

基于BP神经网络的ChatGPT使用满意度分析

张艺秦, 吴应蓉, 蒋红云

(绍兴文理学院, 浙江 绍兴 312000)

摘要: 随着人工智能技术的应用越来越广泛,越来越多的人开始担心机器智能是否会取代传统的人类劳动力。然而,存在担忧的同时,人工智能与传统劳动力之间的关系并非完全的互相排斥,可以在一定条件下进行合作。应用BP神经网络对ChatGPT性能满意度进行综合评价,分析民众对ChatGPT性能的满意度,为ChatGPT未来与人类共同发展提供参考和建议。在杭州市11个地区的高新产业园区和高校在读大学生进行满意度问卷调查,探究民众对ChatGPT的生成语言质量、知识准确度、上下文理解能力、模型的响应速度、可控性、可解释性的满意状况。在民众满意度的预测中,BP神经网络在准确率、召回率、调和平均数和F1-score四个指标在模型检验上表现较好。研究表明,ChatGPT的上下文理解能力、可控性、可解释性是ChatGPT性能满意度的重要影响因素,ChatGPT未来应在这三个方面进一步加强训练和优化。

关键词: ChatGPT; 满意度; 神经网络算法

中图分类号: C811; F062.5; G203 **文献标志码:** A **文章编号:** 1671-1807(2025)14-0240-07

ChatGPT可以追溯到2015年,当时深度学习技术的爆发为自然语言处理带来了一次革命。在2018年,OpenAI推出了第一个GPT模型,在多项自然语言处理任务上取得惊人的成绩。随着人工智能(artificial intelligence, AI)技术的进步,ChatGPT在社会中的应用越来越广泛,通过调研分析ChatGPT的使用现状可以帮助研究者和学者深入了解人们对这种技术的接受程度以及满意度。在此背景下,本文旨在解读当前在人工智能不断发展条件下ChatGPT的发展现状。根据人们对ChatGPT的使用情况,为现有ChatGPT研发的改进方向提出相关建议,推动ChatGPT协调可持续且健康的发展。

本文以杭州市民众为实地调查问卷调查对象,围绕民众对ChatGPT的使用情况、对其满意程度两个方面展开调查,分预调研和正式调研两个阶段。对问卷进行数据处理,结果表明问卷具有良好的信效度。利用神经网络算法对民众对ChatGPT的满意度进行分析,研究民众对ChatGPT的看法。

1 相关文献综述

1.1 关于ChatGPT相关信息的文献综述

国外关于chatGPT与人类社会的相关研究情况表明ChatGPT是一种人工智能生成内容技术,

开创了科学界信息来源调查的新时代。Manik和Samriti^[1]通过研究ChatGPT技术推动远程患者监测表示ChatGPT可以创建模仿真实生活场景的模拟患者交互为护士提供沉浸式和个性化的学习体验。Başaran和Duman^[2]通过在医疗领域利用ChatGPT的研究在学术过程和医疗保健应用中显著的表现介绍了医学生在他们的教育旅程中如何从ChatGPT中受益,以及他们遇到的挑战。基于学生使用ChatGPT体验的内容分析表明,它可能在健康教育方面作为科研活动的资源提供优势。然而,也发现了不良报告,包括道德问题、缺乏个人数据保护和科学研究中的潜在滥用。Li等^[3]通过评估ChatGPT-3.5和ChatGPT-4在回答从德国成人软组织肉瘤(STS)循证S3指南中扣除的问题方面的表现,得出ChatGPT-4的整体表现优于ChatGPT-3.5,在准确性(5.5对5.0)和充分性(5.0对4.0)方面的中位数得分更高。但这两种模型偶尔都会提供误导性的信息。

国内一些学者也研究了ChatGPT目前发展状况,刘宝存和苟鸣瀚^[4]表明,ChatGPT能够在工具层面提升研究者的科研效率与质量,在思维层面提供新的教育研究认知路线,在范式层面拓宽教育科学研究叙事尺度。王轶群^[5]表示ChatGPT作为一

收稿日期: 2025-03-10

基金项目: 作者简介: 张艺秦(2003—),女,云南昆明人,研究方向为社会统计、人工智能统计分析;吴应蓉(2003—),女,云南保山人,研究方向为社会统计、人工智能统计分析;蒋红云(1984—),女,湖南邵阳人,博士,讲师,研究方向为经济统计、消费经济、计量经济。

种自然语言处理模型,能够根据用户需要生成拟人化的文本内容,在新媒体领域展示广泛的应用前景。然而,在公众使用的过程中,ChatGPT 也被发现存在虚假信息、偏见和歧视等问题。冷静等^[6]基于批判性思维框架及五维测评量规,利用基于 ChatGPT 的在线平台对 31 名大学生在 186 篇论证作文中反映的批判性思维水平进行评价,表明 ChatGPT 在捕捉和评估具有直接性和明确性的批判性思维要素方面较为高效,但在评估逻辑推理和论证深度方面仍存在局限。未来可进一步优化提示词设计,引导 ChatGPT 深入理解人机对话中体现的批判性思维水平,提升其在复杂认知任务处理中的适应性和准确性。

1.2 关于 ChatGPT 应用领域的文献综述

国外的研究认为 ChatGPT 可以应用于许多领域,如人机交互、机器翻译、信息总结提取等领域。在机器翻译领域,Yousef 等^[7]收集了关于 ChatGPT 相对于 Google Translate 在翻译和翻译教学领域的优势和挑战的信息,通过半结构化访谈和投射技术评估在阿拉伯语环境下 ChatGPT 对翻译的影响,并表明 ChatGPT 作为一种新近出现的基于人工智能的翻译相关技术,其在提供各种优势的同时,也提出了教育者和利益相关者需要应对的新挑战。在自动客服领域 ChatGPT 还可以通过对话的方式自动回答用户的问题。在教育辅助领域,ChatGPT 可以通过对话的方式帮助学生更好地理解知识。Polakova 和 Ivenz^[8]通过研究调查 ChatGPT 反馈在提高 Z 世代学生写作技能方面的有效性,证明 ChatGPT 作为培养 Z 世代学生写作能力的创新教育工具的潜力,强调了将人工智能驱动的技术整合到语言学习过程中以满足当代学习者不断变化的需求的重要性。

国内学者认为 ChatGPT 的应用领域非常广泛,包括智慧医疗、智能助手、教育辅助等各个领域。在智慧医疗领域,吴英华和罗家锋^[9]在基于理解 ChatGPT 技术原理及优缺点的基础上,分别面向医务管理人员、医生医技人员和患者三类用户群体梳理了以 ChatGPT 为例的新一代人工智能技术在智慧医疗领域的典型应用场景。然后针对新一代人工智能技术发展趋势以及对智慧医疗领域的影响进行讨论与思考,在此基础上提出智慧医疗领域拥抱新一代人工智能技术的实施路径。代进杰等^[10]提出 ChatGPT 在核医学教育领域中存在的指导课程知识学习、辅助课程理论教学、营造科研学术氛围、加强临床实

践能力、平衡教育资源不均,以及提高医学人文素养六大优势因素,为高等医学教育的高质量发展带来巨大的机遇和挑战。张美玲等^[11]表明 ChatGPT 介入高校思想政治工作带来的机遇包括赋能学生提升学习效率,为自主学习提供有利条件,助力教师提升教学效率,为教育革新创造有利空间两个方面。但是 ChatGPT 介入高校思想政治工作也会导致弱化学生主体性和教师权威性,存在学生被错误信息误导、教师产生职业倦怠危机的风险。

综上所述可以看出,ChatGPT 在多个领域皆可以运用其中,且能得到大部分的积极效益,但是也存在相应的一些不足需要研究探讨给出相关的建议与解决方案。

1.3 关于 BP 神经网络模型的研究

与本文研究有关的另一类文献是建立 BP 神经网络模型进行分析,果爽等^[12]发现在患者满意度的预测中,BP 神经网络在准确率、召回率和调和平均数 3 个指标均优于传统的机器学习算法;马建威等^[13]以传媒业上市公司为研究对象,利用 Logistic 回归方法和 BP 神经网络设计商誉减值风险评估模型,表明在传媒业上市公司商誉减值风险评估中 BP 神经网络模型优于传统线性模型。赵彬^[14]通过构建 BP 神经网络模型,利用历史数据和实时环境参数进行训练,实现了对光伏集群功率的准确区间预测。研究表明,该方法能有效捕捉光伏集群功率输出的动态特性,为电力系统的稳定运行和优化调度提供了重要支持。

2 研究框架

2.1 调查思路

本次调查从 ChatGPT 爆火全网、AI 发展层出不穷的背景出发,以 ChatGPT 潜在风险理论为基础,对民众了解使用 ChatGPT 的使用目的、满意程度两个方面进行系统的文献综述,构建以民众特征为调节变量的 ChatGPT 满意度模型。

基于文献研究和调查,设计了四大部分 25 个问题的调查问卷。选取问卷中使用目的和满意程度两个问题进行深入研究,对人工智能时代背景下 ChatGPT 了解认知情况开展系统研究。

2.2 数据来源

杭州市普通民众是我们调查民众对 ChatGPT 的使用情况的重要对象,受访者的基本信息,对 ChatGPT 的使用目的、满意程度是本次调查的重要信息来源,因此对杭州市高校在读大学生进行调查十分必要。

以杭州市 11 个地区的高新产业园区和高校在

读大学生为研究对象,从人工智能相关技术从业人员、编制政府工作人员、普通工作人员、高校编程教授、高校在读大学生五个层面,共发放 750 份问卷,回收 724 份问卷,问卷回收率为 96.5%,其中有效问卷为 704 份,问卷的有效率为 93.8%。

2.3 调查问卷设计

问卷分为四个模块,分别为基本信息、对 ChatGPT 的了解程度及使用意愿、对 ChatGPT 的使用情况以及对 ChatGPT 发展前景的预期。本问卷是在人工智能发展背景下,以杭州市民众为调查对象,调查民众对 ChatGPT 的了解程度及使用意愿,对 ChatGPT 的使用目的以及对其使用过程中的满意度,从而探究民众认为 ChatGPT 带来的机遇和挑战,最终探究民众是否愿意接受与 ChatGPT 共同工作的模式,为未来 ChatGPT 可能与人类“抢饭碗”,人类该如何应对提供建议。

其中,ChatGPT 在杭州市民众中使用过程的满意度具体调研框架如表 1 所示。

表 1 ChatGPT 使用情况具体调研框架

调研模块	题目设计	选项设计
对 ChatGPT 的使用情况	使用 ChatGPT 满意度的具体表现	生成语言质量
		知识准确度
		上下文理解能力
		模型的响应速度
		可控性
		可解释性

3 理论模型

3.1 BP 神经网络

BP(back propagation)神经网络是一种经典的人工神经网络算法,它通过反向传播有效地调整隐含层的连接权值。BP 神经网络具有非线性、分布存储、并行处理、自学习、自组织和自适应等特点,能够很好地模拟人脑,并且不受认知因素干扰。

BP 神经网络由输入层、输出层和一层或多层的隐含层组成。每一层都由多个神经元组成,每个神经元只有一个输出连接其他神经元,但它有多个输入通道,每个输入通道都有相应的连接权重系数。同一层的神经元之间没有连接,但相邻层之间的神经元是全连接状态的。

BP 神经网络的传播过程分为正向传播和反向传播。在正向传播中,输入数据经过隐含层处理后到达输出层。如果在传播过程中出现偏差,就会通过反向传播来调整神经元之间的连接权值和阈值,以使误差减小到可接受范围内。

图 1 为一个经典的三层 BP 神经网络示意图,训练一个 BP 神经网络就是通过网络算法不断地调整权值和阈值,使网络的实际输出逐步逼近网络的期望输出。BP 神经网络的运行主要由以下两部分循环运行。

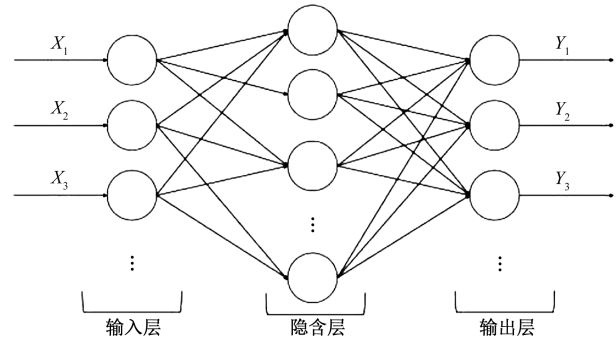


图 1 神经网络示意图

(1)前向传播,数据从输入层进入神经网络,通过隐含层的处理逐层传递至输出层。在每个神经元中,输入和权值的乘积被传递给激活函数,产生神经元的输出。这个过程从输入层一直进行到输出层,形成前向传播。

(2)反向传播,在前向传播之后,通过计算输出层的误差来反向传播误差信号,从输出层向隐含层传递,并最终达到输入层。反向传播的目标是根据误差信号调整网络中各层之间的连接权值和阈值,以减小整体误差。

在训练网络之前,需要随机初始化网络的权值和阈值,之后就开始将数据输入网络,进行信息的前向传输。逆向反馈是从输出层开始,逐步向上一层调整网络的权值和阈值。训练好的 BP 神经网络可以从最后一层的输出值反映数据的类别。

3.2 环境与参数

采用 Python 3.7 进行编程,并在编程中调用了 numpy、pandas、scikit-learn、matplotlib 等第三方辅助软件包。使用的神经网络模型的超参数及对应值如表 2 所示。

表 2 模型超参数及对应值

超参数	对应值
优化器	sgd
学习率	0.1
隐藏层结构	10
正则化参数	10^{-4}
随机状态	42
最大迭代次数	1 000
训练样本与测试样本分割比例	8:2

3.3 性能评价指标

为了评估 BP 神经网络在调查问卷中的预测性能,将数据按比例划分为训练集和测试集,并将民众的满意度得分作为因变量,其他问题得分作为自变量,将分割后的数据集输入模型进行训练。使用准确率(Accuracy)、召回率(Recall)、精确率(Precision)和调和平均数(F_1 -score)4 个指标共同衡量模型的效果,各评价指标的计算方式如式(1)~式(4)所示。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

$$F_1 = \frac{2 \times \text{Accuracy} \times \text{Recall}}{\text{Accuracy} + \text{Recall}} \quad (4)$$

式中:TP 为正例样本被标记为正例的数量;FP 为假例样本被标记为正例的数量;TN 为假例样本被标记为假例的数量;FN 为正例样本被标记为假例的数量。

4 实证结果分析

4.1 调查问卷的检验结果

使用克朗巴哈信度系数检验法,通过 spss 软件计算结果如表 3 所示。

克朗巴哈系数为 0.865,说明问卷数据信度良好。同时采用比较变量间简单相关系数和偏相关系数的检验统计量 KMO(Kaiser-Meyer-Olkin)进行效度检验,通过 spss 软件计算得到结果如表 4 所示。

KMO=0.920,大于 0.7,并且 Bartlett 球形检验对应的 $P=0.000$,因此说明适合进行探索性因子分析。

此案例预期题项可分为 2 个因子,并且 spss 软件生成的两个因子与题项之间的对应关系情况与预期表现出一致性。这 2 个因子与题项的对应关系符合预期,并且对应题项的因子载荷系数值均高

表 3 问卷信度分析结果

克朗巴哈 α	项数
0.865	13

表 4 问卷效度分析结果

KMO 和巴特利特检验		
KMO 取样适切性量数		0.920
巴特利特球形度检验	显著性	0.000

于 0.6,最小为 0.677,最大为 0.746,因而说明题项可以有效地表达因子信息,整体说明效度水平良好。如表 5 所示。

4.2 BP 神经网络在调查问卷中的预测性能结果

为了比较 BP 神经网络与传统机器学习算法的预测性能,分别使用 BP 神经网络、支持向量机、逻辑斯蒂回归、决策树模型在相同的数据上进行实验。实验结果如表 6 所示。

由表 6 可以看出,基于 BP 神经网络的调查问卷预测方法在准确率、召回率和调和平均数 3 个评价指标上均明显优于决策树和逻辑斯蒂回归模型,且相较于支持向量机模型仍有较好的效果,验证了 BP 神经模型的准确率、召回率、调和平均数都较高, F_1 分数也较高说明模型在识别满意度和不满意度的样本方面表现良好。通过模型检验可以看出模型在多元分类问题上表现良好,能够较好地识别大部分实际满意度样本,并且在精确预测满意度方面也有较高的准确性。

表 5 信度分析旋转后的成分矩阵

因子	载荷系数
14-1. 进行文学、媒体等领域的创作,包括小说、文案、报道、广告设计等	0.729
14-2. 修改补充论文,能够得到与论文内容相关的建议和修改意见	0.705
14-3. 编写、调试程序,得到代码片段,算法实现或问题解决方案的建议	0.677
14-4. 快速获取问题的答案,处理简单问题,迅速得到有用的回答	0.677
14-5. 完成程序化的工作,如日程安排、提醒、文档处理、语音文本转换等	0.746
14-6. 语言翻译,进行多种语言的翻译,能查询到广泛的词汇和语法知识	0.691
14-7. 聊天,如进行非正式的对话,提问问题或寻求建议	0.721
16-1. 生成语言质量	0.723
16-2. 知识准确度	0.694
16-3. 上下文理解能力	0.699
16-4. 模型的响应速度	0.699
16-5. 可控性	0.680
16-6. 可解释性	0.698

表 6 BP 神经网络与其他模型比较结果

模型	准确率	召回率	调和平均数
BP 神经网络	0.85	0.96	0.91
支持向量机	0.85	0.83	0.84
决策树	0.85	0.83	0.84
逻辑斯蒂回归	0.85	0.82	0.83

4.3 现有 ChatGPT 的研究结果

根据已有的研究与分析,通过数据挖掘的方法基于 15 个搜索来源,浏览了 12 个网页通过 Python 编码绘制词云图,结果如图 2 所示。

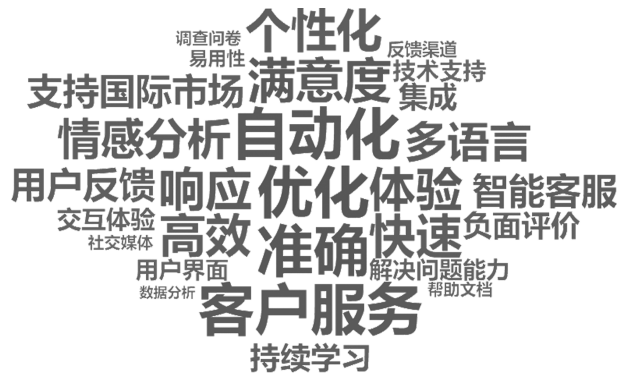


图 2 词云图

根据图 2 可以看出,ChatGPT 高效、自动化性能的满意度较高,表明用户认为 ChatGPT 在提供持续、高效的服务方面表现优异。个性化服务、快速响应、准确这些词也体现民众用户对 ChatGPT 在理解需求并提供定制化、准确回答的能力感到满意。多语言支持、国际市场两个词在图中占比也比较大,表明 ChatGPT 在全球化环境中表现出色,能够服务于不同语言和文化背景的用户,民众对此的满意度也较高。

4.4 民众对 ChatGPT 的满意度

使用 BP 神经网络对数据进行拟合划分,再计算满意度重要性得到评价结果,评价结果如表 7 所示。

由表 7 可知,满意度得分最高的三个选项为可控性、可解释性及上下文理解能力。可控性是指用户能够控制和指导其生成回答的能力,被调查者普遍感到 ChatGPT 在遵循用户指令、适应不同场景和需求以及避免生成不适当内容方面的能力能够较好地符合他们的期望和需求。无论是询问简单事实、寻求建议,还是进行复杂的逻辑推理,ChatGPT 都能够根据用户的指示提供恰当的回答。可解释性是指它能够为用户生成的回答提供合理的解释和依据,被调查者对 ChatGPT 的表现也表示满

表 7 满意度得分

选项	满意度得分
生成语言质量	1.173 6
知识准确度	1.541 1
上下文理解能力	2.518 5
模型的响应速度	2.427 1
可控性	3.397 9
可解释性	3.274 3

意。他们认为 ChatGPT 的输出通常具有足够的透明度,使得用户能够理解其背后的逻辑和依据。即使面对复杂或深奥的问题,ChatGPT 也能够以通俗易懂的方式解释其推理过程,这增强了用户对模型输出的信任感。关于上下文理解能力,被调查者也对 ChatGPT 给予较高的评价。他们认为 ChatGPT 能够较好地捕捉和理解对话的上下文信息,从而做出恰当的回答。无论是在简单的日常对话中,还是在需要维护长时间上下文的复杂场景中,ChatGPT 都能够有效地利用之前的对话内容,使得对话更加连贯和自然。

4.5 ChatGPT 的优缺点

基于使用过 ChatGPT 的民众,对他们进行更深入的调查。在使用过程中,或多或少会发现 ChatGPT 的不足与优点。使用多重响应频率分析的方法,对民众对 ChatGPT 优缺点的看法进行分析,可得每个原因的数量、百分比,将其用图像表示,结果如图 3 所示。

由图 3 可知,民众认为 ChatGPT 的主要优点是它给出的答案更准确,更流畅。ChatGPT 之所以受大众喜欢并且热度不断正是与传统开发的问答式人工智能相比,它所给出的答案更合理,很多技术人员表示它所写出的代码以及回答接近于人经过不断地思考后给出的答案。它不像常用的 AI,靠搜集互联网上的资料拼凑出一份客观的回答,而是集合了所有能搜集到的内容经过“思考”后给出的更加合理流畅的回答。这一优点得到了民众的一致赞同。学术研究也支持这一点,Manik 和 Samriti^[1]通过研究 ChatGPT 技术推动远程患者监测表示,ChatGPT 可以创建模仿真实生活场景的模拟患者交互,为护士提供沉浸式和个性化的学习体验,这表明 ChatGPT 在生成高质量文本方面具有显著优

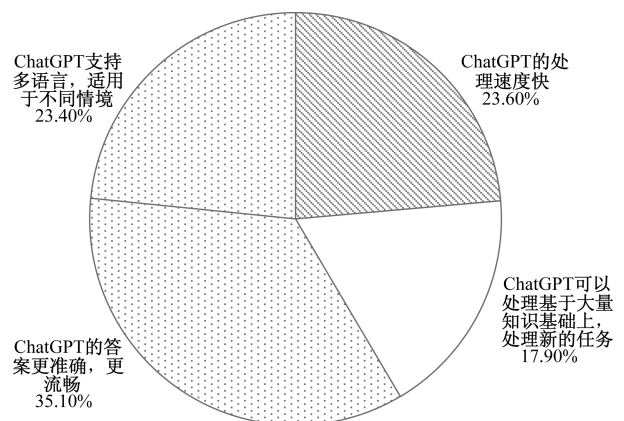


图 3 ChatGPT 优点频率

势。ChatGPT 的处理速度快及支持多语言,适用于不同情境,这一优点在民众反馈中得到了体现,尤其是在全球化环境中表现出色,能够服务于不同语言和文化背景的用户。学术研究也表明,ChatGPT 在多语言环境中的应用前景广阔。例如, Yousef 等^[7]收集了关于 ChatGPT 相对于 Google Translate 在翻译和翻译教学领域的优势和挑战的信息,表明 ChatGPT 作为一种新近出现的基于人工智能的翻译相关技术,其在提供各种优势的同时,也提出了教育者和利益相关者需要应对的新挑战。在短时间内将不同语言的问题思考过滤得到一份完整的回答,这也是以往一些开发的 AI 无法做到的。少部分人认为 ChatGPT 可以处理基于大量知识基础上,处理新的任务,可能在现有的条件下很少有机会尝试让 ChatGPT 处理一些新型任务。但是作为它的优点,相信在未来科技不断发展改进下能让更多的人感受到这一优点。

如图 4 所示,ChatGPT 主要存在的缺点为在特定的专业领域上表现欠佳,尽管 ChatGPT 可以处理自然语言并且回答各种问题,但它仍然有一些局限性,当对于某些非常专业的领域,ChatGPT 可能无法给出详细的专业回答。学术研究也支持这一点,例如, Li 等^[3]通过评估 ChatGPT-3.5 和 ChatGPT-4 在回答从德国成人软组织肉瘤(STS)循证 S3 指南中扣除的问题方面的表现,得出 ChatGPT-4 的整体表现优于 ChatGPT-3.5,但两种模型偶尔都会提供误导性的信息。其次,大众认为 ChatGPT 的缺点还有回答可能出现错误以及 ChatGPT 对新颖的问题会令其受限。ChatGPT 可能在回答问题时产生错误或误导性的信息,此外 ChatGPT 也可能

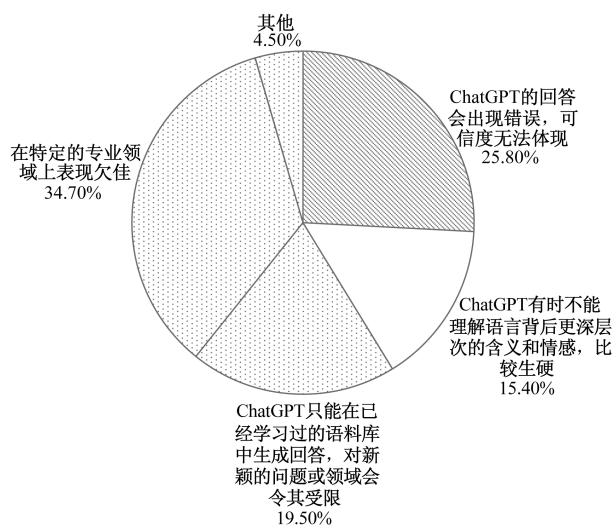


图 4 ChatGPT 缺点频率

受限于其训练数据和算法的范围,导致其在特定领域或任务中的表现不佳,技术存在局限性。学术研究也表明了这一观点,例如,王轶群^[5]表示,ChatGPT 作为一种自然语言处理模型,能够根据用户需要生成拟人化的文本内容,但在公众使用的过程中,ChatGPT 也被发现存在虚假信息、偏见和歧视等问题。ChatGPT 只是人工智能,它通常不会关注用户的情感状态或基于共情来提供建议,因此它给出的回答可能是准确的,也可能是错误的。ChatGPT 还有可能泄露用户隐私,被不法分子利用,实施诈骗等行为。学术研究也支持这一点,例如,刘宝存和苟鸣瀚^[4]表明,ChatGPT 能够在工具层面提升研究者的科研效率与质量,但在思维层面提供新的教育研究认知路线时,也存在一定的局限性,特别是在情感和隐私保护方面。总之,在使用 ChatGPT 时保持理性和警惕力了解其能力和局限性,并结合自身需求进行选择和使用。

5 结论与建议

用户满意度最高的三个选项是可控性、可解释性和上下文理解能力。用户认为 ChatGPT 能够遵循指令、适应不同场景、避免不适当内容,提供合理解释和依据,以及捕捉对话上下文信息,使得对话连贯自然。

ChatGPT 的主要优点包括答案的准确性和流畅性,以及它能够快速处理多语言问题。与传统 AI 相比,ChatGPT 能够提供更合理、流畅的回答,并且能够集合信息经过“思考”后给出回答,这得到了民众的一致赞同。ChatGPT 的缺点主要在于特定专业领域的局限性,可能无法给出详细的专业回答。此外,它可能产生错误或误导性信息,受限于训练数据和算法,且可能不关注用户情感状态,有时可能泄露用户隐私。

被调查者对 ChatGPT 在生成语言质量、知识准确度和模型的响应速度等方面的不满意,实际上为模型的进一步改进和优化提供了明确的方向。

提升生成语言质量方面,可以加强语言模型的训练,通过增加训练数据、优化训练算法,提高 ChatGPT 对自然语言的理解和表达能力。引入更多语言规则和上下文信息,这有助于模型生成更加流畅、自然且符合语境的回答。对于提高知识准确度,定期更新和扩充知识库,确保 ChatGPT 拥有最新、最准确的信息,以回答用户的问题。加强专业领域知识的训练,针对特定领域或行业,提供专门的知识库和训练数据,以提高 ChatGPT 在这些领

域的回答准确度。优化模型的响应速度可以优化模型架构和算法:通过减少模型的复杂度、改进算法效率,提高 ChatGPT 的响应速度。使用更高效的计算资源,如利用高性能计算集群或云计算资源,加速模型的推理过程。

此外,还可以考虑以下几点来综合提升 ChatGPT 的性能:增加用户反馈机制,通过收集用户对 ChatGPT 的回答的评价和建议,不断优化模型的性能;加强多模态交互能力,结合图像、音频等多模态信息,提高 ChatGPT 与用户交互的丰富性和准确性;推广跨领域合作,与其他领域的专家或机构合作,共同研发和优化 ChatGPT 的性能,以满足更广泛的需求。用户在使用 ChatGPT 时需要保持理性和警惕,了解其能力和局限性,并结合自身需求进行选择和使用。

参考文献

- [1] MANIK S, SAMRITI S. A holistic approach to remote patient monitoring, fueled by ChatGPT and Metaverse technology: the future of nursing education[J]. *Nurse Education Today*, 2023, 131: 162-170.
- [2] BAŞARAN M, DUMAN C. Dialogues with artificial intelligence: exploring medical students' perspectives on ChatGPT[J]. *Medical Teacher*, 2024, 13: 11-10.
- [3] LI P C, JAKOB J, MENGE F, et al. Comparing ChatGPT-3.5 and ChatGPT-4's alignments with the German evidence-based S3 guideline for adult soft tissue sarcoma[J]. *Science*, 2024, 27(12): 111493.
- [4] 刘宝存, 苟鸣瀚. ChatGPT 等新一代人工智能工具对教育科研的影响及对策. *苏州大学(教育科学版)*, 2023, 11(3): 54-62.
- [5] 王轶群. 基于 ChatGPT 的误导性舆情研究与治理[J]. *湖南警察学院学报*, 2024, 36(4): 102-110.
- [6] 冷静, 卢弘煊, 代琳. 生成式人工智能赋能批判性思维测评——基于 ChatGPT 的应用实验[J]. *现代远程教育研究*, 2024, 36(6): 102-111.
- [7] YOUSEF S, TALIB M A A, MOHAMMED K J A. A cross sectional study of ChatGPT in translation: magnitude of use, attitudes and uncertainties[J]. *Journal of Psycholinguistic Research*, 2023, 52(6): 2937-2954.
- [8] POLAKOVA P, IVENZ P. The impact of ChatGPT feedback on the development of EFL students' writing skills[J]. *Cogent Education*, 2024, 11(1): 162-182.
- [9] 吴英华, 罗家锋, 林成创. 迈入人工智能新时代: ChatGPT 在智慧医疗应用场景研究与思考[J]. *数据通信*, 2023(4): 33-38.
- [10] 代进杰, 陈艺德, 邓艾琳, 等. 基于扎根理论的 ChatGPT 在核医学教育中的应用优势[J]. *医学教育研究与实践*, 2025, 33(1): 136-143.
- [11] 张美玲, 程魁, 韩雪. ChatGPT 介入高校思想政治工作的机遇、挑战与风险防范[J]. *黑龙江教育(高教研究与评估)*, 2024(12): 13-15.
- [12] 果爽, 程伟, 刘鑫鑫, 等. 基于 BP 神经网络的医疗卫生服务质量满意度研究[J]. *中国医学物理学杂志*, 2023, 40(6): 788-792.
- [13] 马建威, 王晶, 陆珊. 传媒业上市公司商誉减值风险评估模型构建——基于 Logistic 回归和 BP 神经网络的比较[J]. *财会研究*, 2024(11): 35-42.
- [14] 赵彬. 基于 BP 神经网络的分布式光伏集群功率超短期区间预测方法[J]. *自动化应用*, 2024, 65(21): 84-86.
- [15] MIAO J, ZHU W. Precision-recall curve(PRC)classification trees[J]. *Evol Intell*, 2022, 15(3): 1545-1569.
- [16] 马倩倩. 基于 BP 神经网络的农村电商物流服务客户满意度评价[J]. *物流科技*, 2024, 47(7): 66-69.
- [17] AYTAÇ A, ELIF G, EREN T, et al. Integrating artificial intelligence in orthognathic surgery: a case study of ChatGPT's role in enhancing physician-patient consultations for dentofacial deformities[J]. *Journal of Plastic, Reconstructive Aesthetic Surgery*, 2023, 87: 405-407.

Analysis of User Satisfaction of ChatGPT Based on BP Neural Network

ZHANG Yiqin, WU Yingrong, JIANG Hongyun

(Shaoxing University, Shaoxing 312000, Zhejiang, China)

Abstract: With the increasingly widespread application of artificial intelligence (AI) technologies, more and more people are concerned about whether machine intelligence will replace traditional human labor. However, while there are concerns, the relationship between AI and traditional labor is not entirely mutually exclusive. They can cooperate under certain conditions. The BP neural network was applied to conduct a comprehensive evaluation of ChatGPT performance satisfaction, analyze the public's satisfaction with ChatGPT's performance, and provide references and suggestions for the future co-development of ChatGPT with humans. A satisfaction survey was conducted among 11 high-tech industrial parks and university students in Hangzhou, exploring the public's satisfaction with ChatGPT's generated language quality, knowledge accuracy, context understanding ability, model response speed, controllability and interpretability. In the prediction of public satisfaction, the BP neural network performs well in four indicators, such as accuracy, recall, harmonic mean and F1-score during model testing. The result shows that ChatGPT's context understanding ability, controllability and interpretability are important influencing factors of ChatGPT performance satisfaction. In the future, ChatGPT should further strengthen training and optimization in these three aspects.

Keywords: ChatGPT; satisfaction; neural network algorithm