

基于知乎爬虫的话题研究

——以“焦虑、抑郁、迷茫”为例

杜东山*, 乌伟, 徐前进, 武聪

西京学院计算机学院, 陕西 西安

收稿日期: 2025年5月11日; 录用日期: 2025年6月9日; 发布日期: 2025年6月17日

摘要

在信息时代, 知乎作为重要的知识问答社区, 积累了海量数据, 为多学科研究提供丰富资源。本研究通过构建一个高效稳定的知乎数据爬虫系统, 从知乎平台针对“抑郁、焦虑、迷茫”三个话题抽取了701条问题数据, 去除了专栏文章和电子书后, 保留了575个有效问题和220,184条有效回答, 再对每个回答的内容进行解析, 提取问题信息, 包括问题ID、话题、问题描述和回答数等信息, 在此基础上采取关键词分析、情感分析、影响力分析、性别差异分析方法并对结果进行了讨论。本研究的结果不仅对理解网络媒体上心理健康话题的讨论和影响具有重要意义, 还为相关领域的研究提供了实证数据和方法论支持。

关键词

知乎爬虫, 心理健康, 数据分析

Research on Topic Analysis Based on Zhihu Web Crawler

—A Case Study of “Anxiety, Depression, and Confusion”

Dongshan Du*, Wei Wu, Qianjin Xu, Cong Wu

School of Computer Science, Xijing University, Xi'an Shaanxi

Received: May 11th, 2025; accepted: Jun. 9th, 2025; published: Jun. 17th, 2025

Abstract

In the information age, as an important knowledge Q&A community, Zhihu has accumulated massive data, providing rich resources for multidisciplinary research. In this study, an efficient and stable

*通讯作者。

Zhihu data crawler system was constructed to extract 701 question data from the Zhihu platform for the three topics of “depression, anxiety, and confusion”. After removing column articles and e-books, 575 valid questions and 220,184 valid answers were retained. Then, the content of each answer was analyzed to extract question information, including question ID, topic, question description, number of answers, etc. On this basis, keyword analysis, emotional analysis, influence analysis, and gender difference analysis methods were adopted, and the results were discussed. The findings of this study not only hold significant implications for understanding the discussions and impacts of mental health topics on online media but also provide empirical data and methodological support for research in related fields.

Keywords

Zhihu Crawler, Mental Health, Data Analysis

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 绪论

1.1. 研究背景

近年来, 抑郁、焦虑、迷茫等心理问题在青年群体中日益凸显, 而网络平台上的讨论既反映了公众的心理状态, 也影响着社会对心理健康议题的认知。然而, 目前针对知乎等平台心理健康话题的系统性研究仍然有限, 尤其缺乏基于大规模数据的情感、性别差异及影响力分析。传统心理健康研究多依赖问卷调查或临床数据, 而社交媒体数据能够提供更自然、实时的公众心理状态表达。因此, 本研究通过构建高效的数据采集与分析框架, 对知乎平台上的心理健康相关讨论进行挖掘, 以揭示网络语境下公众对抑郁、焦虑、迷茫等问题的认知模式、情感倾向及讨论特征。

1.2. 研究意义

1.2.1. 理论价意义

丰富了心理健康研究的视角, 将传统心理学研究与计算社会科学方法结合, 探索社交媒体数据在心理健康研究中的应用价值。通过关键词分析、情感分析等方法, 揭示公众在讨论心理健康问题时的语言特征和情感倾向, 为网络心理语言学提供实证依据。考察性别差异在心理健康话题讨论中的表现, 有助于理解不同群体在心理困扰表达上的差异。

1.2.2. 实践意义

为心理健康工作者、社会支持机构和政策制定者提供数据支持, 帮助其更精准地把握公众心理需求, 优化心理健康干预策略。本研究构建的数据采集与分析框架可拓展至其他社交媒体平台, 为后续相关研究提供方法论借鉴。

1.3. 国内外研究现状

1.3.1. 国外研究现状

在社交媒体数据分析方面, 针对 Facebook、Twitter 等平台的研究较为深入[1], 凭借文本挖掘以及机器学习等方式, 研究者得以对用户所发布的各类内容进行情感层面的解读和分析[2]、话题建模以及社交网络分析[3], 深入探究并清晰阐释社交网络里信息传递的内在规律以及用户的行为模式。刘洋等[4]通过

分析 Twitter 上的话题标签和用户转发行为, 研究突发事件的传播路径和影响范围。

1.3.2. 国内研究现状

1) 用户行为分析: 国内学者通过爬虫获取知乎用户的提问、回答、点赞等数据, 分析用户的活跃度、社交关系以及知识贡献模式。张静[5]提出知乎成功得益于高质量内容服务, 以及 UGC 模式。

2) 知识挖掘与可视化: 有研究利用知识图谱技术构建知乎话题知识体系, 实现知识的可视化展示与关联分析, 为用户提供更高效的知识检索与学习工具[6]。

3) 用户影响力分析与关键意见领袖挖掘: Aiting 等[7]依据知乎用户的社交网络结构以及问答网络模式, 搭建起用户影响力挖掘的模型。郭博等[8]创新地提出改良的 PageRank 算法与 HITS 算法。

2. 实验过程

2.1. 话题选择

选择“抑郁、焦虑、迷茫”三个话题作为测试对象。这些话题在社交媒体上具有较高的关注度和讨论度, 涵盖了人们在心理和情绪方面的常见困扰, 部分结果见表 1。

2.2. 数据采集

Table 1. Excerpts of some questions under the topics of depression, anxiety, and confusion

表 1. 抑郁、焦虑与迷茫话题下部分问题摘录

序号	问题 ID	话题	问题	回答数
1	389945849	焦虑	你有哪些可以告别焦虑的方法?	31,005
2	367150784	抑郁	你抑郁的时候, 是谁将你拉出了深渊?	11,356
3	354466873	抑郁	你愿意和重度抑郁的人做朋友吗?	11,267
4	20667527	焦虑	想学的东西太多, 而自己能力有限, 为此非常焦虑, 该怎么办?	8743
5	511311505	焦虑	在明白什么道理后, 你不再焦虑了?	7146
6	627701056	迷茫	31 岁了, 多看书还来得及吗?	5861
7	425461256	抑郁	你愿意做高学历却抑郁的人, 还是低学历开心的人?	5336
8	597135671	抑郁	中度抑郁, 32 岁了, 没车没房没工作没孩子, 我还有什么理由活下去?	4839
9	486206053	抑郁	为什么现在那么多小孩得抑郁?	4421
10	434712309	焦虑	87 年大龄剩女感觉嫁不出去, 焦虑到失眠怎么办?	3264
11	491535877	抑郁	我国青少年抑郁检出率达到 24.6%, 为什么青少年抑郁的这么多? 受到哪些因素影响?	3159
12	400632446	抑郁	“忙起来你就没时间抑郁了”这句话有理由么?	3061
13	582549293	焦虑	为什么现在的大学生都很焦虑?	2071
14	301034734	抑郁	为什么近几年自称抑郁者越来越多?	1951
15	620798036	迷茫	人最迷茫的时候应该作什么?	1770
16	608338296	抑郁	重度抑郁了, 学心理学能自救吗?	477

在获取到话题下的问题后, 进一步获取每个问题下的所有有效回答, 部分结果见表 2。通过模拟浏览器操作, 访问每个问题的页面, 滚动页面加载更多回答, 并对回答进行解析和存储。

进一步提取了 220,184 条回答, 去除纯标点、纯字母、纯链接后, 共获得 206,175 条有效回答。

Table 2. Example table of content after data cleaning
表 2. 数据清洗后内容示例表

序号	话题	问题	回答内容	回答分词统计	回答情感分析	点赞	评论	喜欢	时间	性别
321	抑郁	你抑郁的时候,是谁将你拉出了深渊?	重新看各种小时候看过的电视剧家有儿女炊事班的故事武林外传啊四驱兄弟啊啥的听听助眠	{“名词”:{“电视剧”:1,“家”:1,“儿女”:1,“炊事班”:1,“故事”:1,“兄弟”:1},“时间词”:{“小时候”:1},“处所词”:{},“方位词”:{},“动词”:{“看”:4,“有”:1,“外传”:1,“放”:1,“听听”:1,“助眠”:1},“形容词”:{},“区别词”:{},“状态词”:{},“代词”:{“各种”:1,“啥”:1},“数词”:{“四”:1},“量词”:{},“副词”:{“重新”:1,“有时候”:1,“就”:1,“不”:1},“介词”:{},“连词”:{},“助词”:{“过”:1,“的”:3,“啊”:2,“着”:1},“叹词”:{},“语气词”:{},“拟声词”:{“嘎嘎”:1},“简称”:{},“前缀”:{},“后缀”:{“驱”:1},“字符串”:{},“标点符号”:{“,”:1}}	0.9998	40	0	13	2023/9/3	-1

3. 实验结果

3.1. 关键词与情感分析

关键词发现:列出了抑郁、焦虑和迷茫各自话题下出现频率最高的关键词,分别见表 3~6,高频词词云图分别如图 1~3 所示。

Table 3. Summary of Top 15 high-frequency character keywords
表 3. 词频前 15 的人物关键词汇总

迷茫		焦虑		抑郁	
朋友	19%	朋友	20.88%	朋友	22.04%
父母	12.74%	孩子	10.12%	孩子	11.1%
老师	10.5%	父母	9.3%	父母	10.66%
孩子	8.6%	老师	8.08%	医生	9.62%
同学	8.43%	妈妈	6.32%	患者	9.61%
学生	7.43%	同学	6.06%	妈妈	9.07%
妈妈	5.84%	同事	5.86%	老师	5.39%
爸爸	5.39%	医生	5.29%	爸爸	5.02%
普通人	4.48%	对象	5.01%	同学	4.72%
对象	3.47%	学生	4.79%	正常人	2.79%
老板	3.3%	普通人	4.57%	学生	2.6%
同事	3.17%	领导	4.22%	男朋友	1.94%
客户	2.77%	爸爸	3.75%	普通人	1.93%
领导	2.56%	老板	3.44%	咨询师	1.84%
员工	2.32%	咨询师	2.32%	同事	1.67%

Table 4. Summary of Top 15 high-frequency time keywords**表 4.** 词频前 15 的时间关键词汇总

迷茫		焦虑		抑郁	
现在	39.25%	现在	29.55%	现在	33.61%
目前	8.76%	当下	8.82%	当时	10.53%
今天	6.89%	今天	8.47%	晚上	9.36%
过去	5.93%	过去	7.86%	过去	6.53%
今年	5.79%	目前	6.85%	今天	6.26%
当时	5.58%	晚上	5.99%	最近	5.57%
当下	4.88%	最近	5.90%	目前	4.84%
最近	4.40%	当时	5.26%	平时	4.07%
晚上	3.37%	明天	4.73%	早上	3.70%
明天	3.03%	平时	3.60%	小时候	3.02%
平时	2.99%	今年	3.27%	白天	2.92%
当前	2.62%	眼前	2.64%	今年	2.74%
当年	2.19%	当前	2.50%	明天	2.51%
去年	2.18%	早上	2.44%	下午	2.42%
眼前	2.14%	昨天	2.13%	童年	1.93%

Table 5. Summary of Top 15 high-frequency verb keywords**表 5.** 词频前 15 的动词关键词汇总

迷茫		焦虑		抑郁	
学习	11.19%	生活	11.40%	生活	13.25%
工作	10.98%	工作	9.72%	开始	10.09%
生活	10.62%	努力	9.55%	工作	8.59%
开始	8.44%	学习	8.38%	希望	8.19%
选择	8.37%	开始	7.70%	理解	6.89%
努力	7.21%	解决	7.68%	学习	6.47%
读书	6.56%	选择	5.64%	努力	6.43%
思考	5.69%	找到	5.46%	帮助	6.13%
帮助	4.85%	改变	5.41%	治疗	5.58%
解决	4.70%	面对	5.16%	告诉	5.30%
提升	4.58%	运动	5.12%	感受	5.24%
发展	4.49%	缓解	4.98%	自杀	5.17%
毕业	4.37%	帮助	4.80%	接受	4.27%
坚持	4.01%	希望	4.55%	解决	4.22%
改变	3.92%	接受	4.46%	改变	4.18%

Table 6. Summary of Top 12 high-frequency emotion keywords
表 6. 词频前 12 的情绪关键词汇总

迷茫		焦虑		抑郁	
焦虑	20.75%	恐惧	11.54%	开心	19.28%
快乐	9.80%	快乐	11.51%	痛苦	15.26%
痛苦	9.32%	痛苦	11.17%	焦虑	12.53%
稳定	8.41%	开心	10.67%	快乐	11.77%
美好	7.98%	紧张	9.41%	难受	5.53%
轻松	7.76%	轻松	7.46%	美好	5.52%
困惑	7.55%	幸福	7.42%	敏感	5.32%
幸福	7.52%	烦恼	7.08%	难过	5.27%
开心	7.13%	迷茫	6.51%	绝望	5.13%
孤独	5.22%	平静	5.78%	可怕	4.94%
烦恼	4.74%	稳定	5.74%	幸福	4.89%
平静	3.81%	抑郁	5.70%	低落	4.57%



Figure 1. Word cloud of high-frequency confusion terms
图 1. 迷茫高频词词云图

3.2. 情感倾向分析

分析帖子的情感倾向，如负面、中性或正面，以及这些情感如何随时间变化，见图 4。

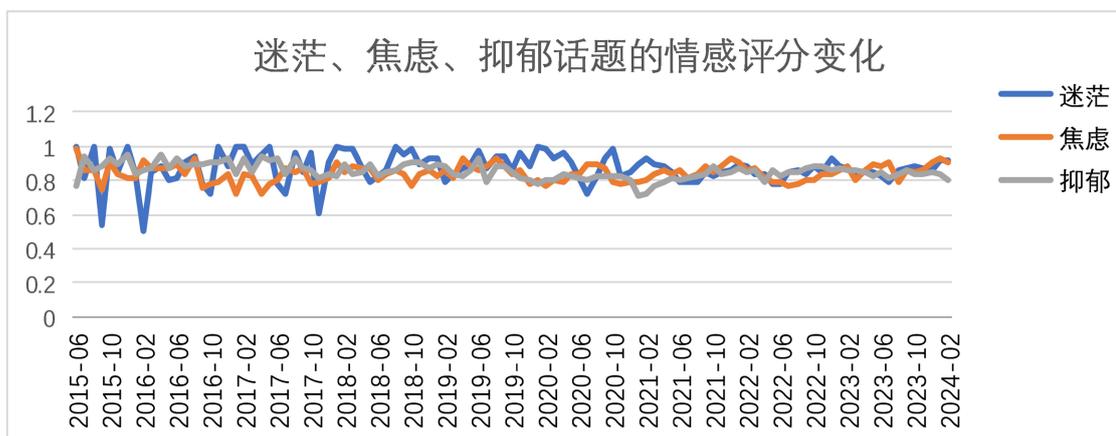


Figure 4. Sentiment score trends for confusion, anxiety, and depression topics

图 4. 迷茫、焦虑、抑郁话题的情感评分

Table 7. Emotional score statistics for confusion, anxiety, and depression topics

表 7. 迷茫、焦虑、抑郁话题的情感评分统计数据

话题	中值	平均值	标准差
迷茫	0.87175	0.873375	0.089042
焦虑	0.842592	0.841528	0.046866
抑郁	0.850869	0.85474	0.045428

Table 8. Multiple comparisons of confusion, anxiety, and depression topics

表 8. 迷茫、焦虑、抑郁话题的多重比较

方差齐性	(I)话题		平均差异(I-J)	显著性	95%信賴區間	
					下限	上限
LSD	迷茫	焦虑	0.0316798964*	0.000431	0.014164	0.049195
		抑郁	0.0206146751*	0.021221	0.003099	0.03813
	焦虑	迷茫	-0.0316798964*	0.000431	-0.0492	-0.01416
		抑郁	-0.01107	0.214799	-0.02858	0.00645
	抑郁	迷茫	-0.0206146751*	0.021221	-0.03813	-0.0031
		焦虑	0.011065	0.214799	-0.00645	0.028581
Tamhane	迷茫	焦虑	0.0316798964*	0.005168	0.007698	0.055662
		抑郁	0.020615	0.107717	-0.00307	0.044301
	焦虑	迷茫	-0.0316798964*	0.005168	-0.05566	-0.0077
		抑郁	-0.01107	0.250701	-0.02679	0.004656
	抑郁	迷茫	-0.02061	0.107717	-0.0443	0.003072
		焦虑	0.011065	0.250701	-0.00466	0.026787

表 7 和表 8 表明不同话题之间存在显著差异，尤其迷茫话题的情感评分波动性明显较大，中值、平均值略高于其他两个话题，可能暗示了在迷茫话题下的回答情感分析中存在一些更积极、更乐观的情感倾向，标准差也略高于其他两个话题，意味着迷茫话题下的回答情感更加分散，即存在更大范围的情感表达。

3.2.1. 影响力分析

表 9 描述了各个话题的社交影响力，包括点赞数、评论数、喜欢数和回答总数，以及随时间的变化趋势，分别如图 5~8 所示。

Table 9. Annual summary of average likes, comments, favorites, and total answers

表 9. 平均点赞数、评论数、喜欢数和回答总数年度汇总

项目	话题	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
平均点赞数	迷茫	18	13.33	58.04	94.96	13.41	42.23	17.77	34.75	17.95	12.16	4.64
	焦虑	134.31	64.78	59.53	67.78	62.76	13.84	26.52	22.29	9.67	5.62	8.3
	抑郁	32.43	51.12	72.01	31.31	32.57	46.17	27.03	23.15	23.34	46.59	18.07
平均评论数	迷茫	9.33	7.33	9.14	7.55	7.12	3.73	2.14	3.89	6.54	2.59	0.51
	焦虑	8.05	4.99	4.54	4.35	4.98	1.63	5.91	4.28	2.17	0.61	1.05
	抑郁	11.43	11.7	13.38	5.95	6.51	8.51	7.06	2.74	2.51	4.66	2.02
平均喜欢数	迷茫	6.33	12.33	29.22	43.37	5.4	16.5	10.73	4.91	5.71	3.38	3.32
	焦虑	60.81	33.5	24.96	28.39	19.72	3.49	9.44	10.04	3.45	2.55	8.37
	抑郁	13	19.87	38.48	10.07	7.72	13.07	7.21	5.31	4.87	10.41	6.75
回答数	迷茫	3	3	120	141	129	315	364	1320	2378	6667	19,827
	焦虑	42	98	1037	578	2413	2779	2544	4964	5881	38,929	24,038
	抑郁	7	67	332	419	616	1574	5273	27,642	14,897	12,931	16,673

平均点赞数

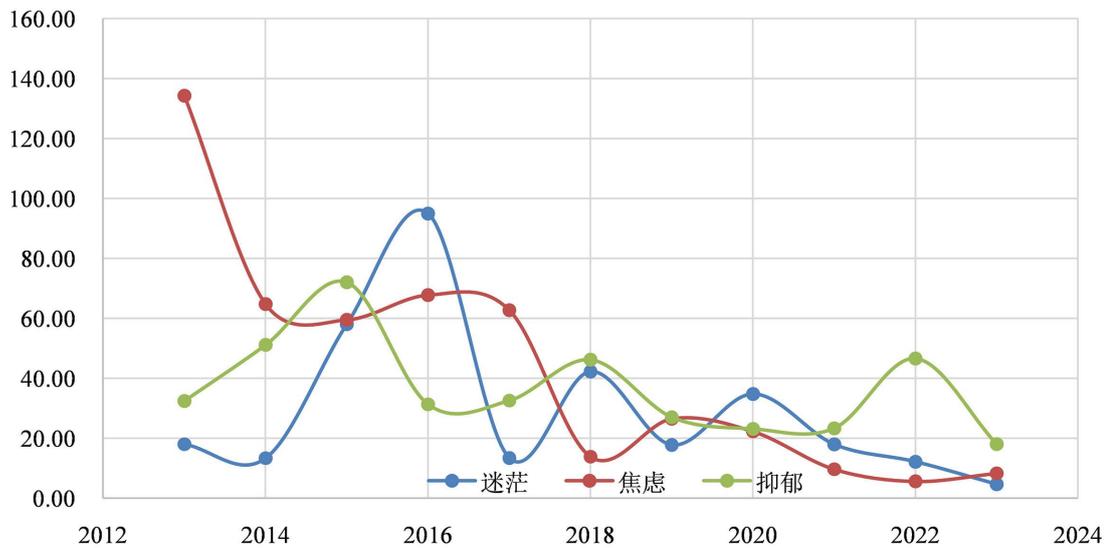


Figure 5. Annual average likes for confusion, anxiety, and depression topics

图 5. 迷茫、焦虑、抑郁话题的年度平均点赞数

平均评价数

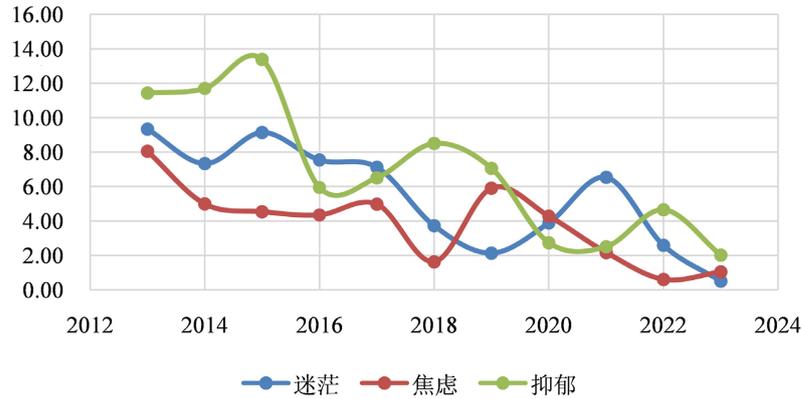


Figure 6. Annual average comments for confusion, anxiety, and depression topics
图 6. 迷茫、焦虑、抑郁话题的年度平均评论数

平均喜欢数

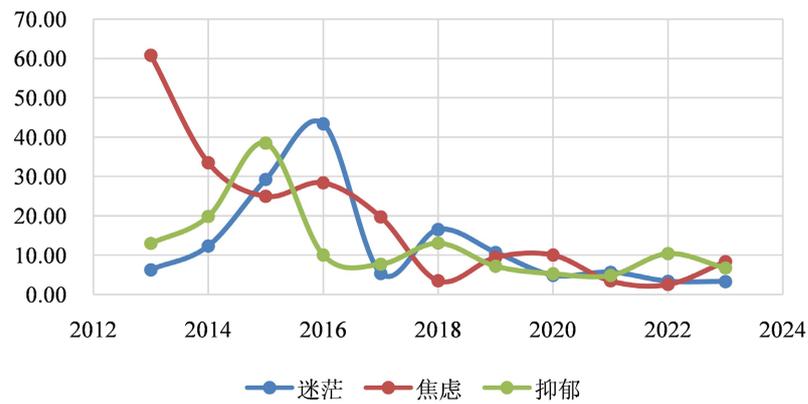


Figure 7. Annual average favorites for confusion, anxiety, and depression topics
图 7. 迷茫、焦虑、抑郁话题的年度平均喜欢数

回答数

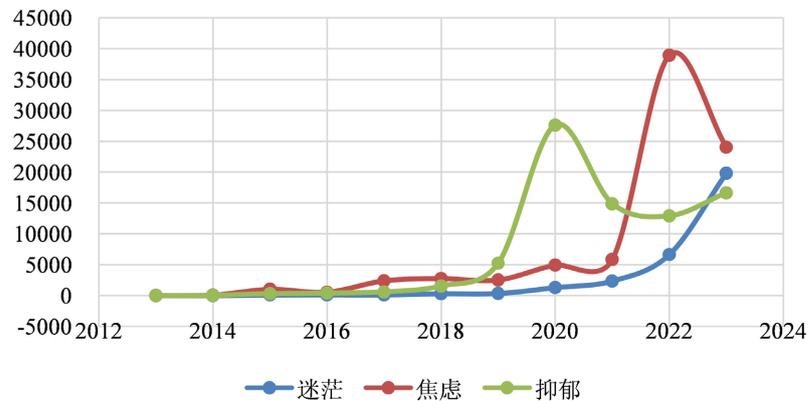


Figure 8. Annual answer counts for confusion, anxiety, and depression topics
图 8. 迷茫、焦虑、抑郁话题的年度回答数

3.2.2. 性别差异

利用 SnowNLP 对回答内容进行情感评分, 分数在 0 和 1 之间, 越接近 1, 情感越正面, 反之越负面, 表 10 展示了不同性别在抑郁、焦虑和迷茫话题下的情感倾向。

Table 10. Excerpt of average emotional scores of comments by gender

表 10. 不同性别平均评论情感分数汇总表节选

话题	性别	2023-01	2023-02	2023-03	2023-04	2023-05	2023-06	2023-07
迷茫	男	0.8937	0.9026	0.8668	0.8672	0.8568	0.8335	0.7962
	女	0.9004	0.8613	0.857	0.8578	0.8787	0.8443	0.8087
焦虑	男	0.8495	0.8922	0.8147	0.8515	0.9097	0.8829	0.9024
	女	0.8953	0.8886	0.8169	0.8851	0.9061	0.9004	0.9038
抑郁	男	0.8794	0.844	0.8489	0.8672	0.8537	0.8396	0.8353
	女	0.8828	0.8759	0.8773	0.848	0.8262	0.8779	0.8065

Table 11. Statistical analysis of average emotional comment scores by gender

表 11. 不同性别平均评论情感分数统计分析

主题	性别	N	平均数	标准偏差	采用相等变異數	F	显著性	T	df
迷茫	男	72	0.88	0.07	是	0.91	0.34	-0.11	142.0
	女	72	0.88	0.11	否			-0.11	123.80
焦虑	男	72	0.84	0.05	是	1.06	0.31	-2.88	142.00
	女	72	0.87	0.05	否			-2.88	140.45
抑郁	男	72	0.85	0.03	是	12.33	0.00	0.38	142.00
	女	72	0.85	0.05	否			0.38	116.27

表 11 显示男女之间在迷茫主题下的评论情感分数没有显著差异。在焦虑主题下的评论情感分数存在显著差异。在抑郁主题下的评论情感分数存在显著差异。

4. 结果讨论

4.1. 关键词分析

在人物关键词汇总表中, 朋友、父母、老师、孩子等关系人物的词频较高, 这反映了个体的社交需求和依赖[9], 以及与这些关系人物的互动对个体心理健康的影响[10]。动词反映了个体在应对迷茫、焦虑和抑郁时采取的行为[11], 例如, 学习、工作、选择等词语表明个体可能在积极寻求解决问题的方法, 而思考、帮助、解决等词语则反映了社会支持和心理咨询的重要性[12]。三个主题下, 都提到了“焦虑”, 说明人们对自己当前的状态具有自知力, 并迫切想要改变。在情绪体验中, 两极分化、两级反转较大较快, 说明了在相应心境下人们的情绪波动大、起伏快[13]。

在“迷茫”主题下, 人们更多地提及了同学、学生、领导、老板、同事, 说明迷茫更多地与学业和职场相关。在“焦虑”和“抑郁”主题下, 除了心理状态外, 人们还提及了更多的社交关系, 如家人、朋友和老师[14]。这表明人们在经历焦虑和抑郁时, 可能会寻求社交支持或受到社交关系的影响[15]。

不同主题下的关键词排名和出现频率也有所不同, 这可能反映了人们在不同心理状态下所关注的焦

点和需求的差异[16]。

4.2. 情感分析

情感分析的意义在于帮助我们理解用户在社交媒体上表达的情绪和态度[17]。

从平均情感评分来看，迷茫主题的评分最高，评论可能更积极，通过 LSD 和 Tamhane 的多重比较，我们可以看到迷茫主题与焦虑主题以及迷茫主题与抑郁主题之间存在着显著差异，而焦虑主题与抑郁主题之间的差异相对较小。这意味着用户对于迷茫主题的评论可能更积极。从整体趋势来看，人们对这些主题的评论普遍具有积极的倾向，其中迷茫主题可能会更受欢迎一些。

4.3. 社交影响力分析

通过数据，迷茫在 2016 年出现一个高峰，焦虑在 2015 年~2017 年都处于高峰，抑郁在 2015 年和 2022 年出现高峰。这些变化，明显带有时间性，与特定年份密切相关，包括经济“新常态”、人口政策调整、新冠疫情、俄乌冲突等，这些事件和变化可能影响到人们的心理状态[18]，导致一部分人感到迷茫、焦虑和抑郁。

4.4. 性别分析

在焦虑和抑郁主题下，男女之间的评论情感分数存在显著差异，而在迷茫主题下，两者之间没有显著差异。这可能表明在这个特定的情境下，男女对于焦虑和抑郁主题的评论情感有明显的差异，可能涉及到性别差异、社会化、生物学因素等方面的考虑[19]。

5. 研究限制和未来研究方向

5.1. 关键词分析

样本代表性：知乎用户群体的局限性可能导致样本不能充分代表整个社会，使得研究结果在推广到其他人群时存在一定的偏差。

信息完整性：未考虑非文本形式的答案，导致对问题和话题的理解不够全面，可能忽略了一些重要的信息。

工具和方法：现有的汉语分词工具和情感分析工具的不足，影响了对文本数据的处理和分析，降低了研究结果的可信度。

5.2. 对未来研究的建议

未来的研究需要从以下方面不断改进和完善，以提高对知乎平台数据的利用效率和研究的深度广度，为相关领域的发展提供更有力的支持。

扩大样本范围：未来的研究可以尝试结合多个平台的数据，或者采用抽样的方法，扩大样本的多样性，包括不同年龄、性别、文化和社会经济背景的用户，以提高研究结果的代表性。

深度数据分析：采用更先进的数据分析方法，如深度学习算法，对文本、图片、视频等多种形式的数据进行深度分析，挖掘更多有价值的信息。

技术改进：致力于改进数据处理和分析工具，提高汉语分词工具、情感分析工具等的准确性和适用性，并进行全面的验证。

长期追踪研究：进行长期的追踪研究，了解知乎平台上话题的演化规律和用户行为的变化趋势，为平台的发展和研究提供更有前瞻性的建议。

参考文献

- [1] 王强. 社交媒体数据分析关键技术和应用[J]. 数据科学, 2023, 15(6): 201-210.
- [2] 李明. 基于 Python 的社交网络情感分析系统研究与实现[D]: [硕士学位论文]. 北京: 北京邮电大学, 2023.
- [3] 张伟. 基于深度学习的情感分析在社交媒体中的应用探究与实践[J]. 计算机应用, 2025, 42(1): 102-108.
- [4] 刘洋. 网络社交媒体数据挖掘与情感分析[J]. 网络安全技术, 2024, 11(10): 120-126.
- [5] 张静. “知乎”用户知识分享行为的“使用与满足”[J]. 青年记者, 2018(8): 94-95.
- [6] 卢晓荣, 张树良. 国内社交媒体用于学术成果 Altmetrics 评价存在问题及对策[J]. 图书情报工作, 2019, 63(21): 87-96.
- [7] Xu, A.T., Wang, F.Y. and Ying, P.T. (2019) Xiaomi Brand Appraisal Research Based on Zhihu by Text Mining Technology. *Proceedings of the 4th International Conference on Big Data and Computing*, Guangzhou, 10-12 May 2019, 221-225.
- [8] 郭博, 许昊迪, 雷水旺. 知乎平台用户影响力分析与关键意见领袖挖掘[J]. 图书情报工作, 2018, 62(20): 122-132.
- [9] Sun, Y.Q., Guo, B., Li, Z.M., Cheng, J.H. and Wang, L. (2019) Leveraging User Profiling in Click-through Rate Prediction Based on Zhihu Data. *2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI)*, Xi'an, 21-22 September 2019, 131-136.
- [10] 雷蒙·威廉斯. 关键词: 文化与社会的词汇[M]. 刘建基, 译. 北京: 三联书店. 2005: 46-48.
- [11] 熊玫. 基于大数据技术的大学生心理危机干预路径研究[J]. 辽宁经济职业技术学院学报, 2023(3): 146-148.
- [12] 付茜茜. 从“内卷”到“躺平”: 现代性焦虑与青年亚文化审思[J]. 青年探索, 2022(2): 80-90.
- [13] 潘泽泉. 网络“陌生人社”行为的心理与本质[J]. 人民论坛, 2020(30): 78-81.
- [14] 吴宁, 孔静漪. 社会加速背景下青年的生存焦虑与消解路径——基于罗萨的社会加速理论[J]. 山东青年政治学院学报, 2024, 40(1): 39-45.
- [15] 靳雅晴, 闫伟娜. 社交媒体时代青年群体身材焦虑现象探析[J]. 全媒体探索, 2024(1): 130-132.
- [16] 黄海荣. Soul 发布《2023 年轻人搭子社交报告》[EB/OL]. 新华网. https://www.xinhuanet.com/fortune/2023-06/21/c_1212236557.htm, 2023-06-21.
- [17] 潘泽泉. 青年群体为何流行治愈系文化[J]. 人民论坛, 2024(1): 84-86.
- [18] 胡伯项, 吴隽民. 新媒体时代泛娱乐主义对我国主流意识形态的冲击及其应对[J]. 思想教育研究, 2021(10): 85-92.
- [19] 韦理强, 李坤. 焦虑和抑郁障碍的性别差异神经机制[J]. 科学通报, 2024(69): 4470-4485.