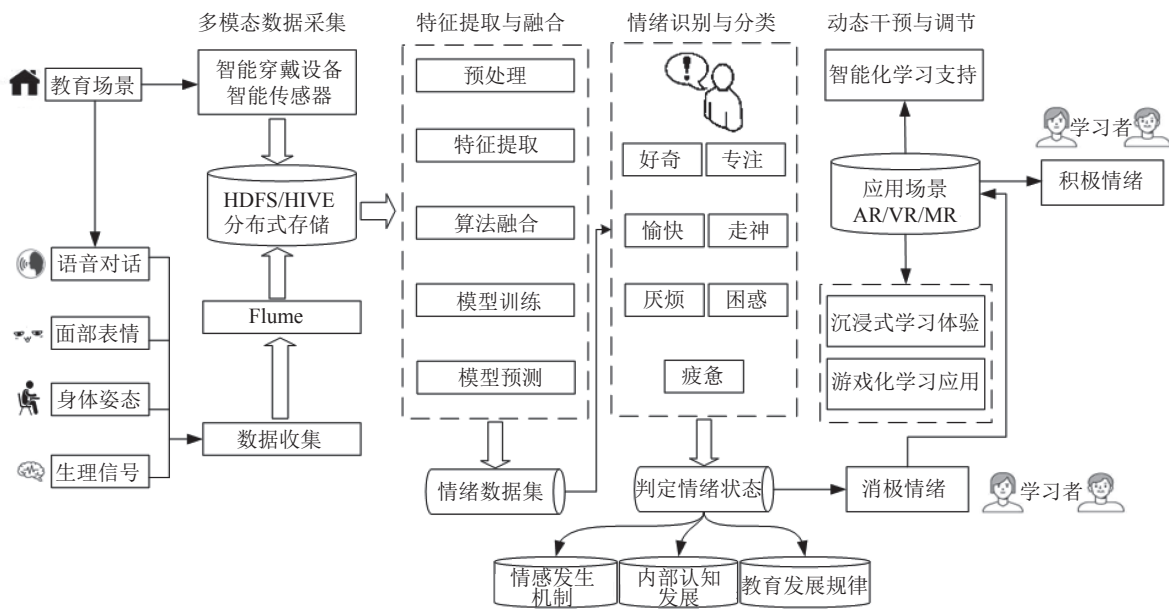


图2 学习者多模态情绪融合分析研究框架

奋脑”“蹲地”“含胸弯腰”等肢体运动表达悲伤情绪;生理信息指通过心跳、眼动、脑电等数据,分析学习者情绪。

2. 以多模态数据融合为依托,实现情绪数据信息互补

学习者情绪源于与环境的相互作用,是学习者直接产生的心理体验,反映学习者多个维度的变化。本研究在采集学习者多种模态情绪数据的基础上,采用 CNN 和 LSTM 等算法提取多模态数据特征,并将多模态数据特征进行融合分析(王一岩等, 2021b)。

并非学习者所有的多模态情绪数据都对情绪识别有积极意义。本研究在特征选择与提取阶段,先对数据进行预处理,即采用贝叶斯方法、随机森林和主成分分析法等降噪处理数据信息;其次组织数据,将单一模态的数据有序排列;在数据融合阶段,采用模型级融合策略,构建深度网络模型,建立多层结构,利用深度学习等技术整合多模态特征数据,尽量保留数据的原始信息,然后将模态数据融合互补,并根据不同模态数据共性特征灵活选取适当位置进行融合,形成学习者多模态情绪数据集,为学习者情绪识别提供数据基础。

3. 以多模态情绪数据集为基础,识别学习者情绪状态

人本主义学习理论认为,学习是情感与认知统

一的精神活动,强调重视学习者的意愿、情感与需要,提升学习成效,激发学习动机,促进知识深层次的意义建构。

本研究采用北京师范大学情绪数据库,将情绪定义为好奇、专注、愉快、走神、厌烦、困惑和疲惫七类,结合收集的语音、面部表情、身体姿态和生理信号的多模态情绪数据集,基于混合神经网络模型精准识别学习者情绪状态,还原教学过程全貌,剖析学习者情绪发生机制,深入研究教育情境和内部认知发展状态影响学习者情绪状态的机理,探究学习者的深层认知发展规律,并动态干预与调节消极情绪,为学习者提供智能化学习支持服务。

4. 以应用场景为调节场域,动态干预学习情绪受学习环境、学习内容和学习体验等因素的影响,学习者会产生焦虑、沮丧、厌烦等消极情绪,导致学习效果不佳。不良情绪若不能得到适当干预和调整,将严重影响学习效果。

应用场景的创建与发展能为减少学习者的消极情绪提供解决方案。沉浸式学习体验、游戏化学习和智能化服务可为学习者的情绪调节提供全面支持。在应用场景的支持下,学习者可以进入教师预设的具有情绪色彩的教学情境中,开展协作式、具身化学习,沉浸式体验实地实景的课堂互动,增强学习参与度,改善学习者的情绪状态;借助 VR/AR/MR 等技术,学习者可以选择或创造游戏场景、

游戏角色与游戏任务等, 调动自身的感官, 通过语音、动作、手势等完成学习任务, 开展游戏化学习, 增强学习愉悦感; 学习者通过表情变化、肢体互动和话语等与应用场景的虚拟助理交流反馈, 虚拟助理结合学习内容、学习者个性特征和学习路径推送智能化资源, 激发学习者的积极情绪。

(二) 学习者情绪融合分析技术

本研究设计了以文本为中心的多模态特征融合方法, 分别构建图片模态、音频模态和生理信息模态之间的共享特征, 采用跨模态注意力机制, 实现多模态之间特征共享, 同时各模态保留自有特征, 达到多模态特征融合的目的(见图3)。

1. 数据采集与清理

首先, 本研究通过智能传感器设备采集学习者语音和面部表情数据, 通过学习者佩戴的智能轻型可穿戴设备采集学习者身体姿态和生理信息(眼动、心跳、脑电)数据, 作为学习者情绪融合分析的数据集; 其次, 对集中的数据进行批量预处理, 包括去噪、端点检测、视频画面和音频语音段切割, 并通过语音识别 API 将切割的语音段批量转换成文本译文; 最后, 采用人工方式对数据集进行文本纠错和情绪标注。

2. 模态信息提取

本研究采用智能传感器设备和智能轻型可穿戴设备, 提取学习者语音文本、面部表情、身体姿势和生理信息四类多模态信息, 为融合多模态特征识别学习者情绪做准备。

1) 针对学习者面部表情信息, 本研究采用

Faster RCNN 网络处理课堂情绪识别数据集, 结合 Dlib 工具库获取眉毛、眼睛、嘴巴和鼻子等关键部位特征点坐标, 识别和定位学习者的面部表情, 实时截取学习者课堂中关键表情图片帧, 主要包括笑脸、困惑、惊讶和无表情四类。

2) 针对学习者身体姿态信息, 本研究采用 YOLOv5 模型动态检测学习者上身姿态, 主要包括抬头、低头、趴下和左顾右盼四种姿态, 完成学习场景下学习者的表情、身体姿态图片数据的预准备。

3) 针对学习者生理信息, 本研究采用 Inside-Out+6DoF 技术的可穿戴设备, 采集学习者的心跳、眼动和脑电等生理信息数据。

4) 针对学习者语音文本信息, 本研究利用课堂交互对话等途径获取文本和声学信息, 利用降噪工具预处理声学信息, 去除音频中的噪声和空白, 利用 Jieba 工具对文本进行分词、清理和标准化, 完成语音数据和文本数据的预准备工作。

3. 多模态特征提取与融合

对上述语音、面部表情、身体姿态和生理信息四种模态的信息分析和现有研究(叶俊民等, 2020)表明, 文本模态研究占主导地位。相对其他模态, 文本模态的语义表达更准确和简洁。因此, 本研究设计采用文本为中心的多模态特征融合方法, 构建其与图片模态、音频模态和生理信息模态之间的共享特征。

三、教育应用场景的实现路向

随着人工智能技术的发展, 越来越多的研究者

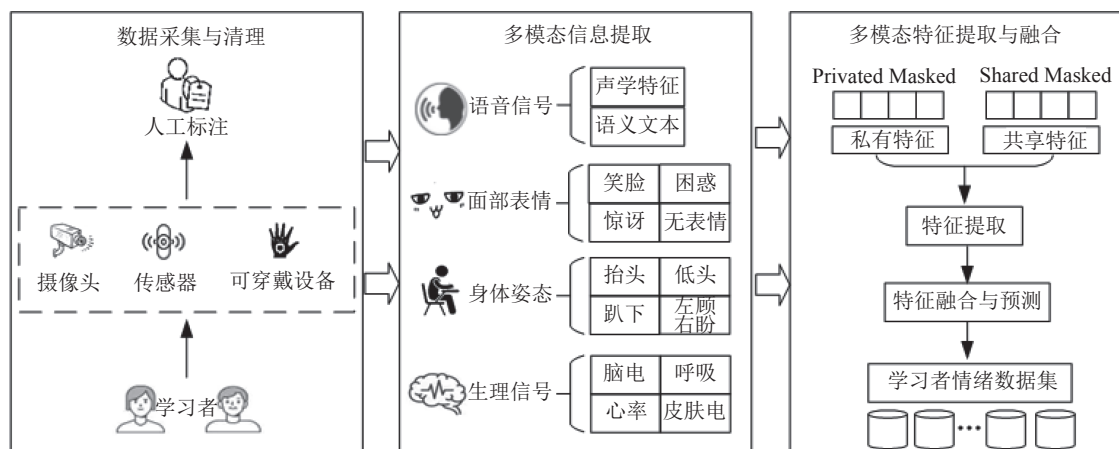


图3 学习者多模态情绪融合分析过程

关注学习者情绪融合分析,挖掘学习者的情感状态,调节学习者情绪,提升学习质量(Hoaiy-Du et al, 2023)。张乐乐等(2022)基于语音、面部表情、身体姿态和脑电等方法识别学习者情绪状态,但对学习者多模态情绪数据的融合互补方面的关注有待加强,对学习者的情绪调节缺乏交互体验,且无法感知调节后学习者情绪是否提升。应用场景的出现让学习者情绪状态的精准识别与动态调节成为可能:首先,基于AR/VR的沉浸式感官体验,有助于剖析学习者情绪状态,探索学习者情绪状态的内在发生机理。其次,基于沉浸式学习环境,有助于通过交互式学习方式调节学习者情绪,激发学习者积极情绪,促进学习者智慧学习。

(一)多模态情绪融合分析支持个性化学习

现有个性化学习研究多从学习者的需求角度出发,教师提供相适应的调节策略,以实现个性化学习(徐碧波等,2023)。这种个性化学习方式仅考虑影响学习的直接因素,忽视情绪这一间接因素对学习的影响,也无法精准支持学习者个性化学习。在教学过程中,情绪能够影响学习者的注意力和学习动机,进而影响学习质量(叶芮杏等,2022)。因此,精准识别学习者情绪有助于实现个性化教学,但教育场景的限制导致学习情绪识别研究是间断非连续的,无法准确还原学习者情绪变化过程(米珍美等,2021)。

应用场景支持的学习者多模态情绪识别融合分析,能够通过智能传感设备和智能轻型可穿戴设备实时动态捕捉学习者多模态情绪数据,准确提取多模态数据特征,并最大限度地保留学习者多模态情绪数据的自有特征,精准识别学习者情绪状态,刻画学习者情绪演变过程,挖掘学习者情绪变化的原因,细粒度地还原教学过程全貌,更精确地支持学习者个性化学习。同时,应用场景的虚实融合特性支持学习者个性化学习的精准、永续发展,结合AR、MR等技术的应用场景,可打破虚拟世界与现实空间的“次元化”壁垒,实现虚拟世界与现实世界的精准叠加和虚实交互,融合内外部能量感知学习者情境,实现学习者资源的动态生成和可持续发展,呈现从预设到生成、从内容到活动、从通用到专用、从千人一面到千人万面的个性化资源推荐。

应用场景能为学习者智能供给资源,实现资源

与认知的高度耦合:1)精准化聚合数据。将学习者的实时动态行为、情绪、心理等数据反馈回系统,实现动态调整;2)动态升级资源,不断更新原始资源、补给新生资源,形成智能实时更新资源库,实现知识的动态循环更替;3)个性化精准服务,贯通知识流转、场景再现、路径更新与资源供给的通道,构建完整的知识体系框架;4)促进虚实融合,达成现实世界与虚拟世界的紧密衔接,实现真正意义的个性化学习(刘革平,2021)。

(二)多模态情绪融合分析促进深度认知加工

认知情感理论认为,情绪影响学习者的认知加工和动机(周进等,2021),主张通过调节认知也可以控制情绪,积极情绪促进认知发展,而消极情绪阻碍正常认知,情绪会通过知觉、注意和记忆等认知活动影响学习。在知觉方面,学习者情绪状态会影响其对事物的感知。在注意方面,相比于中性和消极情绪,积极情绪会减少任务转换的损失,扩大注意范围(Rowe et al., 2007),而注意范围在阅读中具有重要作用。在记忆方面,带有情绪信息的事件容易被精细复述,加强记忆。因此,学习者情绪的精准识别和干预能促进其深度加工学习内容。

高沉浸感是应用场景的关键特征之一。在应用场景中运用VR头盔和基于传感器的可穿戴设备能捕获学习者生理数据,通过获取心电、脑电、眼动等生理信号的时域特征和时频特征,全方位洞察学习者情绪变化,解释学习者认知变化(华子荀等,2022)。同时,基于应用场景的“强在场”性,利用物理空间与数字空间的无缝衔接,可构建虚实融合的“虚拟孪生”场景,为学习者提供视觉、听觉、嗅觉、味觉、情绪、心理、生理等全方位高逼真的多感官深度沉浸体验,实现生命沉浸。在这样多样化的沉浸式互动的虚拟情境场域中,学习者可以获得现实和虚拟学习需求的满足,激发积极情绪,实时动态干预调节自身情绪。在此基础上,深度剖析学习者情绪状态,可以更层次地揭示学习者情绪状态与内在认知过程的内在发生机理,促进学习者的深度学习与认知发展。

(三)学习者多模态情绪融合分析驱动教育决策科学化

在教学过程中,教师需要结合学习者的学习进度、学习状态和学习结果等作出相应的教育决策,

调整教学方案,改进教学策略,提高教学效率与学习质量。因此,把握学习者学习状态是制定教育决策的前提。准确预测学习者情绪对教师优化教育决策、实施有效教育、缩小师生共情鸿沟、促进师生关系和谐具有重要意义(陈宁等,2013)。虽然现阶段学习者情绪识别准确率已达百分之八十以上,但还有待提高,信息特征提取方面也仍待完善,多模态情绪数据融合方面也有待加强。

应用场景的开发为提高学习者情绪识别准确性提供了新的视角,提高了数据捕获的精准性与模态采集的多样性。它能够结合学习者的行为、学习动机、学习投入等,实现对学习者情绪的立体化阐释,完善教育决策和学习各个维度新的定性和定量标准,提供更多元化、完整性、准确性的数据证据,驱动科学教育决策的生成。

同时,应用场景聚合了新兴的智能信息技术。这些技术利用深度学习等算法挖掘情绪的特征规律,凝聚物联网、人工智能、大数据等技术的核心效用,数据化、科学化、智能化地预测学习者的未来趋势,真正建立起基于证据的教育评价,优化学习行为与决策管理,促进科学化测评、系统化反馈、多元化评价的发生,推动教育决策的科学化。

四、反思与建议

应用场景的出现为学习者情绪融合分析研究提供了新的视野,有助于精准挖掘学习者情绪产生的内在机理,为相关问题的解释提供科学依据。学习者多模态情绪融合分析的应用价值主要体现在多模态数据和多融合算法的优化组合,一定程度上提高了情绪识别与融合的准确度,有助于发掘情绪背后隐藏的意义。未来应用场景的学习者多模态情绪融合分析研究可聚焦以下三方面:

1.降低数据采集的侵入性,平衡数据真实性与准确性

采集学习者的情绪数据需要适当的工具,如可穿戴设备等。然而,相关工具的使用易导致学习者的抗拒心理,导致注意力下降、思维闭塞等,影响数据的真实性与准确性。未来的应用场景应将多模态情绪数据采集与意念控制相结合,实现无线连接,伴随式采集多通道信息,尽可能在学习者无意识情况下采集真实数据,保证数据的准确性。

2.综合衡量学习状态,增强个性化学习精准性

目前,多数研究从行为、认知或情绪的单一角度指导个性化学习,缺少不同视角结合的综合关联性分析,无法精准判断学习者的学习状态。未来研究要利用应用场景的情绪融合分析优势,精准感知学习者情绪状态;要将学习者情绪与个体特征、认知能力、学习态度等相结合,利用应用场景的集成功能,融合学习者的情绪数据、认知数据与行为数据等多维数据,提高个性化学习的精准性。

3.融合教育情境信息,推动学习者情绪融合分析的应用探索

在应用场景的推动下,学习者情绪融合分析逐步应用于多元场景中,通过物理世界和虚拟世界的情绪数据结合,全面感知与持续跟踪学习者的情绪状态,并建立情绪与认知之间的连接,深层次解析认知发展规律和情绪发生机制。当前学习者情绪融合分析的应用探索仍存在局限,未来应着力探索基于学习者情绪融合分析的教育生态系统(如共享数据标准、共享学习模型)、基于教育情境感知的动态化学习者模型和面向学习者情绪调节的智能化学习服务等,为教育改革提供参考。

[参考文献]

- [1] Castellano, G., Kessous, L., & Caridakis, G. (2008). Emotion recognition through multiple modalities: Face, body gesture, speech[M]. Berlin: Springer-verlag. 4868: 92-103.
- [2] Chanel, G., Kierkels, J. M., & Soleymani, M. (2009). Short-term emotion assessment in a recall paradigm[J]. International Journal of Human-computer Studies, 67(8): 607-627.
- [3] 陈宁, 卢家楣, 汪海彬(2013). 人际共情鸿沟可以跨越:以教师预测学生情绪为例[J]. 心理学报, 45(12): 1368-1380.
- [4] 陈子健, 朱晓亮(2019). 基于面部表情的学习者情绪自动识别研究——适切性、现状、现存问题和提升路径[J]. 远程教育杂志, 37(4): 64-72.
- [5] Huang, Y. G., Yang, J. H., & Liu, S. Y. (2019). Combining facial expressions and electroencephalography to enhance emotion recognition[J]. Future Internet, 11(5): 105.
- [6] Hoaiy-Du, L., Guee-Sang, Lee., Soo-Hyung, Kim., Seungwo, Kim., & Hyung-Jeong, Yang. (2023). Multi-label multimodal emotion recognition with transformer-based fusion and emotion-level representation learning[J]. in IEEE Access, 11: 14742-14751.
- [7] 华子荀, 付道明(2022). 学习元宇宙之内涵、机理、架构与应用研究——兼及虚拟化身学习促进效果[J]. 远程教育杂志, 40(1): 26-36.
- [8] Koelstra, S., Muhl, C., & Soleymani, M. (2012). DEAP: A data-

base for emotion analysis; using physiological signals[J]. IEEE Transactions on Affective Computing, 3(1): 18-31.

[9] 李锦瑶, 杜肖兵, 朱志亮, 邓小明, 马翠霞, 王宏安(2023). 脑电情绪识别的深度学习研究综述[J]. 软件学报, 34(1): 255-276.

[10] 刘革平, 王星, 高楠, 胡翰林(2021). 从虚拟现实到元宇宙: 在线教育的新方向[J]. 现代远程教育研究, 33(6): 12-22.

[11] 米珍美, 赵恒斌, 高攀(2021). 基于 ConvLSTM 网络的维度情感识别模型研究[J]. 计算机工程与应用, 57(18): 289-296.

[12] Noroozi, F., Marjanovic, M., Njegus, A., Escalera, S., & Anbarjafari, G.(2017). Audio-visual emotion recognition in video clips[J]. in IEEE Transactions on Affective Computing, 10(2): 60-75.

[13] 潘家辉, 何志鹏, 李自娜, 梁艳, 邱丽娜(2020). 多模态情绪识别研究综述[J]. 智能系统学报, 15(4): 633-645.

[14] Rowe, G., Hirsh, J. B., & Anderson, A. K.(2007). Positive affect increases the breadth of attentional selection[J]. Proceedings of the National Academy of Sciences, 104(1): 383-388.

[15] Soleymani, M., Pantic, M., & Pun, T.(2011). Multimodal emotion recognition in response to videos[J]. IEEE Transactions on Affective Computing, 3(2): 211-223.

[16] 孙发勤, 冯锐(2020). 编程教育中的智能导师系统: 架构、设计与应用[J]. 远程教育杂志, 38(1): 61-68.

[17] 王一岩, 郑永和(2022). 多模态数据融合: 破解智能教育关键问题的核心驱动力[J]. 现代远程教育研究, 34(2): 93-102.

[18] 王一岩, 刘士玉, 郑永和(2021a). 智能时代的学习者情绪感知: 内涵、现状与趋势[J]. 远程教育杂志, 39(2): 34-43.

[19] 王一岩, 王杨春晓, 郑永和(2021b). 多模态学习分析: “多模态”驱动的智能教育研究新趋向[J]. 中国电化教育, 41(3): 88-96.

[20] 徐碧波, 裴沁雪, 陈卓, 祝正洲(2023). 国家中小学智慧教育平台推进基础教育数字化转型的现实意义与优化方向[J]. 中国电化教育, 433(2): 74-80.

[21] 叶俊民, 周进, 李超(2020). 情感计算教育应用的多维透视[J]. 开放教育研究, 26(6): 77-88.

[22] 叶芮杏, 孙福海, 黄甫全(2022). AI 全科教师开发: 基础、模型与关键技术[J]. 电化教育研究, 43(3): 42-50.

[23] 张乐乐, 顾小清(2022). 多模态数据支持的课堂教学行为分析模型与实践框架[J]. 开放教育研究, 28(6): 101-110.

[24] 周进, 叶俊民, 李超(2021). 多模态学习情感计算: 动因、框架与建议[J]. 电化教育研究, 42(7): 26-32+46.

(编辑: 李学书)

Learners Multimodal Emotion Fusion Analysis (LMEFA): Motivation, Framework and Direction

TANG Yewei^{1,2}, BU Fanli¹ & ZHAO Yiting¹

(1. School of Information Science and Technology, Northeast Normal University, Changchun 130117, China; 2. Engineering Research Center of Digital Learning Support Technology, Ministry of Education, Changchun 130117, China)

Abstract: In today's diversified education field, how to accurately identify learners' emotions, deeply integrate the emotional data of various modes, and analyze the action mechanism of emotional occurrence has evolved into a key issue in the development of intelligent education. To address it, this study puts forward the concept of Learner Multimodal Emotion Fusion Analytics (LMEFA). Through the three dimensions of the learners' motivation, its framework, and the path, the LMEFA explains the organic fusion among multimodal emotional data, accurate perception learners emotional state, mining the potential value of multimodal emotional data, and establishing the law of development. LMEFA is significant and meaningful in providing a reference guide for both research and practice in the field of emotional recognition.

Key words: Emotion recognition; emotion fusion analysis; learners' multimodal emotion fusion analysis