

亲子冲突类型的智能判别: 基于大语言模型的文本分析方法验证

康春花,王汉秦,倪志浩

(浙江师范大学 心理学院;浙江省儿童青少年心理健康与危机干预智能实验室,浙江 金华 321004)

摘要:研究旨在开发一种基于文本分析的亲子冲突类型智能判别工具,重点验证大语言模型在此类细粒度心理文本分类任务中的性能与可行性。首先,构建包含四种亲子冲突类型的理论框架,并据此开发亲子冲突文本测试框架。随后,应用大语言模型对冲突文本进行自动分类,并从准确率、召回率及语义理解准确性等维度,与人工标注结果进行系统对比。结果表明,相较于传统机器学习方法,大语言模型表现出更稳健的分类性能与更好的语义捕捉能力,尤其是在显性冲突识别方面效果显著;然而,在涉及隐性冲突的复杂场景中,该模型仍存在一定的局限性,需结合人工辅助进行修正。研究证实了大语言模型可作为亲子冲突文本智能判别的核心技术,为其在家庭教育与心理评估领域的应用提供方法依据。

关键词:亲子冲突;大语言模型;文本分析;智能判别;家庭教育

中图分类号:B844.2 **文献标识码:**A **文章编号:**1673-9841(2026)01-0249-12

一、引言

亲子关系指社会关系中的父母与子女以血缘和共同生活为基础,经由互动所形成的人际关系^[1]。亲子关系作为个体成长早期阶段中最为关键的社会联系之一,不仅构成了家庭结构的基础,更在儿童及青少年的心理发展中扮演着至关重要的角色,一直是发展心理学、心理咨询与实践工作者研究的焦点^[2]。亲子冲突是亲子双方发生的不一致状态^[3],可表现为观点分歧、情绪迥然,甚至言行攻击^[4-5]。青少年期(12~18岁)被称为“疾风暴雨期”,超过80%的青少年家庭在此阶段会面临至少一种不良的亲子关系^[6]。这些冲突若处理不当,可能会造成青少年在价值观、社会行为、心理健康等方面的适应失调,难以应对向成年期过渡的压力和挑战,甚至引发极端事件^[7-9]。《中华人民共和国家庭教育促进法》第十五条指出,“未成年人的父母或者其他监护人及其他家庭成员应当注重家庭建设”,“共同构建文明、和睦的家庭关系,为未成年人健康成长营造良好的家庭环境”。因而,在青少年心理发展与家庭动态的交互中,亲子冲突的精准识别与科学干预对促进家庭和谐、青少年健康成长至关重要。

作者简介:康春花,浙江师范大学心理学院,教授,博士生导师。

基金项目:教育部人文社会科学研究规划基金项目“主观题自动评分的机器学习技术及其应用系统开发”(22YJA190005),项目负责人:康春花;浙江省儿童青少年心理健康与危机干预智能实验室重点开放课题“儿童青少年心理危机识别与干预理论与应用研究”(23MHCICAZD04),项目负责人:康春花。

传统的亲子冲突识别主要依赖于自评量表、观察法等主观报告式方法^[10]。罗纳德·普林茨(Ronald J. Prinz)等开发的《冲突行为问卷》(Conflict Behavior Questionnaire, CBQ)包含青少年版(73题)和母亲版(75题),采用是、否作答形式,用于评估母子双方对彼此行为的不满及互动中的冲突表现。被试对负面方向相关条目的认可数量即为其得分,分数越高表明感知到的亲子冲突越严重^[11]。阿瑟·罗宾(Arthur L. Robin)等开发的《问题清单》(Issues Checklist)由44条与亲子冲突相关的题目组成,如家务分配、个人卫生、学校作业、社交活动选择、作息时间等,用来测试冲突发生的频繁程度和冲突中亲子双方关系的紧张程度^[12]。方晓义等开发的《青少年亲子冲突问卷》共16题,考查个体在过去半年内分别与父亲和母亲在学业、家务、交友、花钱、日常生活安排、外表、家庭成员关系和隐私等八个方面发生冲突的频率和强度^[13]。尽管自陈式量表评估便于收集大量数据,但由于量表题目的固定性及其社会赞许效应,难以有效反映亲子冲突在真实生活中的情境特征^[14]。相比而言,基于文本信息和生活情境的评估方式更具生态效度。随着测试朝着情境化和生活化的方向发展^[15],研究者开始使用观察法来评估亲子冲突过程,如使用假想冲突情境^[16]、冲突视频或图画^[17]、冲突话题讨论^[18]、回溯冲突事件的文本或语音分析^[19]等。此类评估需要专业的编码人员对视频中的言语、表情和动作行为进行编码和评分^[20-21]。尽管观察法实现了情境化和生活化评估,具有一定的生态效度,但在人员培训、技术设备等方面需要较大投入,工作可重复性较差,并未实现与之相适应的智能化分析,因此在实际应用中有较多限制^[22-23]。

文本蕴含着丰富的情感信息,文本分析是一种通过对文本数据进行处理、分析和挖掘,以发掘隐含信息和模式的方法^[24-25]。自然语言技术在心理情感文本分析中的应用^[26-28],使得基于情境或文本的智能化亲子冲突识别成为可能^[17,29]。随着深度学习技术的迅猛发展,以其为基础的大语言模型(Large Language Model, LLM)在文本分析领域展现出巨大潜力。LLM不仅在数据整理和分析方面具有巨大优势,能够高效完成诸如规划行文格式、排序任务、数据清洗、统计分析和错误甄别等任务,还能完成文本生成、分类、总结、改写等工作^[30-32]。比如,研究构建的 EmoLLMs 系列通过指令微调技术,在多任务情感分析中表现优异^[33]。在适当的提示词信息或经相关训练数据微调后,LLM可以对法规类型、是否为仇恨言论、疾病类型等进行分类,并且在检测文本语调方面具备相当的稳定性和准确率^[28,34-37]。

可见,随着心理测评的情境化、生活化和智能化发展趋势,基于文本内容的自动判别成为心理测评的主流方向^[38]。为此,本研究基于亲子冲突文本测试框架,并使用LLM实现亲子冲突类型的智能识别,为家庭教育研究的情境化、生活化和智能化探索提供一条初步路径。研究内容包括:(1)通过文献梳理,界定亲子冲突类型的内涵和特征词,构建亲子冲突文本测试框架,并验证其可行性。(2)使用传统机器学习模型和LLM对亲子冲突类型进行智能判别,并使用多元测量指标对比其判别效果,考察LLM在实际应用场景中的优势与不足。

二、亲子冲突文本测试框架与冲突类型界定

本研究旨在构建亲子冲突的文本测试框架,并确立和验证专家标注体系的有效性。具体思路为:参考以往研究,构建亲子冲突的文本测试主体和框架;通过梳理亲子冲突类型的相关理论,界定冲突类型内涵及关键词特征;收集并分析亲子冲突文本信息,采用专家评估一致性和效标量表验证测试框架和专家标注的有效性。

(一)文本测试框架建构

综合参照以往问卷法、观察法和每日日记法的内容和操作过程,本研究建构的亲子冲突文本测试

框架按顺序包括指导语、基本资料、亲子冲突小作文和效标量表四个部分。指导语是说明测试的目的和保密性原则,基本资料是调查个体的一些相关信息,效标量表是用来考察效标关联效度,这三部分与往常自陈式评估量表无异。亲子冲突小作文是文本测试框架的主体部分,其写作框架为:(1)测试内容的确定。本研究通过梳理问卷法^[39]、访谈法^[40]和每日日记法^[16,41]的相关内容,选取冲突频率较高的10个话题作为测试内容,包括家务、外表、个性/行为风格、家庭作业/学业成就、人际关系、人际活动调节、就寝时间和宵禁、健康和卫生、活动调节、财务和其他。(2)写作范式。改编自访谈法^[40]和每日日记法^[16,41]的要求和程序,亲子冲突小作文的写作范式为:要求被试在10个冲突话题中选取与父母发生冲突频率最高的2个话题进行小作文写作(每个话题约200~300字),描述冲突中亲子双方的行为、情绪以及解决冲突的方式等。如果冲突话题不在清单内,可以选择自身最高频的冲突话题进行写作。(3)写作示例。为了防止被试在写作时天马行空,无法聚焦亲子冲突的相关内容,本研究提供了一份写作示例。要求被试仿照示例分三段描述:第一段,描述冲突对象和事情经过;第二段,描述冲突时你和父母的情绪、语言(说的频次最多的词),冲突发生后你和父母分别采取了什么措施,情绪变化如何,采用措施的效果如何;第三段,你对冲突的看法,包括你感受到的父母理解如何等。

总体而言,本研究构建的测试框架,兼顾了实用性与可行性,确保能够在规定时间内顺利完成测试,并获取较充足的文本信息以支持后续分析。

(二) 亲子冲突类型界定

现有研究表明,从不同的视角可以将亲子冲突划分为不同的类型。如温德尔·弗曼(Wyndol Furman)等将亲子冲突分为语言冲突、情感冲突和行为冲突,语言冲突涉及口头争执,情感冲突可能涉及情感的冷漠或疏远,行为冲突可能表现为消极行为或反抗^[42];凯·布拉德福德(Kay Bradford)等认为亲子冲突有公开行为和对立两种形式^[43];樊召锋等将亲子冲突划分为显性冲突与隐性冲突两大类,显性冲突包含语言对抗与行为对抗,隐性冲突涵盖冷暴力(情感冻结)、软暴力(消极抵抗)等非直接对抗形式^[44];吴玉花提出四维分类体系,将亲子冲突分为:(1)教养方式冲突,源自代际教育理念差异(如专制型与民主型);(2)情绪管理冲突,由即时性情感失控引发(如愤怒宣泄);(3)价值观冲突,涉及文化资本代际传递障碍(如职业选择分歧);(4)生活方式冲突,反映现代性冲击下的习惯碰撞(如数字设备使用规范)^[45]。

本研究的核心是基于文本信息的智能判别,LLM的任务是有效提取本文信息中可识别、可编码的关键特征词,并据此实现亲子冲突类型的自动判别。因此,一个直观的、侧重于外在互动行为的分类框架,比一个侧重于内在动因的分类更能直接、有效地服务于本研究目标。弗曼等和樊召锋按照表现形式的分类,其概念直接对应具体的可操作的言行描述,易于从文本中提取关键特征词。吴玉花的根源性分类虽然深刻,但其内涵更为抽象和复杂,往往需要结合上下文深度推理才能判断冲突归属,不适用于进行快速、大规模的文本特征标注与分类。因此,本研究综合分析弗曼等和樊召锋的研究,将亲子冲突类型分为语言冲突、行为冲突、冷暴力和软暴力四种易于区别和识别的类型,其中语言冲突和行为冲突属于显性冲突,冷暴力和软暴力属于隐性冲突。如此,既能有针对性地探索LLM的特征提取能力,也能对LLM在显性冲突、隐性冲突的识别能力方面展开对比探讨。每种冲突类型的内涵及关键特征词如表1所示。

表1 亲子冲突类型的定义及关键特征词

冲突类型	定义	特征词	
显性冲突	语言冲突	双方在冲突发生时的口头对抗性互动,这种冲突通常涉及争论、指责或批评,虽然不涉及身体上的对抗,但其言语的尖锐和敌意同样能够对双方的关系造成负面影响	言语攻击、恶语相向、诽谤、诋毁、逼迫、命令式语气、否定成就、揭短、讽刺挖苦、刻薄评论、强迫服从、威胁抛弃、语言压制、过度批评、羞辱性标签、言语操控等
	行为冲突	双方在冲突发生时的不理智或过激行为,包括身体对抗、过度控制以及其他形式的身体或心理上的强制	撕毁照片、谩骂、打人、没收手机、摔平板、扯被子、体罚、抽打、罚抄课文、责骂、抢夺物品、扔东西、强制、侵犯隐私、监控等
隐性冲突	冷暴力	也称消极沟通,表现为双方在情感和语言上的交流缺失或完全停止,可能通过故意忽视、沉默或回避来表达对对方的不满,从而造成双方情感上的隔阂和沟通上的障碍,是一种非言语冲突形式	冷战、拒绝谈话、沉默、忽视、爱答不理、冷暴力、置之不理、忽视感受、沟通障碍、情绪冷漠、回避、忽视需求、故意疏远、拒绝合作、情感忽视、忽视意见、冷漠回应、拒绝道歉、忽视进步、拒绝理解等
	软暴力	也称隐性对抗,双方通过非直接或非对抗性的方式表达不满或反抗,如拖延、逆反、沉默、回避、情感封闭或离家出走等,通常是为了逃避直接的冲突或表达不满	质问、指责、否定、无视、唠叨、冷漠、嘲讽、比较、压迫、控制、命令、责怪、贬低、忽视、压力等

(三) 基于文本内容和特征词的人工分类效果验证

1. 评定者间一致性系数

首先,招募了5名兼具心理测量和心理健康学术背景的研究生组成专家小组,他们都是曾在中小学进行过心理健康教育或实习的专业人员,具备一定的理论基础和实践经验。基于表1的分类框架,在正式编码前,组织专家小组对分类框架进行深入的讨论和诠释,确保所有专家对分类标准达成共识性理解,这有助于提升编码的内部一致性。在正式分类标注阶段,5名专家独立按照表1对来自浙江省某市的159名中学生的亲子冲突类型进行标注,在标注过程中,如果出现较大分歧,一般遵循最显著、最直接的主导冲突维度原则,依据操作化定义和关键特征词进行客观锚定。在编码框架中,显性冲突(语言、行为)与隐性冲突(冷暴力、软暴力)在互动模式上存在本质区别。显性冲突通常是冲突互动中更直接、能量更强的“第一信号”,它定义了当前互动场域的主要性质,即使之后出现了隐性冲突的特征,但其主导的、定性的冲突维度仍是显性的。当文本中同时出现这两类特征时,规定优先归为显性冲突。当文本中出现的多重特征属于同一维度(显性或隐性)时,则依据其表述的直接性、激烈程度和文本篇幅比重来判断其主导类型。专家标注结果见表2,专家较为一致地把大多数的冲突标注为语言冲突和行为冲突,冷暴力和软暴力则相对较少。5位专家评定一致性 Krippendorff's α 系数为0.878,表明评定者间一致性优秀,结果高度可靠^[46-48]。

表2 5位评定者的分类标注结果

	亲子冲突类型				总计
	语言冲突	行为冲突	冷暴力	软暴力	
专家1	53	56	32	18	159
专家2	50	59	34	16	159
专家3	50	58	35	16	159
专家4	53	57	32	17	159
专家5	50	55	35	19	159

2. 效标关联效度

使用《青少年亲子冲突问卷》^[13]作为效标,量表的内部一致性系数(Cronbach's α)为0.87。于浙江省另一地市的某中学,抽取221名学生,向其发放亲子冲突文本测试工具和《青少年亲子冲突问卷》,要求参与者自愿且如实作答,其中有效问卷212份。随后,请专家小组按照表1对212名参与者的亲子冲突类型进行标注。专家标注和《青少年亲子冲突问卷》分类可类似于两位评定者之间的一致性,因而采用适于两位评定者一致性的 Kappa 系数验证亲子冲突测评框架和关键特征词的效标关联

效度。表 3 是专家标注和问卷分类对比结果,可以看出与表 2 相比,两者规律几乎相似,也是语言冲突和行为冲突占比较大,其他两类相对较少。对表 3 结果进行分析得出专家分类和问卷分类结果之间的 Kappa 系数为 0.728($p < 0.01$),表明分类一致性较高,文本测试的效度较好^[49]。

表 3 专家标注问卷评定结果列表

		问卷结果				总计
		语言冲突	行为冲突	冷暴力	冷暴力	
专家结果	语言冲突	63	1	4	0	68
	行为冲突	0	53	6	12	71
	冷暴力	7	3	31	2	43
	冷暴力	0	7	1	22	30
总计		70	64	42	36	212

三、基于 LLM 的亲子冲突智能判别

在测试框架和关键特征词的有效性得以验证后,再次收集青少年亲子冲突文本,使用 LLM 对其亲子冲突类型进行智能判别,并采用多元视角验证 LLM 在智能判别中的优势与不足,以期为基于文本的心理情感分析提供借鉴。

(一) 文本数据收集与标注

在浙江省三个地市的多所中学收集到亲子冲突文本 1 008 份(其中缺失文本共 21 份),对收集到的文本数据进行清洗,去除无关内容,以保证数据的纯净度。对 987 份文本按表 1 进行人工标注,将标注好的数据集划分为训练集、验证集和测试集,以支持模型训练和性能评估,见表 4。

表 4 数据集描述

序号	类别	训练集	验证集	测试集
1	语言冲突	182	118	73
2	行为冲突	142	91	63
3	软暴力	51	32	21
4	冷暴力	107	66	41
总计		482	307	198

(二) 模型选择

模型的选择要明确任务需求,考虑模型的规模、性能及对中文的支持情况,本研究选取 Qwen2.5-72B-Instruct,即 Qwen2.5 系列 72B 模型。Qwen2.5-72B 基于 18 万亿 token 的预训练数据,覆盖多语言、多领域文本。大规模、多样化的数据增强了模型对复杂语义的解析能力,在多语言文本分类和长文本分类任务中表现突出,尤其是对中文的优化尤为显著,并且支持长达 128K 的上下文输入,可有效捕捉长文档中的全局语义信息,避免传统模型因截断导致的分类偏差。模型采用 Apache 2.0 许可,开发者可自由微调以适应特定领域,在多项基准测试中,以不到 1/5 的参数规模超越 Llama3-405B 等更大模型。这种高效架构使其在文本分类中既能处理复杂任务,又能降低推理成本。模型在指令跟随、JSON 结构化输出等方面表现优异,可精准适配文本分类任务的格式要求。

(三) 模型判别

1. 评估思路

研究关注深度学习模型 BERT、Qwen2.5-72B 基底模型和 Qwen2.5-72B 调优模型这几个模型间的比较,探讨 LLM 在基于文本信息的亲子冲突智能判别中的优势。BERT 模型于 2018 年提出,是预训练—微调范式的里程碑。在文本分类任务中,BERT 及其变体(如 RoBERTa、DeBERTa)长期保持着顶尖的性能。作为一个参数规模适中的强大基线,与 72B 的 LLM 形成有意义的对比,能更有力地揭示 LLM 在特定任务上的潜在优势与局限。按照表 4 划分的数据集,使用训练集($N = 482$)训练

模型,调整模型的所有权重和参数,让模型学习从文本特征到冲突类型的映射关系;验证集($N=307$)用于在训练过程中间接地评估模型性能,调整超参数,如学习率、批次大小等,监控训练过程,防止过拟合(如果验证集性能下降而训练集性能上升,就是过拟合的信号);测试集($N=198$)用于在模型完全训练和调优完成后,提供一个完全独立、未见过的数据上的无偏性能评估,此结果为最终报告结果,代表模型在真实世界中的预期表现。

2. 评估过程

针对具体的亲子冲突文本,需要在阿里云百炼平台对预训练的大语言模型进行微调,调整其权重以更好地适应特定任务。首先,使用标注好的482份训练集文本对模型进行监督式微调(Supervised Fine-Tuning, SFT)训练。在微调过程中,使用指令微调以提高模型在分类任务上的表现,模型在训练中进行多次迭代,不断调整参数数值,逐渐找到一种配置,以最小化特定任务的误差。初步训练完成后,在使用阶段加入专业提示词帮助模型理解任务的性质和预期的输出,确保模型生成与任务相关且准确的回答。接着,使用307份验证集数据对调优模型进行初步评估,验证模型效果,对模型进行优化和更新。最后,使用198份测试集数据考察模型的实际效果,进行对比研究。

图1为微调流程,通过预训练模型与任务特定数据的结合,实现模型从通用语言理解到任务专项能力的转化。可分为三个阶段:首先,选择已经通过大规模语料库进行了预训练的基底大语言模型,该模型具备通用的语言表示与生成能力。其次,通过监督式微调,对预训练模型进行任务适配。模型基于这些配对数据进一步训练,学习如何根据特定指令生成符合要求的响应。最后,微调后的大语言模型具备了任务导向的生成能力,可响应特定类型的指令(如文本总结、翻译等),适应下游任务,提升样本效率与性能。

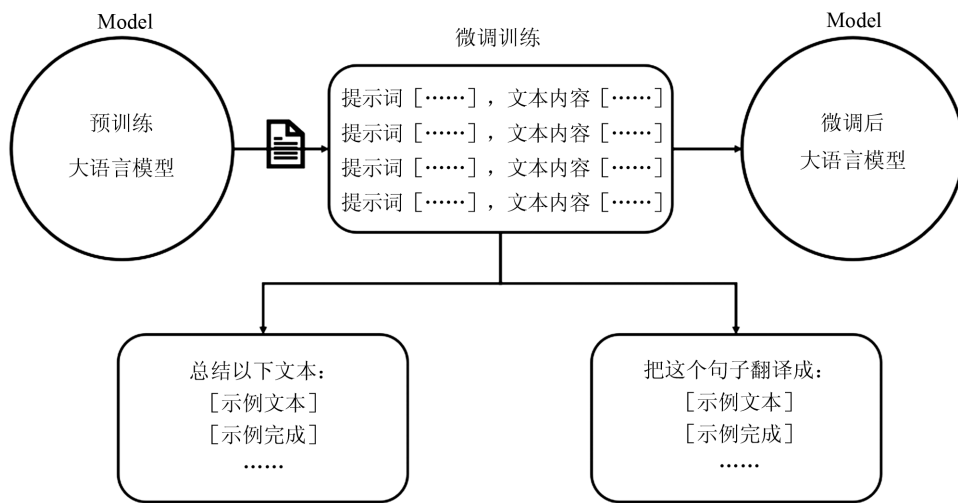


图1 指令微调原理

(四) 结果

1. 模型比较

对经过训练的模型,使用未参与训练的测试集来评估模型的性能,将模型预测结果与测试集的标签进行比较,通过比较不同模型的准确率(A)、精确率(P)、召回率(R)、 F_1 分数等指标来评估模型的准确性。这4个评价指标可用下列四个公式计算。

$$A = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{1}$$

$$P = \frac{TP}{TP + FP} \times 100\% \tag{2}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{3}$$

$$F_1 = \frac{2 \times P \times R}{P + R} \times 100\% \quad (4)$$

由表 5 可知,Qwen2.5-72B 调优模型的准确率、召回率、精确率和 F_1 分数均优于其他模型,分别比 BERT 模型高出 9.04%、10.16%、9.42% 和 5.52%,比基底模型高出 6.94%、6.47%、6.04% 和 2.93%。为进一步考察 Qwen2.5-72B 模型对不同亲子冲突类型的识别能力,对各冲突类型的分类效果进行细粒度分析,结果如表 6 所示,语言冲突的召回率、精确率和 F_1 分别为 79.45%、78.37%、78.91%;行为冲突的召回率、精确率和 F_1 分别为 77.78%、76.56%、77.17%;软暴力的召回率、精确率和 F_1 分别为 62.60%、63.15%、63.05%;冷暴力的召回率、精确率和 F_1 分别为 73.10%、71.88%、68.29%。调优模型能有效捕捉输入序列中的时序信息和上下文关系。

表 5 不同模型分类结果

模型名称	A	R	P	F_1
BERT 模型	65.20%	60.51%	62.18%	65.61%
Qwen2.5-72B 基底模型	67.30%	64.20%	65.56%	68.20%
Qwen2.5-72B 调优模型	74.24%	70.67%	71.60%	71.13%

表 6 Qwen2.5-72B 调优模型各类型分类效果

冲突类型	R	P	F_1
语言冲突	79.45%	78.37%	78.91%
行为冲突	77.78%	76.56%	77.17%
软暴力	62.60%	63.15%	63.05%
冷暴力	73.10%	71.88%	68.29%

提取关键混淆矩阵可分析模型对各个冲突类型的预测准确性,识别模型的强项和不足。深入分析模型预测错误的案例,找出可能的原因,如数据标注错误、模型不足等。图 2 为分类混淆矩阵,横轴为模型预测的类别,纵轴为人工标注的类别,图中对角线上的数据为模型预测和人工标注一致的次数,上下三角线的数据为模型误判次数。由图 2 可以看出,模型对语言冲突和行为冲突的预测较为准确,对软暴力较难判准。

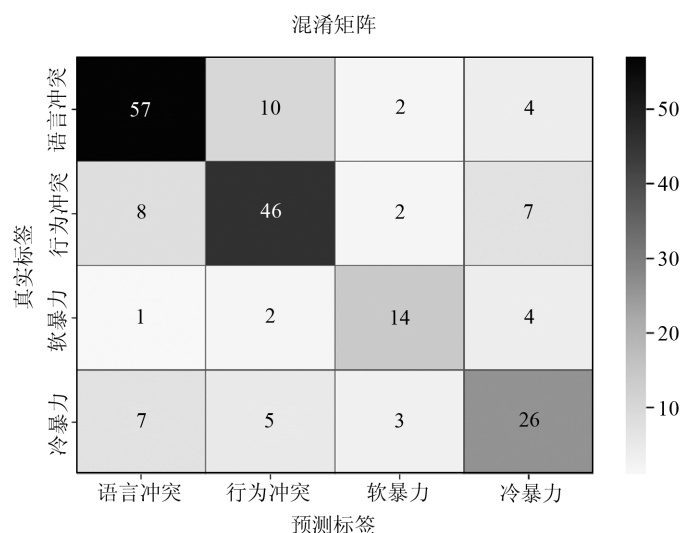


图 2 分类混淆矩阵

2. 质性验证

为进一步明晰 Qwen2.5-72B 在智能判别中的优势与不足,通过典型案例访谈对结果进行质性验证。访谈对象包括参与研究的 6 组家庭(每组含 1 名青少年及其父母),覆盖语言冲突、行为冲突、冷暴力和软暴力四种类型,通过半结构化访谈,探究模型分类结果与家庭实际体验的一致性。访谈中,

请家长与青少年分别描述该家庭亲子冲突事件的经过、情绪体验及应对策略,询问参与者对模型分类的认可度(如“您认为这次冲突属于语言冲突吗?为什么?”)。

访谈结果表明,6对家长中有4对家长完全认可模型快速分类的价值,6个孩子中有5个孩子认可模型的分类结果。比如,一位父亲表示:“如果能早点知道这是行为冲突,可能不会动手。”一名高中男生表示:“模型说这是语言冲突,很准确,我妈总拿我和别人比较,我一听就炸。”在冷暴力的案例中,一名初中女生反馈:“模型把我们的沉默对抗归为冷暴力,确实是这样。”一位母亲说:“孩子低头不说话时,模型没看出她偷偷流泪,但专家访谈注意到了。”在软暴力的案例中,模型将“随便你”(隐含消极抵抗)错误归类为冷暴力。可以看出,模型对语言冲突和行为冲突识别的准确性较高,尤其在标准化场景(如学业指责、家务争执)表现稳健,而对冷暴力和软暴力识别的准确性稍低,主要因模糊语义(如“随便你”)和非语言线索缺失(如沉默中的情感疏离)导致误判,且对冷暴力中的“情感忽视”(如长期回避眼神交流)识别率较低,需结合人工观察予以补充。

结合混淆矩阵与访谈结果可以发现,模型的智能判别与人工判别相比,分类差异主要源于以下因素:(1)语义复杂性,如“你继续玩,别管我死活”被模型归类为语言冲突(表面争执),而人工判定为软暴力(隐含情感操控);(2)多义词歧义,如“算了”在不同语境中可能表示妥协(冷暴力)或消极抵抗(软暴力),模型误判率高达34.8%;(3)代际认知偏差,如父母将“冷战”视为正常管教,人工标注为冷暴力,模型则误判为无冲突;(4)人工能更好地识别反讽、隐喻(如“随便你”隐含的软暴力),并且对地域方言的理解更准确;(5)模型易混淆语言冲突与冷暴力(如沉默中的情绪对抗)。

3. 词云图

词云图通过形成“关键词云层”或“关键词渲染”,实现对文本中高频关键词的视觉凸显,能够直观呈现不同冲突类型的语言特征,从而验证模型分类结果与表1中的定义和关键词特征是否吻合。如图3所示,从左到右依次为语言冲突、行为冲突、软暴力和冷暴力四种冲突的词云图。语言冲突词云图的高频词包括“争执”“沟通”“争吵”等,体现以言语对抗为主的冲突形式;行为冲突词云图突出“没收”“推搡”“夺走”等动作词汇,反映肢体对抗或物品损毁等显性行为对抗;软暴力词云图以“拖延”“敷衍”“否定”“无视”为核心词,呈现消极抵抗、表面妥协但实际回避的隐性冲突特征;冷暴力词云图的高频词为“沉默”“借口”“不想”“不说话”等,显示无语言互动或情感冻结的长期对抗模式。



图3 词云图

四、讨论

本研究旨在通过开发一个标准化的文本测试工具,并利用大语言模型的智能识别能力,实现对语言冲突、行为冲突、冷暴力和软暴力四种亲子冲突的自动识别。研究结果基本证实了研究路径的可行性与有效性,同时揭示了其中存在的挑战与未来改进的方向。

(一) 亲子冲突四维分类体系的理论与实践价值

本研究摒弃了传统基于策略(如合作与对抗)的分类方式,在吸收朱迪思·斯梅塔娜(Judith G. Smetana)提出的冲突情境理论^[40],融合格蕾丝·钟(Grace H. Chung)的冲突内容分类^[41],并结合了樊召锋等对东方家庭冲突特征质性观察^[44]的基础上,从冲突表现形式的本质出发,构建了语言冲

突、行为冲突、冷暴力和软暴力的四维分类体系,并详细界定了其内涵和特征词,从而形成兼顾文化特殊性与理论普适性的测量框架,具有重要的理论和实践意义。

首先,这种分类方式更精准地捕捉了当代中国家庭亲子互动,特别是冲突情境中的复杂性与隐蔽性。冷暴力(如冷漠、忽视)和软暴力(如讽刺、贬低)的提出,极大地提升了对非肢体、非直接辱骂类精神伤害的识别能力。这与当前青少年心理健康领域愈发关注“心理虐待”和“情感忽视”的趋势高度吻合。

其次,这种分类方式更贴近家庭教育指导的实际需求。相比于判断策略是“合作”还是“妥协”,家长和教育工作者更迫切需要了解的是“冲突具体是什么行为”以及“它的危害等级如何”。本研究的分类直接指向可观察、可干预的具体行为表现,如将“频繁贬低孩子”判别为软暴力,其干预方案(如改变沟通语言)与将“发生推搡”判别为行为冲突的干预方案(如情绪管理、行为约束)是截然不同的。因此,亲子冲突四维体系的界定和测量框架对于家庭教育诊断与指导具有较大的理论和实践意义。

(二)大语言模型在亲子冲突判别中的优势与局限性

本研究验证了 LLM(以 Qwen2.5-72B 为例)在执行细粒度、高专业度文本分类任务上的强大能力,其优势主要体现在:(1)对复杂和隐含语义的强大理解能力。Qwen2.5-72B 能够识别“沉默”“敷衍”等隐性冲突的文本特征^[33],在处理软暴力和冷暴力这类依赖语境和隐含意图的冲突类型时,展现出近乎专家水平的判别力,这对于传统基于词典或简单机器学习的模型而言是极其困难的任务。这证明了 Qwen2.5-72B 在理解人类复杂情感和心里互动方面具有巨大优势。(2)高效率与标准化。人类具备理解每句话背后的潜台词的能力,并能够对语义进行准确解读,从而做出精确判断,但也存在劣势,包括可重复性差、处理海量文本的能力有限以及多人协同工作时难以统一标准等。然而,LLM 一旦训练完成,便能在瞬间完成对海量文本的批量处理与判别,远超人工编码的效率,且评判标准高度一致,避免了人工编码中因疲劳、主观情绪波动造成的标准浮动问题。

以 Qwen2.5-72B 为例,LLM 的局限性也同样明显:(1)尽管 LLM 在亲子冲突分类任务中表现出较高的准确性和效率,但在处理模糊语义和隐性冲突时仍存在一定的局限性。质性访谈表明,LLM 对非结构化语义(如反讽、方言)和隐性情感线索的敏感性仍需提升。这与斯梅塔娜提出的“亲子冲突的语境依赖性”理论一致,即冲突类型的判定需结合具体情境的社交规则与文化背景^[40]。(2)伦理与偏差风险。LLM 的训练数据本身可能包含社会偏见。例如,模型可能会将某些文化背景下常见的、实则属于软暴力的指责性语言(如“你看看别人家的孩子”)误判为语言冲突。这要求我们在使用 LLM 时必须非常谨慎,需要建立人工审核机制,防止技术的误用和对某些家庭群体的歧视。

(三)智能判别结果的进一步优化

针对 Qwen2.5-72B 在智能判别中的局限性,结合访谈研究发现,本研究进一步提出了模型比较和优化思路。首先,鉴于隐性冲突(如冷暴力、软暴力)高度依赖语境、语气和非语言线索从而导致 Qwen2.5-72B 在判别上的不足,本研究进一步引入上下文语义增强机制和优化特征工程,即结合历史对话,通过同义词替换、句式变换、重点捕捉否定词密度(如“不”“没”“讨厌”)和情绪词强度(如“气炸”“绝望”)等方式,以提高模型的稳健性。结果显示,Qwen2.5-72B 优化模型的准确率、召回率、精确率和 F_1 分数分别为 78.28%、76.47%、75.88%、76.17%,相较于原调优模型分别高出 4.04%、5.8%、4.28%和 5.04%,均有提升。其次,考虑到当前研究只使用了一个 LLM(Qwen2.5-72B)进行评估,本研究进一步补充多个 LLM 间的比较探索。补充的模型包括:(1)KIMI-2 模型,该模型结合了多任务学习和深度卷积神经网络,并且具备强大的跨领域迁移能力;(2)DeepSeek 模型,是由国内科研机构开发的模型,通常用于数据挖掘、推荐系统和自然语言理解任务;(3)ChatGLM3 模型,是一种多模态生成语言模型,主要用于中文自然语言处理任务。它在对话生成、文本理解和多轮对话等方面有显著优势,且在中文处理上比其他类似模型具有更优的性能。在模型训练后,应用测试集比较各个模型的评估效果。结果显示,KIMI-2 模型的准确率、召回率、精确率和 F_1 分数分别为 76.26%、74.92%、72.78%、73.83%;DeepSeek-V3 模型的各项指标分别为 73.74%、70.07%、70.61%、70.34%;Chat-

GLM3模型的各项指标分别为69.19%、66.28%、66.70%、66.49%。其中KIMI-2模型各项指标均优于其他模型,相较于未优化的Qwen2.5-72B调优模型,分别高出2.02%、4.25%、1.18%、2.70%。最后,本研究初步探索了人机协同下,即模型完成自动判别冲突类型后,再由专家着重针对标签与预测不一致的情况,结合冲突文本进行适当的人工修正,将修正后的标签再次进行模型训练和调优,结果显示各个模型的多维指标均有所提升。KIMI-2的准确率、召回率、精确率和 F_1 分数分别为84.77%、85.10%、83.18%、84.13%,相较于人机协同前的各指标分别高出8.51%、10.18%、10.40%、10.30%;DeepSeek-V3模型的准确率、召回率、精确率和 F_1 分数分别为84.34%、84.38%、82.06%、83.21%,相较于人机协同前的各指标分别高出10.60%、14.31%、11.45%、12.87%;ChatGLM3模型的准确率、召回率、精确率和 F_1 分数分别为79.29%、80.82%、77.95%、79.36%,相较于人机协同前的各指标分别高出10.10%、14.54%、11.25%、12.87%;Qwen2.5-72B优化模型的准确率、召回率、精确率和 F_1 分数分别为87.37%、86.92%、85.01%、85.95%,相较于人机协同前的各指标分别高出9.09%、10.45%、9.13%、9.78%。这一结果初步验证了人机协同的可行性。

(四)未来研究方向与实践启示

尽管本研究取得了一些有意义的结果,但也存在一些不足。例如:(1)在样本代表性方面。本研究仅在浙江省三个地市的多所中学(初中和高中)取样,可能限制了研究结论的普适性。在后续研究中,可扩大样本采集范围,纳入不同省份、城乡背景、不同家庭结构(如单亲家庭、多子女家庭)的学生数据,以增强样本的代表性与多样性。(2)构建领域专用的LLM。有学者指出,大模型在情感分析任务中可通过对抗训练增强对模糊语义的稳健性,未来可以在通用LLM的基础上,使用高质量的亲子互动、心理学咨询语料进行进一步的精调训练,打造一个更懂家庭教育心理的“领域专家模型”,以减少误判和偏差^[33]。(3)通过实验深入探索人机协同机制,是未来智能判别值得重点思考的方向之一。此外,在实践应用方面,本研究工具可作为核心引擎,集成于在线家庭教育平台、心理咨询APP和家校沟通系统中,为家长提供即时的冲突分析报告和沟通改善建议,实现从测评到干预的闭环。

综上,本研究将亲子冲突划分为语言冲突、行为冲突、冷暴力和软暴力四种类型,搭建了亲子冲突文本测试框架,并证实了文本测试和大语言模型进行自动判别的可行性。LLM在标准化场景中表现稳健,展现出其理解复杂情感语义的巨大潜力,显著优于传统机器学习方法,但在处理反讽和非语言线索等复杂场景时还需人工辅助修正。本研究为亲子关系研究提供了一个崭新的视角,也为人工智能技术在家庭教育领域的深度应用开辟了新的路径。

参考文献:

- [1] 王云峰,冯维. 亲子关系研究的主要进展[J]. 中国特殊教育,2006(7):77-83.
- [2] Bush K R, Peterson G W. Parent-child relationships in diverse contexts[M]// Handbook of Marriage and the Family. Boston, MA: Springer US, 2013: 275-302.
- [3] Laursen B, Coy K C, Collins W A. Reconsidering changes in parent-child conflict across adolescence: a meta-analysis[J]. Child Development, 1998(3): 817-832.
- [4] Weymouth B B, Buehler C, Zhou N, et al. A meta-analysis of parent-adolescent conflict: disagreement, hostility, and youth maladjustment[J]. Journal of Family Theory & Review, 2016(1): 95-112.
- [5] Marshall S K, Tilton-Weaver L, Shankar S, et al. Parent-adolescent conflict processes and their measurement: a systematic review[J]. Journal of Family Theory & Review, 2023(1): 118-132.
- [6] 吴念阳,张东昀. 青少年亲子关系与心理健康的相关研究[J]. 心理科学, 2004(4): 812-816.
- [7] Lindell A K, Killoren S E, Campione-Barr N. Parent-child relationship quality and emotional adjustment among college students: the role of parental financial support[J]. Journal of Social and Personal Relationships, 2021(2): 459-481.
- [8] Davies P T, Pearson J K, Coe J L, et al. Beyond destructive and constructive interparental conflict: children's psychological vulnerability to interparental disorganization[J]. Developmental Psychology, 2021(12): 2192-2205.
- [9] LoBraico E J, Brinberg M, Ram N, et al. Exploring processes in day-to-day parent-adolescent conflict and angry mood: evidence

- for circular causality[J]. *Family Process*,2020(4):1706-1721.
- [10] Grych J H,Seid M,Fincham F D. Assessing marital conflict from the child's perspective:the children's perception of interparental conflict scale[J]. *Child Development*,1992(3):558-572.
- [11] Prinz R J,Foster S,Kent R N,et al. Multivariate assessment of conflict in distressed and nondistressed mother-adolescent dyads [J]. *Journal of Applied Behavior Analysis*,1979(4):691-700.
- [12] Robin A L. Problem-solving communication training;a behavioral-family systems approach to parent-adolescent conflict[J]. *The American Journal of Family Therapy*,1979(2):69-82.
- [13] 方晓义,张锦涛,刘钊. 青少年期亲子冲突的特点[J]. *心理发展与教育*,2003(3):46-52.
- [14] 陈美芬,张敏,陈蕾,等. 长时间相处是亲子关系的保护性因素吗? ——疫情背景下基于微博大数据的质性分析[J]. *心理学*,2022(3):687-694.
- [15] Fondren K,Lawson M,Speidel R,et al. Buffering the effects of childhood trauma within the school setting;a systematic review of trauma-informed and trauma-responsive interventions among trauma-affected youth[J]. *Children and Youth Services Review*,2020,109:104691.
- [16] Tamm A,Tulviste T,Urm A. Resolving conflicts with friends;adolescents' strategies and reasons behind these strategies[J]. *Journal of Adolescence*,2018,64:72-80.
- [17] Lopez-Larrosa S,Sánchez-Souto V,Losada D E,et al. Using machine learning techniques to predict adolescents' involvement in family conflict[J]. *Social Science Computer Review*,2023(5):1581-1607.
- [18] Bird A L,Russell S,Pickard J A,et al. Parents' dispositional mindfulness,child conflict discussion,and childhood internalizing difficulties;a preliminary study[J]. *Mindfulness*,2021(7):1624-1638.
- [19] Wiebe W T,Zhang Y B. Conflict initiating factors and management styles in family and nonfamily intergenerational relationships;young adults' retrospective written accounts[J]. *Journal of Language and Social Psychology*,2017(3):368-379.
- [20] Koss K J,Cummings E M,Davies P T,et al. Patterns of adolescent regulatory responses during family conflict and mental health trajectories[J]. *Journal of Research on Adolescence*,2017(1):229-245.
- [21] Swerbenski H G,Sturge-Apple M L,Koven M,et al. Strengths-based spillover models;constructive interparental conflict,parental supportive problem solving,and development of child executive functioning[J]. *Journal of Family Psychology*,2023(7):1060-1071.
- [22] 丁美琦,许梦雪,朱宗涵,等. 早期亲子互动常用观察法评估量表研究进展[J]. *中国儿童保健杂志*,2018(6):641-644.
- [23] 李春涛,闫续文,张学人. GPT在文本分析中的应用:一个基于Stata的集成命令用法介绍[J]. *数量经济技术经济研究*,2024(5):197-216.
- [24] Sailunaz K,Dhaliwal M,Rokne J,et al. Emotion detection from text and speech;a survey[J]. *Social Network Analysis and Mining*,2018(1):28.
- [25] Renjith S,Abraham A,Jyothi S B,et al. An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms[J]. *Journal of King Saud University-Computer and Information Sciences*,2022(10):9564-9575.
- [26] 冯翔,邱龙辉,郭晓然. 基于LSTM模型的学生反馈文本学业情绪识别方法[J]. *开放教育研究*,2019(2):114-120.
- [27] Cheng J C,Chen A L P. Multimodal time-aware attention networks for depression detection[J]. *Journal of Intelligent Information Systems*,2022(2):319-339.
- [28] 骆方,姜力铭,田雪涛,等. 小学生羞怯特质预测及语言风格模型构建[J]. *心理学报*,2021(2):155-169.
- [29] Zhao W X,Zhou K,Li J,et al. A survey of large language models[EB/OL]. (2025-03-11)[2026-01-22]. <https://doi.org/10.48550/arXiv.2303.18223>.
- [30] Malhotra A,Jindal R. Deep learning techniques for suicide and depression detection from online social media;a scoping review [J]. *Applied Soft Computing*,2022,130:109713.
- [31] Bin Sawad A,Narayan B,Alnefaie A,et al. A systematic review on healthcare artificial intelligent conversational agents for chronic conditions[J]. *Sensors*,2022(7):2625.
- [32] Jaimovitch-López G,Ferri C,Hernández-Orallo J,et al. Can language models automate data wrangling? [J]. *Machine Learning*,2023(6):2053-2082.
- [33] Liu Z,Yang K,Zhang T,Xie Q,et al. EmoLLMs;a series of emotional large language models and annotation tools for comprehensive affective analysis[C]//KDD'24:Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Barcelona Spain. ACM,2024:5487-5496.
- [34] Liga D,Robaldo L. Fine-tuning GPT-3 for legal rule classification[J]. *Computer Law & Security Review*,2023,51:105864.

- [35] Chiu K L, Collins A, Alexander R. Detecting hate speech with GPT-3[EB/OL]. (2022-03-24)[2026-01-22]. <https://doi.org/10.48550/arXiv.2103.12407>.
- [36] Leippold M. Sentiment spin: attacking financial sentiment with GPT-3[J]. *Finance Research Letters*, 2023, 55: 103957.
- [37] Kheiri K, Karimi H. Sentiment GPT: exploiting GPT for advanced sentiment analysis and its departure from current machine learning[EB/OL]. (2023-07-23)[2026-01-22]. <https://doi.org/10.48550/arXiv.2307.10234>.
- [38] 潘可淇, 曾洁霖, 刘意娆, 等. 大数据与人工智能赋能新时代应用型心理测评转型升级[J]. *社会科学前沿*, 2021(7): 1839-1844.
- [39] 方晓义, 张锦涛, 孙莉, 等. 亲子冲突与青少年社会适应的关系[J]. *应用心理学*, 2003(4): 14-21.
- [40] Smetana J G. Adolescents' and parents' reasoning about actual family conflict[J]. *Child Development*, 1989(5): 1052-1067.
- [41] Chung G H, Flook L, Fuligni A J. Daily family conflict and emotional distress among adolescents from Latin American, Asian, and European backgrounds[J]. *Developmental Psychology*, 2009(5): 1406-1415.
- [42] Furman W, Buhrmester D. Age and sex differences in perceptions of networks of personal relationships[J]. *Child Development*, 1992(1): 103-115.
- [43] Kay B. When there is conflict[J]. *Journal of Family Issues*, 2008(6): 780-805.
- [44] 樊召锋, 俞国良. 青少年的亲子冲突及其应对策略[J]. *天津师范大学学报(基础教育版)*, 2008(4): 36-39.
- [45] 吴玉花. 青春期亲子冲突的类型及应对策略[J]. *中小学心理健康教育*, 2018(14): 72-75.
- [46] Hughes J. Toward improved inference for Krippendorff's Alpha agreement coefficient[J]. *Journal of Statistical Planning and Inference*, 2024, 233: 106170.
- [47] Kozina S, Kowalski M, Vlastelica M, et al. Traumatic memory of one's son gone missing in war: content analysis using Krippendorff's Alpha[J]. *SAGE Open*, 2019(1): 1-9.
- [48] Cohen J. A coefficient of agreement for nominal scales[J]. *Educational and Psychological Measurement*, 1960(1): 37-46.
- [49] Landis J R, Koch G G. The measurement of observer agreement for categorical data[J]. *Biometrics*, 1977(1): 159-174.

责任编辑 蒋秋高阿蕊

网 址: <http://xbjbjb.swu.edu.cn>