

特色专题

大语言模型的情感智能及其心理学应用

伍海燕¹, 何翠琳¹, 曲由之², 刘泉影^{2*}

摘要 探讨了大语言模型(LLMs)是否能够通过语言任务的预训练而产生情感智能,并从心理学的角度提出了测试大语言模型的情感智能水平的方法,以及情感智能大模型在心理学的应用。从情感智能的定义出发,回顾了大语言模型的发展历程,总结出目前LLMs相关研究主要集中于提升模型在自然语言处理任务中的性能,而忽视了模型的情感智能,即感知、理解、使用和管理情绪的能力。尽管LLMs在处理语言任务方面表现出色,但在理解和生成情感文本方面仍存在局限性,因此,提出对大模型的情感智能水平进行全面评估的必要性。从心理学角度出发,总结了LLMs情感智能评估相关的工作,倡议心理学把大语言模型作为新的研究对象,重点关注大模型的情感道德问题。从情感认知科学的角度,论述大语言模型作为一种工具,对新时代的情绪心理学研究的重要价值。

关键词 情感智能;大语言模型;多维社会认知能力;情绪心理学

中国古代经典论述中已有情感智能机器相关的描述。以《列子·汤问》中的一则故事为例,匠人偃师制造了一个机器人献给周穆王,这个机器人能歌善舞,简直与真人一样。周穆王便邀请自己的妃嫔一同观赏机器人的表演。表演快要结束时,机器人突然眨了眼睛,还做出挑逗妃嫔的举动。这令周穆王非常愤怒,他以为偃师用真人冒充机器人来欺骗他,便下令要处死偃师。偃师便将机器人拆开,周穆王检查后发现这确实是由材料制成的机器人。这则故事侧面说明古代对于用人工材料组装成机器人的想象,其不仅外貌和动作模仿真人,而且还能表达感情与情欲,甚至与真人无异。时至今日,这种类人水平的情感智能机器人还未完全实现。但是最近一年如

生成式预训练模型(GPT)、大型语言模型元架构(LLaMA)等大语言模型(LLMs)的出现,重新塑造了人们对于人工智能(AI)能力范畴的认识。回顾近期大语言模型情感智能的发展轨迹,大语言模型是否已经具备了真正的情感智能呢?如何评估大模型的情感智能水平?如果大语言模型具备了情感智能,人类将会面临哪些新顾虑?将如何拓展情感心理学研究范畴?本文将针对这些问题展开论述。

1 情感智能:以人类为对象和以大模型为对象的研究

1.1 情感智能的定义

情感智能,也被称为情绪智能,其一般包含4个维度,即个体感知、理解、使用、管理自己和他人情绪的能力^[1-2]。其中,情绪感知是指识别自己和他人的

1. 澳门大学认知与脑科学中心及心理学系,澳门 999078

2. 南方科技大学生物医学工程系,深圳 518055

收稿日期:2023-09-05;修回日期:2024-02-20

基金项目:深港科技计划(C类)(SGDX2020110309280100)

作者简介:伍海燕(通信作者),助理教授,研究方向为社会神经科学,电子信箱:haiyanwu@um.edu.mo;刘泉影(共同通信作者),助理教授,研究方向为用于神经科学的人工智能,电子信箱:liuqy@sustech.edu.cn

引用格式:伍海燕,何翠琳,曲由之,等. 大语言模型的情感智能及其心理学应用[J]. 科技导报, 2025, 43(3): 47-58; doi:10.3981/j.issn.1000-7857.2023.09.01352

情绪,以及理解这些情绪背后的含义,不仅包括识别语言中的情绪,也包括识别非语言信号,如面部表情和身体动作。情绪理解涉及理解情绪的复杂性,以及不同情绪之间的关系。例如,焦虑可能会转变为愤怒,失落可能会发展为抑郁。情绪使用涉及利用情绪来促进思考和解决问题,例如,利用自己的情绪帮助自己进行决策,或利用他人的情绪帮助其改变看问题的角度。情绪管理则指能适当地回应和管理自己和他人的情绪,包括在情绪激烈的情况下仍能保持冷静,以及帮助他人处理情绪。

情感智能是人类社交互动中的重要组成部分,对个体的生存和发展都发挥着重要作用。在以人为对象的研究中发现,情感智能较高的个体的生活满意度、自我接受度和自尊水平都更高^[3],同时,情感智能也能有效预测社交关系满意度^[4]。已有研究表明,情感智能与学生的学术表现之间存在正相关关系^[5]。情感智能还被视为职业成功的一个重要预测因子,尤其是在工作满意度方面^[6]。情感智能是人类个体能力的重要特性。

在情感心理学领域,已经发展出一系列量化评估人类不同维度的情感智能的方法,主要分为“量表”和“任务测试”2种评估方法。量表是测量人类情感智能运用最为广泛的工具,主要由被评估者进行自我报告,有时候也会加入他人报告,以减轻自我意识带来的偏见。Bourne等^[7]开发了情感能力量表(emotional competency inventory, ECI),这一360°评估工具,可以评估个体或组织情感智能水平;Schutte等^[8]开发了自我报告情感智能测试(self-report emotional intelligence test, SREIT),其建立在个体对自我和他人情绪的意识、评价、表达、调节和利用其解决问题的基础上。任务测试则通过个体在实际任务中的表现,来量化评估情感智能,目前最为经典的是Mayer等^[9]开发的Mayer-Salovey-Caruso情感智能测试(Mayer-Salovey-Caruso emotional intelligence test, MSCEIT)。该测试包括一系列情绪相关的问题,通过测试个体在情感智能的4个维度上的表现,最终生成每个维度的分数,以及情感智能的总分。此外,心理学实验中的一些范式也可以反映个体情感智能的一些维度,如情绪加工任务(emotional processing)^[10]、情绪注意偏向任务(emotional bias task, EBT)^[11]。情

绪加工任务中,参与者需要感知和处理不同类型的情绪刺激,反映了个体的情感智能中的感知能力维度。情绪注意偏向任务,如情绪Stroop范式^[12],则不仅要求参与者感知情绪,还涉及个体对情绪的控制。这些评估方法为理解人类个体的情感智能水平提供了工具。

1.2 以AI为对象的情感智能研究

有关AI情感智能研究的历史可以追溯到20世纪90年代。当时学术界开始关注人类的情感计算和情感智能,但是主流智能算法为支持向量机和决策树等经典机器学习算法,性能有限,因此,AI的情感智能在当时并没有受到广泛关注。随着深度学习的发展,AI在机器视觉(如图像分类、目标识别、语义分割等)和自然语言处理(如文本分类、提问和回答、机器翻译等)等众多领域上取得性能上的突破。作为自然语言处理的任务之一,文本的情感分析(sentiment analysis)主要关注对带有情感色彩的主观性文本进行情感分类、情感分析和情感抽取的过程,也受到了更多的关注^[13-17]。有趣的是,近期也有研究人员开始以大模型为对象,研究其价值取向、人格特质等^[18-19]。

在AI的情感分析领域,产生了一系列的带情绪标签的本文数据集,包括Yelp-5数据集^[20](包含5个评级标签,每一类有650000条训练样本和50000条测试样本)、Yelp-2数据集^[21](包含正负情感标签,每一类有560000条训练样本和38000条测试样本)、CBET数据集^[22](包括9种情绪,共76860条推文)、ISEAR数据集^[23](包含7种情绪,收集于多位受访者)、SST数据集^[24]等英文数据集。中文的数据集有OCEMOTION^[25](包含7种情绪类别的中文细粒度情感性分析数据集,其中7种情绪类别分别为悲伤、幸福、厌恶、愤怒、喜欢、惊讶、恐惧),以及带情绪标注的微博语料库^[26]。研究人员使用ERNIE Tiny深度学习模型对中文文本中的情绪进行正负情绪二分类,实验结果表明,在中国歌曲评论中,正性情绪和负性情绪的得分分别可以达到93.00%和90.14%^[27]。这也证明了AI模型已经在一定程度上具备了识别情绪的能力。

大语言模型的情感智能有更为高级的体现,例如,LLMs可以根据用户的情感倾向,为其提供个性化的推荐和建议,如金融建议^[28]。研究表明,通过

LLMs与用户之间的多轮对话互动,LLMs可以识别用户的情感状态、喜好和兴趣^[29],也可以根据用户的情绪状态推荐适合用户情感的音乐和电影等^[30]。但是,LLMs的情感分析能力更多是反映其情绪感知和情绪理解2个维度的情感智能,对LLMs的情绪使用和情绪管理2个维度的研究依然不足。LLMs是否具有自身情绪的意识,是否具备管理自己情绪的能力和共情能力,是否可以利用情绪来促进思考和决策,尚无定论^[31-32]。

1.3 开发具有情感智能的大模型

开发具有情感智能的智能体是未来目标之一。具有情感智能的AI,也被称为情感AI,是指让AI具备4个维度的情绪智能,即理解、模拟人类情感,并具备对自身进行情绪管理和情绪利用的能力。拥有情感智能可以让AI更好地适应人类的社会环境,更高效地与人类进行交互,从而更好地服务人类、融入人类社会。大语言模型的出现为情感AI的实现带来曙光。

大语言模型的情感智能可以通过多种技术途径来增强。一种方法是构建更丰富的带情感标签的文本数据集,利用有监督学习的方法微调预训练好的大模型,从文本数据中学习人类的情绪表达和情绪反应^[33]。另一种方法是利用基于人类反馈的强化学习(RLHF)方法^[34],让大语言模型与人类的情感价值

函数对齐,使其具备与人类相似的情绪感知、情绪理解、情绪管理和情绪利用的能力。

赋予大语言模型情感智能并不是一项简单的任务。常用的大模型的预训练任务不能确保训练出来的LLMs具备情感智能。首先,情绪是一种非常复杂的心理现象,它涉及认知、生理和行为等多个层面。因此,要让机器理解和模拟人类的情绪,就需要让机器理解和模拟这些复杂的神经心理过程。其次,情绪是一种高度个体化和社会化的现象,每个人和每个文化对情绪的理解和表达都是不同的,而让机器理解和模拟这种多样性是非常困难的。最后,因为目前的LLMs尚没有具身性,没有感知觉输入的LLMs很难对人类情绪“感同身受”。总之,LLMs的情感智能是一个重要的、充满挑战的研究领域,它需要情绪心理学理论的发展,需要深入理解情绪的复杂性和多样性,更对大模型的训练方法和理论提出了新要求。图1概括了具有情感智能的大模型及其与情绪心理学的交互作用,重点关注3个方面,情感智能研究、情感智能评估,以及新的工具和对象。一方面,以大模型为研究对象的情感心理学研究,可帮助应对大语言模型情感智能评估需求和对伦理和社会规则的挑战;另一方面,将大模型作为情感心理学的研究工具,可拓展现有的人类情感的心理学研究领域和方法。

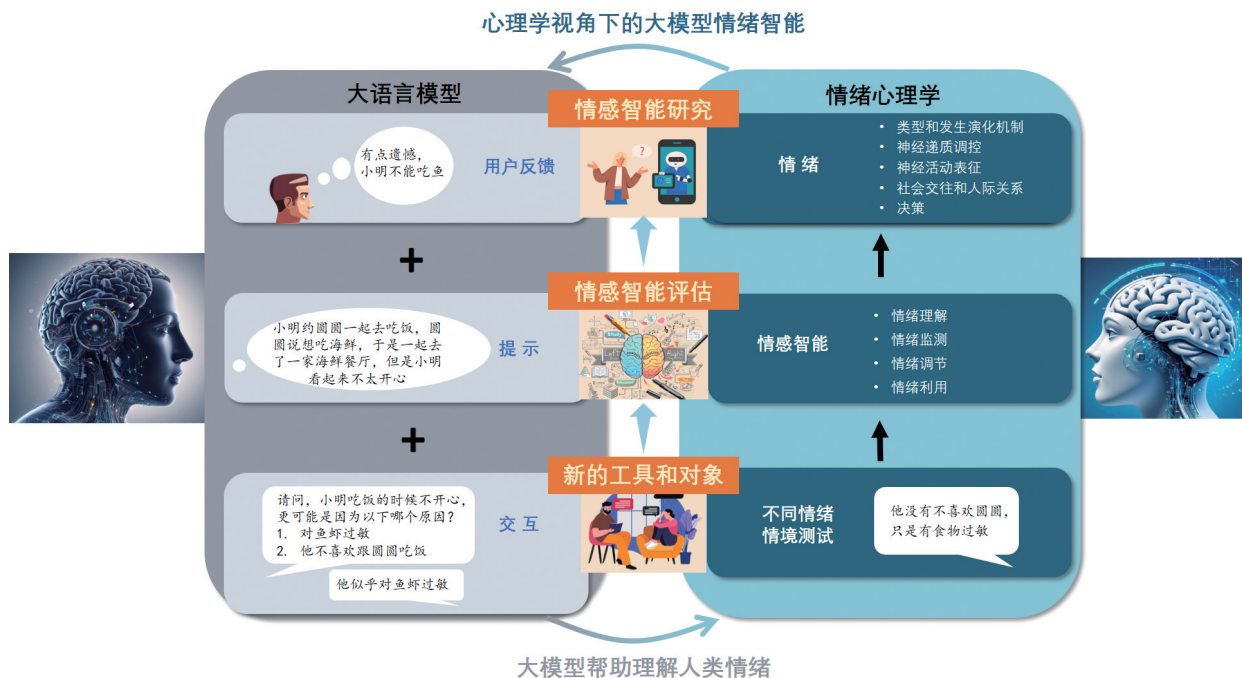


图1 具有情感智能的大模型及其与情绪心理学的交互作用

1.4 情绪具身性和大模型对具身性的模拟

情绪通常被定义为对个体重要事件的复杂心理和生理反应^[35],不仅包括主观的感受,还包括生理激活、行为表达等成分。Wilson^[36]因此提出了情绪的具身性观点,认为情绪体验和表达是通过身体的感觉运动系统实现的。Niedenthal等^[37]通过各种实验设计,如身体动作操纵、面部表情反馈等,揭示了身体状态如何影响情绪的处理和体验,也为情绪的具身理论提供了支持。当前的大模型,已经在多个领域展现出了卓越的性能,但是在模拟人类情绪的具身性方面存在明显的限制^[38-42]。首先,其缺乏对生理信号的直接处理能力,例如,无法模拟心跳加速或肌肉紧张等生理反应对情绪体验的影响;其次,大模型通常依赖于预先定义的特征和标签来训练,而不是实时的身体或者生理指标反馈,这限制了它们在动态环境中模拟情绪反应的能力。未来的研究需要提高大模型在模拟情绪具身性方面的能力。例如,集成多模态数据,包括生理信号、面部表情和身体动作等,以构建更全面的大模型;开发新的算法和技术来模拟身体状态对情绪体验的动态影响。考虑到个体差异和文化因素在情绪体验中的重要性,未来的情绪LLMs应该能够适应不同个体和不同文化背景下的情绪表达和情绪体验。

2 评估大模型的情感智能

评估大模型的情感智能具有重大意义。从计算角度,研究者可能首要关注大模型在情感计算中的表现,如情感分类任务。一系列研究表明,与其他基线小模型(如RoBERTa-base、Word2Vec、BoW等基础语言模型)相比,大模型在文本情感分类与语音情感识别任务中均展现出较好的情绪计算表现^[43-46]。此外,从心理学角度去评估大模型的情感智能,也具备多方面的意义。(1) 有助于提升大模型性能:增强情感智能水平,大模型可以更好地理解和表达人类的情感。这不仅可以提高大模型与人类的交互质量,使其更符合人类的交流习惯,也可以使大模型在处理涉及情感的问题时表现得更加人性化。(2) 促进大模型在心理学领域的应用:大模型的情感智能可以使其在许多实际应用中发挥更大的作用。例如,在

客户服务、心理咨询、教育等领域,具有情感智能的大模型可以帮助人们更好地处理和理解复杂的情感问题。(3) 提高大模型的心理可信度和接受度:如果大模型不能有效理解和表达人类情感,人们可能会对其产生怀疑和不信任。从心理学角度评估和增强大模型的情感智能水平可提高大模型可信度和接受度。总之,基于心理学的大模型情感智能评估,不仅可以帮助更好理解和处理人类情感,大模型也能在多种场景下更有应用价值,还为构建更可信、更人性化的AI奠定了重要基础。

2.1 现有大模型已具备一定情感智能

大语言模型通过海量文本进行自监督预训练,这些模型可以学习并理解词汇、语法和语义关系,并已经展现出一定的情感智能。大语言模型表现出识别和理解文本中情绪的能力,Wang等^[41]的研究评估了大语言模型的情感智能,包括情绪识别、解释和理解。研究人员开发了一种新的心理测量评估方法,重点关注情绪理解,并使用超过500位成年人类的测试结果作为参考框架,测试了多种主流的大语言模型。大多数大语言模型的得分都高于平均水平,其中GPT-4的情商得分超过了89%的人类参与者。大语言模型还能基于上下文对情绪进行推理,Li等^[47]研究发现大语言模型中不同的人工神经元群组可以代表情感的不同概念属性,经过特定操纵后,大语言模型在大多数情绪推理任务中的表现下降。其中特定属性的表现下降与人类心理空间中不同属性的重要程度有关。此外,大语言模型具备生成情绪文本的能力,Huang等^[48]基于心理学的情绪评价理论(emotion appraisal theory)研究了大语言模型的共情能力,具体为大语言模型的情感在特定情况下如何发生变化。研究发现大语言模型通常可以对特定情境做出适当的反应。这些研究表明,LLMs已经在一定程度上展现出情感智能。

然而,尽管大语言模型在一定程度上具备了理解、生成和响应人类情感的能力,但这并不意味着它们真正理解了情感的本质。事实上,这些模型可能只是通过学习大量的文本数据,学会了识别和生成包含情绪的语言。大语言模型的情感智能还可能会受到数据偏见的影响。研究者发现,大模型在焦虑调查问卷中具有比一般人类更高的焦虑评分,大模

型的决策判断会愈来愈趋向于冒险^[49]。需要更多的研究来深入理解大模型的“情感智能”，并确定它是否真的等同于人类的情感认知能力。

2.2 从心理学角度评估大模型的情感智能水平

目前,关于大模型的情感智能的评估方法主要分为2类:一是通过观察模型输出来评判其情感智能,但这种方法的准确性有待提高;二是通过让模型预测人类反馈以评估情感智能,但该方式的实用性有限。由于情感的复杂性和主观性,以及当前测试方法具有的局限性,应从心理学和认知科学的角度引入一套更准确有效的方法来测试大模型的情感智能^[32]。为了应对这些挑战,需要从心理学角度出发全面评估大模型的情绪智能水平,主要从以下几个方面展开讨论。

1) 大模型的情感智能评估方法。对大模型的情感智能水平进行多角度、全方位的评估,需要采用多种方法和指标,来分别评估模型的情绪感知、理解、表达、生成等能力。对大模型的情感智能评估方法涵盖人工评估、自动化的情感评估指标和结合多模态数据等方法。其中,人工评估是一种直观的情感智能检测方法,可以提供关于大模型情感智能的初步认识。例如,专家可以对模型生成的情感文本进行评估,并将其与人类判断进行比较。评估人员可以根据情感的准确性、一致性和流畅度等指标对大模型的情感生成能力进行评分。虽然此方法具有一定的主观性和较高的时间成本,但其评估结果更直观且贴近人类判断。利用自动化的情感评估指标,是一种更为高效和可扩展的量化大模型情感智能的方式。这些指标主要是针对情感领域的文本模式、情感词汇和情感语义的分析。例如,通过使用情感词典来计算文本中正负面情感词汇的比例,并根据比例推断情感色彩。通过多模态数据对大模型的情感智能进行评估可以更全面地理解和评估大模型的情感智能,例如,通过结合文本、图像、音频和视频等多种模态的数据。这种多模态学习的方法可以通过将情感文本与相关的图像、音频或视频进行关联,以更准确地捕捉和理解情感的多样性和复杂性。例如,中国科学院自动化研究所和清华大学合作提出的 AffectGPT 模型是一个音频-视频-文本对齐的情绪识别大模型^[50]。此外,他们还提出了“可解释的多

模态情感推理”(EMER)的任务,与之前传统的情绪识别任务不同,EMER 更进一步要求为这些情绪预测提供了解释^[50]。

为评估大语言模型的情感智能,计算机方向通常采用基于语言的情感分析任务(例如,识别、分类文本中的情感)。但是,这类情感分析任务更偏向于评估模型对于文本语言的理解能力,而非其情感智能水平。因此,有必要从心理学角度引入大模型的情感智能评估方法。清华大学刘嘉团队开发了关于情感理解的新型评估方式,其中包括经过精心设计的学校、家庭等情景,旨在引发积极和消极情绪^[41]。不仅大语言模型的情感智能评估方式存在新变化,情绪相关理论也为大语言模型的情感智能的提升发挥了重大作用。基于自我监控(self-monitoring)、社会认知理论(social cognitive theory)、认知情绪调节理论(cognitive emotion regulation theory),通过设计使用带有情绪提升相关的提示词(emotion prompt),大语言模型回答更真实、更具有责任感,在情感相关任务中的性能有明显的提升^[51]。

2) 大模型情感智能评估的结果分析与理解。理解和解读大模型的情感智能评估的结果,对于了解大模型的情感智能水平十分关键,尤其是大模型的评估结果可能会有假阴性和假阳性的问题。首先,如果模型能够生成与指定情感相符的文本,则该模型在情感智能方面具有良好的表现。然而,由于情感是一个主观且多样的概念,在不同的人和情境下可能有不同的解读。因此,解读结果时应该考虑到不同人的主观评价差异及上下文信息。其次,分析模型的情感理解能力。情感理解是指模型对输入文本中的情感信息进行识别和理解的能力。如果模型能够准确地识别文本中的情感倾向和情感表达,并根据上下文做出恰当的回答,这表明该模型具有理解文本上下文复杂性和准确解析情感信息的能力。此外,需要综合多种评估方式结果,考虑数据偏见和模型的局限性,减少因模型的特定应用场景和目标的限制,以获得更全面、准确和客观的情感智能检测结果^[52-53]。综上所述,解读大模型在情感智能方面的检测结果涉及多个方面,包括情感生成和理解能力、评估方法的多样性、潜在的数据偏见和模型局限性,以及应用场景的特殊需求。获取更全面、准确和有

意义的解读结果,对于情感智能的进一步研究和改进具有指导意义。

3) 大模型的情感智能评估的价值。大模型的情感智能评估对于大模型的开发和应用具有重要的价值与意义。对大模型进行情感智能评估,了解模型对情感信息的理解、识别和生成能力,进而为优化模型和构建更高质量的数据集提供指导。大模型训练过程所采用的大规模数据库中可能包含主观偏见、文化差异和情感倾向等问题,导致大模型对情感信息处理的准确性受到影响。通过情感智能评估,可及时发现并纠正问题,有助于模型的改进和数据集质量的提升,从而提高模型情感智能水平。具有高水平的情感智能大模型,在情感分析、情感机器人、智能对话系统等应用领域将发挥突出的作用。例如,西湖大学开发的心理健康支持对话陪伴机器人,可以提供情感支持和心理疏导^[54]。随着 LLMs 在工作和生活中日益普及,人们对其情感智能能力的期望也越来越高。用户希望可以与大模型进行更加自然流畅的情感交流,并得到具有情感的回应。情感智能评估可以确保大模型能够迎合人们的情感需求,提供令人满意的情感智能服务。这对于提升用户体验、增加用户忠诚度和推动智能技术的广泛应用具有重要意义。通过对大模型的情感智能评估,可以深入了解大模型对情感信息的理解和生成,为进一步研究和改进大模型的情感智能水平提供启发。大模型的情感智能评估不仅可以促进 LLMs 技术的发展,还能使 LLMs 为人类提供更人性化的情感智能服务。

3 大语言模型作为情绪心理学的研究对象

3.1 大模型的情感引发的伦理道德问题及其对拓展道德研究领域的贡献

2023年4月,斯坦福大学以人为本人工智能研究所发布了《2023年人工智能指数报告》,指出:2022年,生成模型成为时代思潮的一部分,这些模型表现良好,但同时也伴随着道德上的挑战。人们产生的伦理道德担忧可能主要集中在大模型无法承担法律责任,却可能诱导违背社会规范的行为^[55]。这是因为大语言模型在训练过程中,往往会接触到大量的网络文本,其中包含各种观点和立场。这些观点和立

场在很大程度上是多元化的,但也可能存在一些错误、偏见或歧视等。因此,当大语言模型生成内容时,很可能会导致内容的失真和偏见。已有研究表明,LLMs 普遍存在对于男性和女性职业的有偏见的假设,LLMs 的观点更符合大众的普遍看法,而不是官方就业统计数据所反映的事实^[56]。Coda-Forno 等^[57]发现,当存在引发焦虑的文本提示时,GPT3.5 生成的回答中显示出偏见的大幅增加。因此,需谨慎地对待 LLMs 生成的情感文本,并对其可信度进行审慎评估。另外,大语言模型在处理用户输入的文本时,可能会涉及用户的隐私信息。例如,在情感分析任务中,大语言模型需要分析用户输入的文本中所表达的情感倾向。然而,这些文本可能包含了用户的个人信息、家庭状况、健康状况、价值取向等敏感信息,可能导致隐私泄露和其他伦理道德问题。

此外,在道德情绪和道德决策领域,大模型既更新了已有研究的范式,也拓展了其研究的子领域。例如,矣晓沅等^[58]从交叉学科的角度提出了大模型与人类道德价值对齐的新的概念性范式,并从道德对齐角度、评估方式和实际方法3个层面进行阐述。通过道德对齐的方式有望提升大模型的安全性能。也有研究人员利用道德机器(moral machine)框架来调查几种 LLMs (如 GPT-3.5、GPT-4、PaLM 2 和 Llama2) 的道德决策倾向。比较它们对人类偏好的反应发现,虽然 LLMs 和人类的偏好大体一致(例如优先考虑人类而不是宠物,以及倾向于拯救更多生命),但与人类较温和的倾向相比,LLMs 可能倾向于更不妥协的决策^[59]。Zhou 等^[60]的研究则主要关注以下问题:LLMs 能理解并遵循不同的道德理论吗?哪种理论能指导 LLMs 更好地与人类注释者的道德判断相符合?是什么原因导致了提倡的方法与现有资源的不一致?研究最终提出了一个灵活的框架来引导 LLMs 利用来自跨学科研究的成熟道德理论进行道德推理。

大语言模型虽然可能引发伦理道德的顾虑,也能在另外一个侧面拓展已有道德研究的疆域,尤其是道德情绪研究范畴。传统道德情绪研究涉及个体的道德判断、情感体验和社会认知等多个层面,却鲜少涉及当 LLMs 作为道德判断或者决策中的部分主体时,应如何做出责任和情绪的推断。通过训练 LLMs,

使其学会理解道德问题、推理和判断,可以在一定程度上模拟人类,以及人类和大模型交互中的道德决策过程,从而揭示个体在道德问题上的思维模式和偏好^[52]。例如,通过大量人类语言样本预训练的语言模型可以用来预测人类的道德判断。研究人员使用预先训练的语言模型对从流行的 Reddit 用户论坛检索到的帖子进行道德判断,获得了 87% 的正确率^[61]。此外,通过输入特定的情境和话题,LLMs 可以生成相应的道德情绪表达语句,这可以帮助研究者更好地理解个体在不同道德情境下的情感体验和态度变化^[62]。

3.2 大模型的“情绪障碍”及“治疗”方案

利用大模型来预测和辅助诊断人类情绪障碍的研究已有很多^[63-64],但探索大模型的“情绪障碍”的研究尚少。随着人类与大模型的互动频繁且日益常态化,大模型的情绪状态(如稳定性)可能会深刻影响人类情绪状态。例如,当人类个体和大模型交互时,可能面临 LLMs 输出不当的情绪表达而致情感伤害产生。LLMs 可能会以负面或冷漠的方式回应用户的情感。这类 LLMs 共情缺失可能导致用户感到不理解或心理受伤。那么,大模型是否也会出现“情绪障碍”(如焦虑、抑郁、躁狂类症状)、“情绪障碍”以及 LLMs 该如何“治疗”就成为重要的研究内容。想象这样一种 LLMs 的情绪障碍的情景,由于输入数据的不一致性或预训练过程中的噪声,LLMs 表现出情绪的不稳定性:当输入数据中包含负面情绪或争议性内容时,大模型可能会产生消极或矛盾的回答。这一方面要求情绪心理学家发展出一套检测 LLMs 情绪状态的评估方法,另一方面要求大模型的算法设计者在 LLM 算法中加入情绪障碍监测机制和自纠、自救机制。

未来大模型很可能出现情绪障碍问题,对其心理健康的关注与维持势在必行。面对这样“情绪障碍”的大模型时,LLM 算法设计者可以针对性地筛选与调整训练数据,加入多样化的情感和观点,确保涵盖不同群体的声音,这有助于模型更好地理解和回应不同情绪的输入^[65]。进一步地,如果检测到某个大模型存在严重的情绪障碍,开发者可以调整参数权重,来减轻其负面情绪、偏见等问题,进而改善模型的情绪表达能力^[66]。此外,用户的反馈也是改进大模型情绪处理能力的重要途径。通过收集和分析用户

数据,可以对大模型进行针对性的“高情商”训练。综合来看,通过数据平衡、调整模型权重和用户反馈教育等治疗方案,可以有效减轻大模型的情绪障碍,并提高其回应用户情感的能力。

4 情感智能大模型在心理学中的应用

4.1 大模型可用于情绪心理学问卷的生成和自动化分析

情绪心理学中,问卷评估是研究个体情绪状态和心理健康的一种常用工具。传统的问卷设计和分析需要大量的时间和人力投入,但 LLMs 有强大的文本生成和分析能力,将其应用于心理学问卷的生成和自动化分析,将有助于提高数据质量、分析效率,还能对个体和群体情绪状态进行精准分类和评估^[67]。

大语言模型的强大生成能力可以用于生成情绪心理学问卷,包括问题的设计、选项的制定。通过在大量的心理学和情绪学文献中学习,大语言模型根据指示(prompt),设计具有一定逻辑性和多样性的问题及与之对应的选项,以让被试者更好地表达其情绪状态。与传统的问卷生成和论证方式相比,使用 LLMs 进行问卷生成不仅更加高效,还能实现更多多样化的设计,并具备自适应调整的优势。例如,传统问卷设计需要经过多轮修改和多领域专家评定,而 LLMs 可以在较短的时间内生成具有应用潜力的问卷。这将大大缩短问卷设计周期,降低人力成本。此外,LLMs 可以学习多种问题形式和表达方式,生成多样化的问题,提高问卷的吸引力和个性化,提高被试的参与度。LLMs 可以根据被试的回答动态调整后续的问题,实现个性化的问卷设计。根据被试的情绪状态和回答模式,可以实时调整,提高问卷的有效性。例如,Kjell 等^[68]探讨了 ChatGPT 在心理治疗应用中的前景,通过在一系列的场景中进行验证,结果表明最新的人工智能技术有望将心理健康评估从传统的评分量表模式转变为更贴近自然语言交流的方式。Götz 等^[69]开发了一个基于 GPT-2 的心理测量项生成器(psychometric item generator, PIG),旨在为心理学研究领域简化开发评估量表的过程。无需编程知识仅需简单的语言提示,PIG 便可生成量表,并在实际测试中表现与当前评估的金标准相当的水平。

LLMs还可以用于自动化分析和解释问卷数据。通过分析问卷的回答,LLMs可以自动处理和清洗数据,剔除无效或重复的回答,提高数据质量。LLMs能够对问卷数据进行情绪识别与分类。通过识别分析被试回答的情绪,LLMs可以自动将回答划分为不同的情感类别,如喜悦、悲伤、焦虑等,从而提供对被试情绪状态的客观评估。多模态大模型可以根据问卷数据进行统计分析,生成相关的报告和图表。通过文本理解并自动计算统计指标,如均值、标准差和相关系数,使分析结果更加准确和可靠。

4.2 多模态大模型可用于情绪实验任务中刺激材料的生成

多模态大语言模型(multimodal large language models, MLLMs)是近年来兴起的一个新的研究热点,具有处理多种类型数据输入(如文本、图像、声音等)的能力^[70-72]。MLLMs具有理解和生成多模态类型数据的能力,从而完成更加复杂的任务,如基于文本生成图像、包含图像与问题的视觉问答等^[73]。研究报道对比语言-图像预测训练(contrastive language-image pretraining, CLIP)多模态大模型中具有与人脑相似的多模态神经元,而且还有专门响应情绪的神经元可以识别情绪^[74-75];改进的EmotionCLIP仅使用未经整理的数据即可从言语和非言语交流中提取视觉情感表征^[76]。在心理学领域,情绪实验旨在研究个体情绪体验和情绪加工的过程。为了激发受试者的情绪,实验任务通常需要使用不同情绪类刺激材料。MLLMs具有多模态数据处理与生成能力,可以生成多模态的刺激材料,以更有效地激发被试的各种情绪。多模态大模型可以通过文本等多类型的指示,生成包含如文本、图像等多模态且具有情绪性的刺激材料。例如,在情绪实验中,研究人员可以提供一个描述情感事件的文本,多模态大模型可以通过生成相关的图像来补充和激发被试的情绪。多模态大模型可以根据被试个体的特征和偏好生成个性化的刺激材料。通过学习被试的历史数据和反馈,多模态大模型可以生成更符合被试独特喜好和情绪体验的刺激材料,提高实验的个体化程度。多模态大模型可以模拟各种情境和情感情景,帮助研究人员设计实验条件。例如,在研究情绪加工的注意偏向时,多模态大模型可以生成不同情境下的图像和声

音刺激,模拟真实世界中不同情感状态下的注意分配。

多模态大模型不仅增加了情绪实验刺激材料的真实性和多样性,还可以提供更加灵活的刺激材料,并提高实验的可控性和个性化水平。研究人员通过调整各种类型的输入指示,如文本和图像,使多模态大模型生成更加符合特定实验设计的刺激材料,提高实验的多样性与可控性。多模态大模型还可以通过学习被试的反馈和情感反应,以实现实验设计的个性化和自适应。

4.3 对大模型的情绪操控为理解人类情绪障碍提供新思路

情绪障碍如抑郁症、焦虑症等对人的心理健康和生活质量有着重大的负面影响。大模型的快速发展可能为情绪障碍的研究和临床应用开辟了新的可能性。通过采用包含情绪状态的文本构建大模型,研究人员可以将大模型作为研究的对象,对大模型的情绪状态进行分析与操控。通过观察大模型在不同情绪状态下的反应和表现,可以推断出它在不同情绪状态下如何理解和解释信息,有助于理解情绪的形成过程、情绪与认知之间的关系,以及情绪障碍的产生机制。

中国科学院软件研究所、微软公司,以及威廉与玛丽学院的研究人员提出Emotion Prompt方法,利用心理学的知识对大语言模型进行引导,发现积极的暗示不仅可以调节大模型的情绪,还能够提升大模型回答的质量^[51]。通过对大模型的情绪操控(例如,设计特定的指令以增强或减弱情绪),研究形成一套有效的情绪调节策略。这不仅有助于个体管理其情绪,也为干预情绪障碍提供了新的途径。进一步地,通过分析情绪对大模型决策的影响,可以更准确地了解情绪与认知之间的相互作用。这将有助于识别情绪障碍的情感和认知异常,从而为情绪障碍的治疗提供更有效的治疗方案。大模型也有助于预防情绪障碍及其相关的自伤自杀行为的发生。通过分析庞大的数据集,大模型可以识别与情绪障碍或自伤和自杀行为有关的特定模式和趋势,帮助社会组织和相关机构更准确地把握情绪障碍和自伤自杀行为的普遍性和地理分布,以便更有效地分配资源和提供定制的援助和干预措施。

5 结语

针对大模型的情感智能的主题,总结了现有研究中大模型体现出的情绪分析和理解能力,提出从心理学角度来评估大模型的情感智能水平。一方面,希望更多以大模型为研究对象的情感心理学研究出现,以应对大语言模型对伦理和社会规则的挑战;另一方面,应将大模型作为情感心理学的研究工具,以丰富现有的人类情感的心理学研究方法。此外,大模型在心理健康领域的应用将增加,如预测精神疾病的风险因素,帮助实现早期识别和干预,降低治疗成本等。伴随这一应用,对于数据隐私、算法偏见等问题的关注也会随之增长。未来的研究将更加注重心理学家、计算机科学家、伦理学家等多领域专家的合作,开发符合伦理标准的情绪AI系统。

参考文献(References)

- [1] Mayer J D, Roberts R D, Barsade S G. Human abilities: Emotional intelligence[J]. *Annual Review of Psychology*, 2008, 59: 507–536.
- [2] Salovey P, Mayer J D. Emotional intelligence[J]. *Imagination, Cognition and Personality*, 1990, 9(3): 185–211.
- [3] Carmeli A, Yitzhak-Halevy M, Weisberg J. The relationship between emotional intelligence and psychological wellbeing [J]. *Journal of Managerial Psychology*, 2009, 24(1): 66–78.
- [4] Lopes P N, Salovey P, Straus R. Emotional intelligence, personality, and the perceived quality of social relationships[J]. *Personality and Individual Differences*, 2003, 35(3): 641–658.
- [5] Petrides K V, Frederickson N, Furnham A. The role of trait emotional intelligence in academic performance and deviant behavior at school[J]. *Personality and Individual Differences*, 2004, 36(2): 277–293.
- [6] Urquijo I, Extremera N, Azanza G. The contribution of emotional intelligence to career success: Beyond personality traits[J]. *International Journal of Environmental Research and Public Health*, 2019, 16(23): 4809.
- [7] Bourne P, Hanna I, Solomon T, et al. Primal leadership: Realizing the power of emotional intelligence: A book review [J]. *COJ Nursing & Healthcare*, 2017, 1(2): 46–50.
- [8] Schutte N S, Malouff J M, Bhullar N. The assessing emotions scale[M]//*Assessing Emotional Intelligence*. Boston, MA: Springer, 2009: 119–134.
- [9] Mayer J D, Salovey P, Caruso D R, et al. Measuring emotional intelligence with the MSCEIT V2.0[J]. *Emotion*, 2003, 3(1): 97–105.
- [10] Keightley M L, Winocur G, Graham S J, et al. An fMRI study investigating cognitive modulation of brain regions associated with emotional processing of visual stimuli[J]. *Neuropsychologia*, 2003, 41(5): 585–596.
- [11] Koster E H W, Crombez G, Verschuere B, et al. Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage[J]. *Behaviour Research and Therapy*, 2004, 42(10): 1183–1192.
- [12] Williams J M, Mathews A, MacLeod C. The emotional Stroop task and psychopathology[J]. *Psychological Bulletin*, 1996, 120(1): 3–24.
- [13] Agarwal A, Xie B, Vovsha I, et al. Sentiment analysis of twitter data[C]//*Proceedings of the Workshop on Language in Social Media*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2011: 30–38.
- [14] Feldman R. Techniques and applications for sentiment analysis[J]. *Communications of the ACM*, 2013, 56(4): 82–89.
- [15] Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: A survey[J]. *Ain Shams Engineering Journal*, 2014, 5(4): 1093–1113.
- [16] Minaee S, Kalchbrenner N, Cambria E, et al. Deep learning: Based text classification[J]. *ACM Computing Surveys*, 2022, 54(3): 1–40.
- [17] Zhang L, Wang S, Liu B. Deep learning for sentiment analysis: A survey[J]. *WIREs Data Mining and Knowledge Discovery*, 2018, 8(4): e1253.
- [18] Cahyawijaya S, Chen D L, Bang Y J, et al. High-dimension human value representation in large language models[EB/OL]. [2024-03-05]. <https://arxiv.org/abs/2404.07900v3>.
- [19] Serapio-García G, Safdari M, Crepy C, et al. Personality traits in large language models[EB/OL]. [2024-03-05]. <https://arxiv.org/abs/2307.00184v3>.
- [20] Xie Q Z, Dai Z H, Hovy E, et al. Unsupervised data augmentation for consistency training[C]//*Proceedings of Advances in Neural Information Processing Systems*. New York: Curran Associates, 2020: 6256–6268.
- [21] Zhang X, Zhao J B, LeCun Y. Character-level convolutional networks for text classification[C]//*Proceedings of Advances in Neural Information Processing Systems*. New

- York: Curran Associates, 2015: 649–657.
- [22] Shahraki A G, Zaiane O R. Lexical and learning-based emotion mining from text[C]//Proceedings of the International Conference on Computational Linguistics and Intelligent Text Processing. Switzerland: Springer, 2017: 24–55.
- [23] Scherer K R, Wallbott H G. Evidence for universality and cultural variation of differential emotion response patterning[J]. *Journal of Personality and Social Psychology*, 1994, 66(2): 310–328.
- [24] Socher R, Perelygin A, Wu J Y, et al. Recursive deep models for semantic compositionality over a sentiment treebank [C]//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA, USA: Association for Computational Linguistics, 2013: 1631–1642.
- [25] Li M, Long Y, Qin L, et al. Emotion corpus construction based on selection from hashtags[C]//Proceedings of the Tenth International Conference on Language Resources and Evaluation. Portorož, Slovenia: European Language Resources Association, 2016: 1845–1849.
- [26] 姚源林, 王树伟, 徐睿峰, 等. 面向微博文本的情绪标注语料库构建[J]. *中文信息学报*, 2014, 28(5): 83–91.
- [27] Huang Y X, Jusoh S. Sentiment detection through emotion classification using deep learning approach for Chinese text [C]//Proceedings of 15th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). Piscataway, NJ: IEEE, 2023: 1–6.
- [28] Xing F. Designing heterogeneous LLM agents for financial sentiment analysis[EB/OL]. [2024–03–05]. <https://arxiv.org/abs/2401.05799v1>.
- [29] Sun X F, Li X Y, Zhang S Y, et al. Sentiment analysis through LLM negotiations[EB/OL]. [2024–04–06]. <https://arxiv.org/abs/2311.01876v1>.
- [30] Kang I, Ruan S K, Ho T, et al. LLM-augmented preference learning from natural language[EB/OL]. [2024–05–05]. <https://arxiv.org/abs/2310.08523v1>.
- [31] Cao S Q, Fu D, Yang X, et al. Pain recognition and pain empathy from a human-centered AI perspective[J]. *iScience*, 2024, 27(8): 110570.
- [32] Qu Y Z, Du P H, Che W X, et al. Promoting interactions between cognitive science and large language models[J]. *The Innovation*, 2024, 5(2): 100579.
- [33] Liu Z W, Yang K L, Xie Q Q, et al. EmoLLMs: A series of emotional large language models and annotation tools for comprehensive affective analysis[C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM, 2024: 5487–5496.
- [34] Ouyang L, Wu J, Xu J, et al. Training language models to follow instructions with human feedback[C]//Proceedings of Advances in Neural Information Processing Systems. New York: Curran Associates, 2022: 27730–27744.
- [35] Albert S. Cognition and emotion[M]. Oxford: Oxford University Press, 2000.
- [36] Wilson M. Six views of embodied cognition[J]. *Psychonomic Bulletin & Review*, 2002, 9(4): 625–636.
- [37] Niedenthal P M, Barsalou L W, Winkielman P, et al. Embodiment in attitudes, social perception, and emotion[J]. *Personality and Social Psychology Review*, 2005, 9(3): 184–211.
- [38] Romera-Paredes B, Barekatin M, Novikov A, et al. Mathematical discoveries from program search with large language models[J]. *Nature*, 2024, 625(7995): 468–475.
- [39] Thirunavukarasu A J, Ting D S J, Elangovan K, et al. Large language models in medicine[J]. *Nature Medicine*, 2023, 29(8): 1930–1940.
- [40] Demszky D, Yang D Y, Yeager D S, et al. Using large language models in psychology[J]. *Nature Reviews Psychology*, 2023, 2: 688–701.
- [41] Wang X N, Li X T, Yin Z, et al. Emotional intelligence of large language models[J]. *Journal of Pacific Rim Psychology*, 2023, 17: 1–12.
- [42] Xi Z H, Chen W X, Guo X, et al. The rise and potential of large language model based agents: A survey[EB/OL]. [2024–05–05]. <https://arxiv.org/abs/2309.07864v3>.
- [43] Amin M M, Cambria E, Schuller B W. Will affective computing emerge from foundation models and general artificial intelligence? A first evaluation of ChatGPT[J]. *IEEE Intelligent Systems*, 2023, 38(2): 15–23.
- [44] Zhang Z, Peng L, Pang T, et al. Refashioning emotion recognition modelling: The advent of generalised large models [J]. *IEEE Transactions on Computational Social Systems*, 2024, 11(5): 6690–6704.
- [45] Yang H, Zi Y, Qin H, et al. Advancing emotional analysis with large language models[J]. *Journal of Computer Science and Software Applications*, 2024, 4(3): 8–15.
- [46] Xu Y, Chen H, Yu J, et al. Secap: Speech emotion captioning with large language model[C]//Proceedings of the AAAI Conference on Artificial Intelligence. California,

- USA: Association for the Advancement of Artificial Intelligence, 2024: 19323–19331.
- [47] Li M, Su Y S, Huang H Y, et al. Language-specific representation of emotion-concept knowledge causally supports emotion inference[EB/OL]. [2024-05-05]. <https://arxiv.org/abs/2302.09582v5>.
- [48] Huang J T, Lam M H, Li E J, et al. Emotionally numb or empathetic? evaluating how LLMs feel using EmotionBench [EB/OL]. [2024-06-10]. <https://arxiv.org/abs/2308.03656v6>.
- [49] Liu J L, Liu C, Zhou P L, et al. Is ChatGPT a good recommender? A preliminary study[EB/OL]. [2024-06-10]. <https://arxiv.org/abs/2304.10149v3>.
- [50] Lian Z, Sun H Y, Sun L C, et al. Explainable multimodal emotion recognition[EB/OL]. [2024-06-10]. <https://arxiv.org/abs/2306.15401v6>.
- [51] Li C, Wang J, Zhu K J, et al. Large language models understand and can be enhanced by emotional stimuli[EB/OL]. [2023-08-01]. <https://arxiv.org/abs/2307.11760v7>.
- [52] Schramowski P, Turan C, Andersen N, et al. Large pre-trained language models contain human-like biases of what is right and wrong to do[J]. *Nature Machine Intelligence*, 2022, 4(3): 258–268.
- [53] Motoki F, Pinho Neto V, Rodrigues V. More human than human: Measuring ChatGPT political bias[J]. *Public Choice*, 2024, 198(1): 3–23.
- [54] Qiu H C, He H L, Zhang S, et al. SMILE: Single-turn to multi-turn inclusive language expansion via ChatGPT for mental health support[EB/OL]. [2024-06-10]. <https://arxiv.org/abs/2305.00450v3>.
- [55] Krügel S, Ostermaier A, Uhl M. The moral authority of ChatGPT[EB/OL]. [2024-09-10]. <https://arxiv.org/abs/2301.07098v1>.
- [56] Kotek H, Dockum R, Sun D. Gender bias and stereotypes in large language models[C]//Proceedings of The ACM Collective Intelligence Conference. New York: ACM, 2023: 12–24.
- [57] Coda-Forno J, Witte K, Jagadish A K, et al. Inducing anxiety in large language models can induce bias[EB/OL]. [2024-09-11]. <https://arxiv.org/abs/2304.11111v2>.
- [58] 矣晓沅, 谢幸. 大模型道德价值观对齐问题剖析[J]. *计算机研究与发展*, 2023, 60(9): 1926–1945.
- [59] Takemoto K. The moral machine experiment on large language models[J]. *Royal Society Open Science*, 2024, 11(2): 231393.
- [60] Zhou J Y, Hu M D, Li J N, et al. Rethinking machine ethics: Can LLMs perform moral reasoning through the lens of moral theories? [EB/OL]. [2024-09-11]. <https://arxiv.org/abs/2308.15399v2>.
- [61] Alhassan A, Zhang J, Schlegel V. "Am I the Bad One"? Predicting the Moral Judgement of the Crowd Using Pre-trained Language Models[C]//Proceedings of the Thirteenth Language Resources and Evaluation Conference. Marseille, France: European Language Resources Association, 2022: 267–276.
- [62] Scherrer N, Shi C, Feder A, et al. Evaluating the moral beliefs encoded in llms[C]//Proceedings of Advances in Neural Information Processing Systems. New York: Curran Associates, 2023: 51778–51809.
- [63] Yang K L, Ji S X, Zhang T L, et al. Towards interpretable mental health analysis with large language models[C]//Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA, USA: Association for Computational Linguistics, 2023: 6056–6077.
- [64] Van Heerden A C, Pozuelo J R, Kohrt B A. Global mental health services and the impact of artificial intelligence-powered large language models[J]. *JAMA Psychiatry*, 2023, 80(7): 662–664.
- [65] Bender E M, Gebru T, McMillan-Major A, et al. On the dangers of stochastic parrots: Can language models be too big?[C]//Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. New York: ACM, 2021: 610–623.
- [66] Liu Y, Yao Y S, Ton J F, et al. Trustworthy LLMs: A survey and guideline for evaluating large language models' alignment[EB/OL]. [2024-09-11]. <https://arxiv.org/abs/2308.05374v2>.
- [67] Li C, Fung M, Wang Q, et al. MentalArena: Self-play Training of Language Models for Diagnosis and Treatment of Mental Health Disorders[EB/OL]. [2024-09-11]. <https://arxiv.org/abs/2410.06845v1>.
- [68] Kjell O N E, Kjell K, Schwartz H A. Beyond rating scales: With targeted evaluation, large language models are poised for psychological assessment[J]. *Psychiatry Research*, 2024, 333: 115667.
- [69] Götz F M, Maertens R, Loomba S, et al. Let the algorithm speak: How to use neural networks for automatic item generation in psychological scale development[J]. *Psychological Methods*, 2024, 29(3): 494–518.

- [70] Huang S H, Dong L, Wang W H, et al. Language is not all you need: Aligning perception with language models[C]// Proceedings of Advances in Neural Information Processing Systems. New York: Curran Associates, 2023: 72096–72109.
- [71] Linear-probe A. Learning transferable visual models from natural language supervision[C]//Proceedings of International Conference on Machine Learning. New York: ACM, 2021: 8748–8763.
- [72] Zhu X X, Guo C P, Feng H Y, et al. A review of key technologies for emotion analysis Using Multimodal information [J]. Cognitive Computation, 2024, 16(4): 1504–1530.
- [73] Vaiani L, Quatra M L, Cagliero L, et al. Viper: Video-based perceiver for emotion recognition[C]//Proceedings of the 3rd International on Multimodal Sentiment Analysis Workshop and Challenge. New York: ACM, 2022: 67–73.
- [74] Goh G, Cammarata N, Voss C, et al. Multimodal neurons in artificial neural networks[J]. Distill, 2021, 6(3): e30.
- [75] Bondielli A, Passaro L C. Leveraging CLIP for image emotion recognition[C]//Proceedings of the 5th Workshop on Natural Language for Artificial Intelligence. Aachen, Germany: CEUR-WS, 2021: 172–185.
- [76] Zhang S T, Pan Y M, Wang J Z. Learning emotion representations from verbal and nonverbal communication[C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2023: 18993–19004.

Emotional intelligence of large language models and its psychological applications

WU Haiyan¹, HE Cuilin¹, QU Youzhi², LIU Quanying^{2*}

1. Centre for Cognitive and Brain Sciences and Department of Psychology, University of Macau, Macau 999078, China

2. Department of Biomedical Engineering, Southern University of Science and Technology, Shenzhen 518055, China

Abstract This article discusses the possibility for large language models (LLMs) to develop emotional intelligence and the applications of emotional LLMs in psychology. We introduce ways to evaluate the emotional intelligence level of LLMs from a psychological perspective. We begin by defining human emotional intelligence and discussing the history of LLMs' development. We highlight that the focus of research of LLMs has been on optimizing their performance in natural language processing tasks, such as text understanding and generation, while neglecting the ability to identify, understand, and express emotions. Although LLMs excel in language processing they still have limitations in comprehending and generating emotional texts. To address this, we argue for the necessity of conducting comprehensive assessments of the emotional intelligence level of LLMs. From a psychological standpoint, we summarize existing work on evaluating the emotional intelligence of LLMs and advocate for psychology to consider LLMs as a new research subject, particularly in emotional and moral aspects. Furthermore, we explore the important role of LLMs as tools for studying emotional psychology in the new era from the perspective of emotional cognitive science.

Keywords emotional intelligence; large language model; multidimensional social cognitive ability; emotional psychology ●



(责任编辑 王微)