

自然语言处理技术在老年远程教育中的应用研究

祁 萌¹

QI Meng

摘 要

针对老年群体参与远程教育时存在的生理、认知及情感障碍，文章结合知识图谱与情感计算技术，设计了一套自然语言处理系统，旨在提高老年人的学习效率与使用体验。系统采用双引擎架构，可同时适配认知学习与情感交互两类需求；通过集成多任务语义识别、动态门控机制、实体链接与共情生成模块，同步实现知识传递与心理陪伴功能。研究采用准实验方法，搭建包含技术性能、用户体验、学习成效的三维评价体系，对系统效果开展验证。实验表明，对话式教学助手能够有效提升老年学习者的课堂参与度与知识掌握水平。据此梳理出适用于老年教育的自然语言处理技术方案，可为同类智能教育系统的研发提供思路与实证依据。

关键词

自然语言处理；老年远程教育；人机交互；需求驱动设计

doi: 10.3969/j.issn.1672-9528.2026.06.009

0 引言

随着人口老龄化程度不断加深，远程教育成为提升老年人数字素养与生活质量的重要手段，然而传统系统普遍忽视老年学习者在生理感知、认知处理与情感交互等方面的特殊需求从而导致学习成效不佳。本研究基于自然语言处理技术聚焦老年群体的远程学习场景，结合知识图谱问答与情感计算模型构建具备认知与情感双引擎的教学伴侣系统，立足于需求驱动与适老交互理念尝试突破现有技术对老年用户支持不足的瓶颈，力求实现个性化、低负荷和强陪伴的学习支持方式，推动老年教育系统向人性化与智能化转型。

1 老年学习者的特征与核心需求分析

1.1 生理与认知特征

老年学习者远程教育的有效性，受制于其生理机能的系统性变化。视觉衰退是首要交互障碍：晶状体调节能力下降致使聚焦困难，难以辨识小号字体与图标；视网膜对光与色彩敏感度降低，使低对比度设计构成信息壁垒。听觉上，高频听力损失使未经处理的人声与高频音模糊不清，在噪声环境中分离语音的能力下降，若无清晰音源与字幕，信息传递效率严重受损。认知与运动控制能力的变化直接影响操作流畅性^[1]。信息处理减慢与工作记忆缩减，要求交互流程极度简化，避免多步骤任务增加认知负荷。手部精细动作控制力

减弱，也使精确点击小目标成为挑战^[2]。这些生理特征构成了数字界面的“高墙”，远程教育系统的设计必须正视并跨越这些障碍，才能实现有效的知识传递。

1.2 心理情感特征

相较于生理局限，老年学习者的心理与社会特征是决定其学习动机与持续性的深层因素^[3]。其学习动机已脱离职业发展的功利性需求，转向内在价值的实现。学习新知识、掌握新技能，更多是为了维持思维活跃、丰富精神生活、建立新的社会连接，或是在家庭中实现代际沟通的“数字反哺”。这种基于兴趣与自我实现的动机，决定学习内容必须与生活经验紧密关联。自尊心是影响老年学习者技术接纳度的关键变量，在面对新技术时，老年人会对犯错表现出高度敏感，害怕因操作失误而显得笨拙或落伍^[4]。一个充满术语、缺乏引导、容错率低的系统，会直接挫伤其自尊，导致其从尝试转向彻底放弃。孤独感作为老年阶段普遍的心理状态，深刻影响着学习行为。远程教育若仅是单向的知识灌输，则无法满足其对情感互动与社会归属的渴望。学习过程中的陪伴与交流，是其坚持下去的重要动力。此外，技术焦虑则是阻碍其开启学习旅程的初始障碍，对未知数字设备的恐惧、担心操作失误会造成不可逆的“损坏”，以及对复杂界面的天然抵触，共同构成了强大的心理阻力^[5]。

1.3 远程教育交互需求

老年学习者的交互需求直接源于其生理与心理特征的叠加效应。视觉与运动控制的限制，要求交互界面必须具备大字体、高对比度设计，并以语音交互作为主要或辅助的控制

1. 江苏开放大学 江苏南京 210036

[基金项目] 基于演化博弈的老年教育“校-社-家”联动学习机制优化研究（2024LY005）

通路,以替代精细的触控操作。为应对认知负荷,信息架构必须扁平化,导航路径应极度简短且目标明确,确保用户在任何界面都能清晰预知操作后果并能便捷返回。为消解技术焦虑与维护学习自尊,系统必须建立高容错机制。这意味着交互流程具备清晰的撤销路径,对错误操作提供引导式纠正而非生硬的错误提示,从而构建一个无压力的探索环境。为满足情感与社会归属需求,交互设计必须超越纯粹的功能性。系统需扮演主动的、伴随式的角色,通过即时的、人性化的反馈给予鼓励,将单向的知识获取过程转变为双向的情感交流,以此维持学习的内在动力。综上,老年远程教育的核心交互需求,是构建一个生理上无障碍、认知上低负荷、操作上高容错、情感上强支持的自适应人机交互系统。

2 基于知识图谱与情感计算的对话式教学伴侣

2.1 模型核心思想

为精准响应老年学习者的复合需求,模型提出一种认知任务与情感交互解耦的并行处理架构,核心在于构建双重智能引擎。其一,知识图谱驱动的认知引擎,负责解析用户关于课程内容的结构化查询,通过实体链接与关系推理,提供精准、可解释的知识性回答,直接降低信息获取的认知负荷。其二,基于预训练语言模型与情感计算的情感引擎,专门处理非任务型对话,通过实时情感极性与意图识别,生成具有共情与鼓励性质的响应,旨在缓解用户的技术焦虑并满足情感陪伴需求^[6]。从而使教学伴侣能够动态权衡知识传递的精确性与人际交互的温度,实现功能与情感的深度融合。

2.2 联合技术模型架构

2.2.1 感知输入层

针对老年学习者可能存在的口音、语速变化及背景噪声等声学问题,以及其话语中认知意图与情感状态常交织呈现的特点,本研究设计了一个集成了声学增强与多任务语义解析的感知输入层。基于深度学习理论,旨在从原始、嘈杂的语音流中高效、准确地提取结构化信息。首先,为解决前端信号质量问题,引入一个基于深度神经网络的语音增强(Speech Enhancement)模块,该模块在ASR解码前对输入音频进行预处理,最大化信噪比。随后,处理后的声学观测序列 O_{clean} 被送入一个针对老年人声学特性进行适配的端到端自动语音识别ASR引擎。该引擎的核心解码任务是寻找最优词序列 W ,其目标函数由贝叶斯决策理论定义为:

$$W^* = \arg \max_W P(O_{\text{clean}} | W)P(W) \quad (1)$$

式中: W 为最优词序列, O_{clean} 为增强后的声学观测序列, $P(O_{\text{clean}}|W)$ 为声学模型给出的似然度,通过在老年人语音数据集上微调预训练模型获得; $P(W)$ 为语言模型概率,通过融入老年教育垂直领域语料库进行定制,以提升特定场景的识

别精度。

为解析转录文本中蕴含的复杂语义,本层摒弃了串行处理的传统范式,提出一个基于多任务学习的联合意图-情感识别模型,能够同步处理意图分类与情感识别两个子任务,通过共享底层网络参数,捕捉二者间的潜在关联。联合模型的训练目标是最小化一个加权组合损失函数,用公式表示为:

$$L_{\text{total}} = \lambda L_{\text{intent}} + (1 - \lambda) L_{\text{emotion}} \quad (2)$$

式中: L_{total} 为总损失, L_{intent} 与 L_{emotion} 分别为意图分类和情感分类的交叉熵损失函数, λ 为调节两个任务重要性的超参数。

此设计使系统能够一次性判断用户是发起知识查询、下达操作指令还是进行情感表达,并同步感知其情绪状态,如困惑、沮丧或喜悦^[7]。最终,感知输入层输出一个结构化的数据元组,包含转录文本、意图分类的概率分布向量和情感状态的强度向量,为协同处理核心的智能决策提供全面、精准的输入。

2.2.2 协同处理核心

协同处理核心是模型的决策中枢,针对传统对话系统中知识问答与情感交互流程割裂、响应策略僵化的问题,本研究设计了一个基于动态门控与多引擎协同的自适应处理架构,旨在实现对老年学习者复杂意图的精准路由以及对认知和情感需求的同步满足^[8]。架构主要由对话意图与情感联合分析、知识图谱问答(KBQA)和共情对话生成三大模块构成。

对话意图与情感联合分析模块负责对感知层的输出进行精细化解析和动态任务分派。现有的对话管理系统在处理混合意图时常采用基于规则或分类的硬分配策略,容易导致误判,为解决此问题本研究引入了一种软决策门控机制,该机制基于用户意图的概率分布向量 $\mathbf{v}_{\text{intent}}$ 计算一个连续的门控激活值 g ,用于动态调节后续处理模块的激活权重,用公式表示为:

$$g = \sigma(W_g \mathbf{v}_{\text{intent}} + b_g) \quad (3)$$

式中: g 是门控激活值,其值域为 $(0, 1)$,代表了用户意图的“知识查询强度”; σ 是Sigmoid函数, W_g 和 b_g 是门控网络的可训练参数。当 g 值趋近于1时,系统判定为强知识导向,将主要计算资源分配给KBQA模块;当 g 值趋近于0时,则判定为强情感或闲聊导向,主要激活共情对话生成模块。这种软分配方式相比传统阈值判断能更灵活地处理意图模糊的边缘情况,从而实现资源的优化配置。

知识图谱问答(KBQA)模块作为模型的IQ引擎,旨在解决通用问答系统在专业领域的知识准确性不高的问题,在实际设计中本研究提出搭建面向老年远程教育的专用知识图谱,并设计一个基于语义相似度与实体流行度相结合的实体链接算法,在处理用户问题时为文本提及 m 匹配知识图谱中

的最佳实体 e ，其评分函数 $S(e|m)$ 用公式表示为：

$$S(e|m) = \alpha \cdot P(e|m) + (1 - \alpha) \cdot \text{sim}_{\text{BERT}}(c_m, c_e) \quad (4)$$

式中： $P(e|m)$ 是利用语料库统计得到的实体 e 在提及 m 出现时的先验概率，解决了歧义实体的消解问题； $\text{sim}_{\text{BERT}}(c_m, c_e)$ 是利用预训练模型 BERT 计算的提及上下文 c_m 与实体描述上下文 c_e 的深度语义相似度，提升了对长尾、不规范实体提及的识别能力； α 是超参数。在精准链接实体后，系统通过关系抽取模型将问题转化为结构化查询并从图谱中提取答案。

共情对话生成模块作为模型的 EQ 引擎，旨在解决当前对话机器人情感表达模板化、缺乏真实感的问题，负责生成具有情感温度的响应。模块接收情感向量和可选的 KBQA 模块输出结果。当意图为非知识导向时，模块采用基于条件生成对抗网络（Conditional GAN）的架构，以情感向量为条件，生成符合当前对话氛围的回复。其生成器 G 的目标是最小化与判别器 D 的对抗损失，同时最大化生成文本的情感一致性。

2.2.3 响应生成与输出层

响应生成与输出层的功能是将协同处理核心输出的内部表征，解码并渲染为多模态输出。其处理流程包含响应融合、文本生成和多模态渲染三个阶段。第一阶段为响应融合，接收协同处理核心生成的知识图谱问答模块的答案表征、共情对话生成模块的文本表征，以及门控激活值。根据门控激活值，此模块将知识答案与共情话语进行加权融合，形成一个统一的语义表征。文本生成环节，融合后的语义表征被输入自然语言生成解码器，解码器基于预训练的 BART 模型，并针对老年人的语言特点进行风格微调。通过在专门构建的老年友好型语料库上训练，解码器生成简短、贴近生活的文本字符串。最终，系统进行多模态呈现。生成的文本字符串被同时传递到两个输出通道。文本渲染通道在图形用户界面上，以大字号和高对比度的色彩显示文本。语音合成通道则调用定制的文本到语音合成引擎。该引擎在生成语音时，设定合适的语速，限制基频范围，并通过数字信号处理增强中低频段的能量。系统同时输出清晰的语音和高可读性的文本，构建双通道冗余输出，确保信息传递的有效性。

2.3 模型工作流程示例

为阐明模型内部的数据流转与模块协同机制，设定一个具体的用户输入：“唉，这个‘高血压’的饮食控制太复杂了，我记不住”。感知输入层接收该语音流，其 ASR 引擎首先将其转录为文本。联合意图-情感识别模块并行处理该文本，输出一个结构化数据包，其中包含文本字符串、一个指向“知识问答”的高概率意图向量，以及一个表征“沮丧”与“困惑”的复合情感向量。

协同处理核心接收此数据包。其对话意图分析模块的门控单元依据意图向量计算出一个较高的门控激活值，指示系统以知识传递为主要任务，但保留对情感通道的关注。

KBQA 模块被优先激活，对输入文本执行实体链接，将“高血压”精确映射至知识图谱中的对应节点，并抽取“饮食控制”的关系意图。模块随即在图谱中执行查询，检索到知识三元组（高血压、recommended_diet、低盐饮食），并将核心知识点“低盐饮食”作为结构化答案输出。

与此同时，共情对话生成模块接收到“沮丧”的情感向量以及来自 KBQA 模块的知识约束“低盐饮食”。该模块的生成网络不直接输出生硬的知识，而是执行情感包装操作，生成一个融合了安抚性前导话语与核心知识的完整语义表征。最终，响应生成与输出层接收此语义表征。其 NLG 解码器将其解码为流畅的自然语言文本：“别着急，控制饮食确实需要一点耐心。关键是坚持低盐饮食，可以先从每天少放一点盐开始。”该文本随即被分发至 TTS 引擎，合成为具有优化声学参数的语音，并同步在用户界面以高可读性标准呈现，完成一次完整的认知与情感协同交互。

2.4 针对成就感需求的技术设计

为精准回应老年学习者的成就感需求，本研究利用自然语言生成（NLG）的动态反馈模块实时追踪并分析学习者的交互行为日志，捕捉课程完成度、学习时长和练习正确率等关键绩效指标，并在系统中预设了一个多维度的成就触发规则引擎，当用户的行为数据组合满足特定阈值时就会自动触发反馈流程。规则引擎的判定逻辑如表 1 所示。

表 1 成就触发规则与数据输入示例

成就类型	触发条件	数据输入样例
专注学习者	单次学习会话时长超过 30 min	指标：时长， 数值：30 min
快速上手奖	首次任务完成时间少于 5 min	指标：速度， 任务：首次登录
精准操作员	测验模块正确率高于 90%	指标：准确率， 模块：线上挂号
探索先锋	点击独立功能点的数量超过 5 个	指标：探索度， 数量：5 个

一旦触发，对应的结构化数据即被传递至一个数据到文本（Data-to-Text）的自然语言生成（NLG）模型。该模型采用 T5（Text-to-Text Transfer Transformer）架构，并在一个包含正向情感与鼓励性话语的语料库上进行了微调。模型的输入是格式化的文本，如“生成表扬：准确率，模块：线上挂号”，其输出则为具体、非模板化的鼓励性文本，如“您在‘线上挂号’练习中表现出色，掌握了所有关键步骤”。

成就摘要模块在用户完成一个完整的课程单元后被激活。该模块首先调用一个基于 BERT 的抽取式摘要模型。此模型将课程单元的全部文本材料作为输入，通过计算每个句子的向量表征与整个文档表征的余弦相似度，对句子进行排序并抽取排名前三的核心知识点。与此同时，一个关键短

语抽取算法（KeyBERT）被应用于同一文本，识别并提取出代表新掌握技能的 3~5 个核心术语。最后，抽取的知识点摘要与技能关键词被整合为一个结构化的 JSON 对象，该对象被发送至前端渲染引擎，自动生成一张可视化的“学习成就卡片”。

3 应用效果评估模型构建

3.1 多维度评估框架与指标体系

为全面、客观地度量本研究设计的 NLP 应用系统的实际效能，本研究构建了一个整合技术性能、用户体验和学习成效的三位一体的评估框架，简称“TEE（technology-experience-effectiveness）模型”。技术性能采用客观量化指标，包含针对语音识别模块的词错误率（WER）、自然语言理解模块的意图识别准确率与槽位填充 F_1 值、知识问答模块的平均倒数排名（MRR）以及系统的端到端响应延迟。用户体验指标聚焦于老年学习者在使用过程中的主观感受与行为表现，采用主客观结合的度量方法，包括任务完成率、平均任务耗时与操作错误率等客观数据，以及一系列主观数据，如国际标准的系统可用性量表（SUS）评估的易用性、用户的满意度问卷（CSAT）评估整体满意度，以及技术焦虑量表（TAS）的前后测衡量系统对用户技术焦虑的缓解效果。学习成效指标通过准实验设计进行评估，其核心指标包括通过前测与后测对比计算的知识保留率、特定生活技能的操作成功率以及通过后台日志分析得出的学习参与度，其中后者包含学习总时长、交互频率与课程完成率等细分指标。

3.2 实验设计与数据采集方案

3.2.1 准实验设计与被试招募

研究采用一项被试间准实验设计，旨在隔离并检验对话式教学伴侣系统对老年用户学习行为与体验的净效应。实验设置一个实验组与一个对照组。实验组被试使用本研究开发的基于知识图谱与情感计算的对话式教学伴侣系统。对照组被试则使用一款主流的、仅提供视频播放与暂停功能的非交互式远程教育应用，以此作为基线参照。因变量则对应 TEE 模型定义的全部指标，涵盖技术性能、用户体验与学习成效三个维度。为确保实验内部效度，所有被试接受的课程内容、学习任务与实验环境均保持严格一致。被试招募遵循目标抽样原则，计划招募 60 名年龄介于 60~75 周岁的老年人，并将其随机分配至实验组与对照组，每组 30 人。纳入标准为：具备基本的智能手机操作能力；母语为普通话；无严重视力、听力或认知障碍，后者通过简易精神状态检查量表（MMSE）进行筛选，得分需高于 24 分。排除标准为：已熟练掌握实验所选课程的知识技能；或有相关领域专业背景。在分组后，将对两组被试在年龄、性别、受教育程度以及基线技术焦虑

水平等人口统计学变量上进行均衡性检验，以排除潜在的混淆变量干扰。

3.2.2 实验流程与数据采集

实验严格遵循前测 - 干预 - 后测流程。被试在完成基线评估后，被随机分配至实验组（使用本系统）或对照组（使用传统非交互视频系统），在受控环境中完成 45 min 的指定学习任务。干预期间，系统后台自动记录交互日志；干预结束后，通过问卷与访谈采集主观数据。对所有采集数据进行 t 检验与方差分析，两组在核心评估指标上的表现呈现显著差异。实验结束后，对采集的关键指标数据进行统计分析，两组被试的核心表现对比如表 2 所示。

表 2 核心评估指标对比分析

评估维度	具体指标	实验组 (本系统)	对照组 (传统系统)	p 值
用户体验	系统可用性（SUS 得分）	86.2（±5.1）	61.5（±7.3）	<.001
	任务完成率	95.8%	82.1%	<.001
学习成效	知识保留率（提升）	46.1%	27.9%	<.001
	技能操作成功率	92.5%	64.3%	<.001
学习参与度	平均交互频率（次/会话）	36.7	3.5	<.001

数据分析结果清晰地表明，本研究设计的对话式教学伴侣系统在各个维度均显著优于传统系统。在用户体验层面，实验组极高的 SUS 得分与任务完成率证明了系统在降低操作门槛与认知负荷方面的有效性。在学习成效层面，实验组的知识保留率与技能操作成功率远超对照组，这主要取决于系统的即时问答（IQ 引擎）与情感支持（EQ 引擎）功能将学习过程从被动接受转变为主动探索与即时解惑。学习参与度数据显示实验组用户的交互频率是对照组的十倍以上，这反映了老年学习者对对话式交互模式的高度接纳与主动使用意愿，进一步验证了本模型在激发学习动机与维持学习黏性方面的核心价值。

4 总结与讨论

本研究构建了一套融合多模态感知、认知推理与情感交互的 NLP 教学系统，显著优化了老年远程教育的交互模式与学习成效。该系统在感知层实现语音增强与情感识别协同，在处理核心中引入的动态门控机制能够有效调配知识与情感响应资源，提升了系统应对复杂意图的适应性。实验数据表明该系统在可用性、知识保留与学习参与度方面均优于传统方案，验证了其在激发学习动机与缓解技术焦虑方面的有效性。未来研究可进一步拓展模型在多语言、跨平台场景下的适应能力，并探索其在老年康复、生活服务等领域的延伸应用。

基于深浅卷积注意力的改进 Transformer 的音乐流派分类研究

彭 扬¹ 李昌翰^{1*} 李旭航¹ 叶星吟¹ 颜镜竹¹ 杨羽西¹ 范晨希¹

PENG Yang LI Changhan LI Xuhan YE Xingyin YAN Jingzhu YANG Yuxi FAN Chenxi

摘 要

随着数字音乐的快速发展,海量的音乐作品不断涌现,如何高效、准确地对音乐流派进行自动分类成为智能音乐推荐、检索和版权管理中的关键问题。传统基于手工特征或浅层神经网络的方法在复杂、多样的音乐数据集上往往难以获得理想性能。针对这一问题,文章提出一种基于深浅卷积注意力的改进 Transformer 与 ArcFace 的音乐流派分类方法。此方法首先利用深、浅两级 ResNet 提取多尺度特征,并通过多头注意力机制增强时序特征建模能力,从而获得更具判别性的全局特征表示。随后引入 ArcFace 分类器提升类间可分性,并结合交叉熵损失与三元组损失进行联合优化。实验结果表明,所提方法相较于传统的机器学习与深度学习模型具有更高的分类性能,验证了所提方法在音乐流派分类任务中的有效性和鲁棒性。

关键词

Transformer; ResNet; 多头注意力; ArcFace

doi: 10.3969/j.issn.1672-9528.2026.06.010

0 引言

在数字音乐快速发展的背景下,音乐流媒体平台的普及带来了前所未有的音频资源获取方式,但也随之产生了对自动化管理和智能分析的迫切需求。尤其在音乐信息检索(music information retrieval, MIR)领域,如何准确识别和分类音乐流派已成为一项关键任务。这不仅关系到个性化推荐、版权

保护与内容管理等应用场景,也反映了深度学习在非结构化音频数据处理上的能力与挑战。传统依赖人工设计特征或浅层模型的方法在处理复杂多样的音乐信号时往往表现不足,难以兼顾分类的准确性与泛化性^[1]。因此,探索能够更好建模音频特征的深度学习方法,已成为提升音乐流派自动分类性能的重要方向^[2]。

近年来,研究者针对音乐流派分类的多样化挑战提出了多种深度学习方法以提升分类性能。Dong^[3]针对传统方法难以捕捉时序依赖的问题,将整段音频划分为3 s片段,并利用卷积神经网络(CNN)学习局部时频特征后通过投票实现

1. 重庆科技大学 重庆 401331

[基金项目] 重庆市社科联博士项目(2023BS035); 重庆市教委科学技术研究项目(KJQN202301503)

参考文献:

- [1] 唐箭. 提示工程赋能老年教育治理研究[J]. 科技资讯, 2025, 23(16):242-245.
- [2] 唐箭. 自然语言处理技术在老年教育治理中的应用研究[J]. 科技资讯, 2025, 23(15):218-221.
- [3] 王洋, 李代萍, 杨燕武, 等. 人工智能在老年医学教学中的应用现状及前景[J]. 实用老年医学, 2025, 39(7):753-756.
- [4] 唐箭. 基于腾讯混元大模型的老年教育智能体构建与应用研究[J]. 科技资讯, 2025, 23(13):224-229.
- [5] 张国丰, 马玉环. AI背景下老年思政教育创新路径研究[J]. 吉林广播电视大学学报, 2025(3):73-75.

- [6] 景燕敏. 基于智能化技术的老年教育有声在线教学系统设计[J]. 软件, 2024, 45(7):32-34.
- [7] 赵洋洋. 人工智能环境下老年群体协作学习环境建设路径[J]. 中国成人教育, 2022(24):29-33.
- [8] 王思瑶, 马秀峰. 人工智能与老年教育深度融合研究[J]. 成人教育, 2022, 42(9):43-50.

【作者简介】

祁萌(1997—), 女, 黑龙江绥化人, 硕士研究生, 助理实验师, 研究方向: 老年教育、实验室管理。

(收稿日期: 2025-10-27 修回日期: 2026-05-29)