

引用格式:朱家明,席浩楠,杨文松,张伟杰.微博情感视角下中国新能源消费的满意度评价[J].中国传媒大学学报(自然科学版),2023,30(06):37-43.

文章编号:1673-4793(2023)06-0037-07

微博情感视角下中国新能源消费的满意度评价

朱家明^{1*},席浩楠²,杨文松²,张伟杰²

(1.安徽财经大学大数据学院,蚌埠 233030;2.安徽财经大学金融学院,蚌埠 233030)

摘要:针对中国新能源消费满意度评价,从微博情感分析的视角,利用网络爬虫程序获取新浪微博评论数据,来研究社会公众的满意度评价。从微博评论数据情感值分析发现,社会公众对“社会发展”、“环保”、“比亚迪新能源”、“GDP”、“新能源消费”的满意度依次递减;从词云图分析来看,“新能源”、“人均GDP”、“经济发展”、“环境”等是公众关注的热门议题。研究得出:将互联网社群的民意与政府的施政措施有机结合,可以有效改进并提升社会公众的满意度。

关键词:新能源消费;满意度评价;情感值分析;词云图分析;中国

中图分类号:F426 **文献标识码:**A

Satisfaction evaluation of new energy consumption in China from a microblog emotion perspective

ZHU Jiaming^{1*}, XI Haonan², YANG Wensong², ZHANG Weijie²

(1. School of Big Data, Anhui University of Finance and Economics, Bengbu 233030, China; 2. School of Finance, Anhui University of Finance and Economics, Bengbu 233030, China)

Abstract: In order to evaluate the satisfaction of China's new energy consumption, we used a web crawler program to obtain Sina Weibo comment data from the perspective of microblog sentiment analysis to study the public's satisfaction evaluation. From the analysis of the sentiment value of the microblog comment data, it is found that public satisfaction with "social development", "environmental protection", "BYD new energy", "GDP" and "new energy consumption" is decreasing in order. From the word cloud analysis, "new energy", "GDP per capita", "economic development" and "environment" are the most popular issues of public concern. The study concludes that combining the public opinion of the Internet community with the government's administrative measures can effectively improve and enhance the satisfaction of the public.

Keywords: new energy consumption; satisfaction evaluation; sentiment value analysis; word cloud graph analysis; China

1 引言

能源消费与国家经济发展关系密切,进而影响到国民的生活水平。然而煤、石油、天然气等能源在消

耗过程中会排放大量的温室气体,反而对经济发展产生负面影响,不利于经济的稳定发展。制定合理的能源政策来推动经济发展成为当前重要话题之一^[1]。在经济增长保稳同时,降低常规能源依赖,减慢能源消

基金项目:国家社会科学基金年度项目(22BTJ048);安徽省高等学校省级一流课程“高等数学”(2021xxkc002)

作者简介(*为通讯作者):朱家明(1973-),男,副教授,硕士,主要从事环境经济统计研究。Email:zhujm1973@163.com;席浩楠(2003-),男,本科生,主要从事金融科技研究。Email:Xihn2003@163.com;杨文松(2003-),本科生,主要从事金融研究。Email:2842441496@qq.com;张伟杰(2003-),男,本科生,主要从事金融研究。Email:2063567707@qq.com

耗速度,关键在于优化能源消费结构,增大新能源消耗量占比。本文从三个方面展开研究:从研究内容来看,现有文献主要研究能源消费结构的影响因素^[2-8],很少涉及到新能源消费的满意度评价;从研究视角看,已有文献大多是利用统计年鉴的客观数据来分析新能源发展,缺少站在公众角度上的对中国新能源消费的满意度评价;从研究方法来看,现有研究多是采取数学折线图、数据表以及数学公式等,采用大数据技术和python语言将微博评论文本分析用于新能源消费的满意度评价的方法较少。

为了研究社会公众对新能源消费的了解,本文利用新浪微博评论数据,采用情感分析和词云图分析方法归纳公众对中国新能源消费的满意度评价。采用定量和定性研究方法对发布在新浪微博上的相关评论进行分析,旨在深入了解公众对新能源消费的态度和看法,以揭示公众在新能源消费方面的意见和建议,同时为政府制定新能源政策提供一定的参考。

2 构建中国新能源消费的满意度评价指标体系

2.1 研究思路

本文旨在利用情感分类模型对爬取的微博文本中的大众情绪进行数据挖掘。拟以新浪微博平台数据为基础,参考邢万里等^[9]的研究,从经济增长、产业发展、能源消费、能源经济四个维度来考察中国新能源消费的满意度评价指标体系,研究社会公众对中国新能源消费分维度的情感态度。

2.2 数据收集

(1)关键词选取

选用数据来源于新浪微博用户对已发表微文进行的评论。在研究新能源消费的满意度评价时,首先需要确定相关微文话题,即关键词的选取^[10-11]。关键词选取原则是确保其能够准确反映中国新能源消费基于不同维度的特点,与每个维度相关的可选取关键词具体见表1。

表1 关键词选取

维度	关键词
经济增长	GDP、劳动效率、消费需求、技术进步、产业结构、新基建
产业发展	技术创新、国民经济、高能级创新、光伏产业、产业链、比亚迪新能源
能源消费	能源生产、能源利用效率、中国经济、环保、能源消费结构
能源经济	配额制、新能源发展、可持续发展战略、社会发展、和谐社会、节能减排

对表1中的能源消费、经济增长、产业发展、能源经济四个维度所提取关键词进行搜索,并查找微博对于相关关键词的话题讨论度、评论用户群体分布、涵盖的范围(需要具有针对性、代表性),发现部分关键词存在不能令人满意的地方,如“能源消费结构”、“配额制”、“高能级创新”等词专业性较强,导致相关话题热度较低、讨论群体较单一,无法获取足够的样本;个别关键词涵盖的范围过大,如搜索“新基建”会出现较多“沙盒类游戏”的商业广告,导致收集的样本不纯,不具备切题性;还有一些关键词涵盖的范围过窄,如“光伏产业”、“节能减排”等词,无法准确反映所对应的评价维度。鉴于“GDP”、“比亚迪新能源”、“环保”和“社会发展”的微文在质量和切题度方面相对其他关键词更为突出,综上所述,本文选定其作为四个评价维度中的关键词,并选取“新能源消费”作为中国新能源消费

整体评价的关键词。

(2)数据爬取

使用八爪鱼采集器对微博网页端数据进行爬取,相关爬取方法如下:登录新浪微博网页端,在新浪微博上进行高级搜索,输入关键词,复制新浪微博网址到八爪鱼采集器,设置采集步骤,按类别依次生成相应网址,收集并导入网址,再次设置采集步骤,设定目标爬取字段并设置采集数据的范围,收集相关微博评论。采集的字段为微文评论内容。

2.3 数据预处理

数据预处理的目的是确保数据的质量和一致性,以便更好地对后续结果进行分析。数据预处理的质量对后续分析的准确性和可靠性有着至关重要的影响。具体步骤如下:(1)利用八爪鱼采集器将爬

取的数据导入 Excel 表格;(2)数据清洗,识别和处理包含无效信息、列名不规范、格式不一致、异常值、缺失值、重复值、不一致值等问题,从不同数据源中整合数据,消除重复数据,通过定义好的数据清洗策略和清洗规则,获取满足质量的数据,提高分析效率,以便后续分析;(3)数据转换,将数据转换为适合分析的格式,包括数据类型转换、缩放、归一化、标准化等;(4)中文分词,由于中文语言的特殊性,相比于英语等拉丁语系语言,中文语言中的词汇没有明显的单词边界,也没有诸如空格之类的分隔符,这导致了在进行自然语言处理(NLP)任务时,需要对中文文本进行分词处理,以便更好地进行文本处理和分析。为方便后续的文本处理和分析,本文使用 SnowNLP 文本分析库基以 Character-Based Generative Model 算法实现中文分词。

2.4 情感分析

情感分析是针对非结构化的网络评论数据进行分析,以确定其中表达的情感状态的技术。它通常将情感分为积极和消极两种类型,并通过对文本等数据的分析,基于层面的情感分析将深入一个层次,以确定哪些特定特征或层面正在产生正向或负向的情感。本文基于处理中文的自然语言的方法,将文本信息表示成向量后通过机器学习来判断情感类型,采用 Python 中的一个自然语言处理库——SnowNLP 库进行情感分析。SnowNLP 情感分析是基于情感分类的技术。它首先从一系列文本中提取情感特征,然后使用机器学习模型对文本进行分类。具体步骤包括情感特征提取、情感分类和结果分析。

(1)情感特征提取

从文本中提取情感特征。它可以使用语义分析,如词性标注和词汇表,以及自然语言处理方法,如术语提取和情感分析,从文本中提取语义特征以用于训练情感分类器的过程。在 SnowNLP 情感分析中,常用的情感词典包括 How Net 情感词典、Baidu LAC 情感词典、NTUSD 情感词典等。同时, SnowNLP 还提供了自定义情感词典的功能,允许用户根据具体应用场景和需求自定义情感词典,以提高情感分析的准确性。

(2)情感分类

使用已标记好的数据集训练朴素贝叶斯情感分类器对文本进行划分,以确定文本中情感的正负类

型。需要注意的是,在进行情感分类时,特征提取的质量对分类器的准确性有很大的影响,因此需要选择合适的特征提取方法。同时,情感分类器的训练数据集也需要具有代表性和充分性,以提高情感分类器的可靠程度。

(3)结果分析

SnowNLP 文本分析库通过朴素贝叶斯分类器将文本分为积极和消极两类,返回值处于[0,1]区间。对文本数据在每一种情感类型下的得分进行分析,以确定文本数据所表达的情感类型。通过对情感得分进行分析,了解文本数据在不同情感类型下的情感强度和倾向,以确定文本数据所表达的情感类型和情感强度。同时也可以比较不同文本数据在情感得分上的差异和相似度,以了解不同文本数据的情感倾向和重点内容。

3 基于 SnowNLP 算法对微文评论的结果分析

3.1 研究思路

使用八爪鱼采集器爬取新浪微博用户的微文评论数据,将数据进行预处理后得到关于“GDP”话题的微文评论 1560 条、关于“比亚迪新能源”话题的微文评论 3383 条、关于“环保”话题的微文评论 1803 条、关于“社会发展”话题的微文评论 3079 条、关于“新能源消费”话题的微文评论 3297 条,共计 13122 条。

3.2 研究方法

SnowNLP 文本分析库通过朴素贝叶斯分类器返回的情感值处于[0,1]区间,不利于直观判断正负情感类型,为了能够从情感值上更直观地展示出正负情感类型,本文将 SnowNLP 算法计算出来的情感值减去 0.5,使得情感值处于[-0.5,0.5]的对称区间内。具体的情感类型分类如表 2 所示。

表 2 情感分类准则

分类准则	消极	负向	中立	正向	积极
情感值	[-0.5, -0.3]	[-0.3, -0.1]	[-0.1, 0.1]	[0.1, 0.3]	[0.3, 0.5]

分别计算 5 个关键词的情感均值,计算方法如(1)所示:

$$\bar{E}_i = \sum_{h=1}^n e_h / n, n = 1, 2, 3, \dots, \infty \quad (1)$$

关于公式, E 为第 i 个关键词的平均情感值, e 为第 i 个关键词的第 h 个样本的情感值, n 为第 i 个关键词的微文评论条数。

3.3 情感值统计特征结构分析

统计各个关键词的情感值的统计特征,其结果如表3所示。

表3 各关键词情感值的统计特征

统计量	均值	标准差	最小值	25%分位数	中位数	75%分位数	最大值
GDP	0.0994(中立)	0.3278	-0.5000	-0.1649	0.1347	0.4217	0.5000
比亚迪新能源	0.1131(正向)	0.3141	-0.5000	-0.1381	0.1662	0.4111	0.5000
环保	0.2028(正向)	0.3166	-0.5000	-0.0236	0.3508	0.4764	0.5000
社会发展	0.2151(正向)	0.3221	-0.5000	-0.0200	0.3693	0.4976	0.5000
新能源消费	0.0341(中立)	0.3027	-0.5000	-0.2169	0.0262	0.3130	0.5000

从表3可以看出:(1)GDP,公众对于GDP的社会评价情感均值为0.0994,属于中立的情感类型,是所有维度中最低的,中位数为0.1347,表明超过一半的微博发言用户对GDP都持有正向看法,同时标准差为0.3278,说明整体变化幅度较大。(2)比亚迪新能源,公众对其情感态度表现出正向,均值为0.1131,中位数为0.1662,呈现出一半以上微博发言用户的正面态度。然而,对比亚迪新能源的微博发言中仍然存在负面情感倾向,最小值为-0.5000,标准差为0.3141,显示情感倾向具有较大浮动。(3)环保,针对环保问题的公众情感均值为0.2028,具有积极的倾向。此外,中位数为0.3508,表明超过一半的微博用户持有积极的态度。然而,标准差为0.3166,说明公众对环保问题的感情倾向存在较大的变化。因此,我们需要更加系统和全面地了解公众对环保问题的态度及其变化趋势,以便更好地推动环保事业的发展。(4)社会发展,公众对于社会发展的情感均值为0.2151,持正向态度,接近积极态度,中位数为0.3693,说明一半以上的微博发言用户持积极的态度,标准差为0.3221,说明情感值波动幅度较大。(5)新能源消费,在新能源消费领域,公众对新能源消费的情感均值为0.0341,表现出中性的情感类型。通过情感分布的观察,可以发现75%的用户对新能源消费的评价情感值超过了0.3130,这表明超过25%的用户对新能源消费的评价是积极的。这些结果说明了公众对新能源消费仍保持观望态度,而同时存在对此的积极评价。

总体而言,微博用户对“社会发展”方面的评价是最积极的,相较而言,对“GDP”则持理性态度。另外,用户对“环保”和“社会发展”的满意度普遍较高,但对“比亚迪新能源”的评价尚有待提高。总的来说,四个维度的评价差异并无过于显著的特点,均存在一定的提升余地。为了更好地推广和发展新能源消费和其相关领域,必须对公众的需求和评价有更深入的了

解,并相应地采取措施和加以改善。

为了更加直观地呈现情感值的变化趋势,绘制“新能源消费”总体以及不同维度下的情感值变化图,深度分析各关键词的情感变化趋势,旨在揭示新能源消费领域的情感态势和发展趋势,为进一步研究提供有益的启示。

(1)对“GDP”的总体情感分析

“GDP”的情感值变化如图1所示。从图1可以看出,公众对于“GDP”的情感值有相当一部分处于较高的水平,且大部分集中在[0.3, 0.5]区间,虽然大部分情感值都属于正向评价,却也存在一定数量的负面评价,说明微博发言用户对于我国GDP持中立态度。结合实际来看,公众对于我国GDP的积极评价主要集中在两个方面:一是民营经济高速发展,并且此趋势备受公众欢迎;二是超额储蓄向消费转化,公众对此表示积极支持。消极评价主要体现在对我国贫富差距问题的关切上。

(2)对“比亚迪新能源”的总体情感分析

“比亚迪新能源”的情感值变化如图2所示。分析图2可知,公众对于“比亚迪新能源”的负面看法主要集中在对比亚迪新能源汽车自燃事件的不满和担忧。这些事件凸显技术的不成熟性可能会引发一连串的安全问题。作为汽车生产企业,比亚迪新能源在电动汽车领域拥有良好的口碑。然而,在面临电动汽车自燃事件之后,消费者的信心开始动摇。这不仅严重影响了比亚迪新能源的形象,也引发了大众对于整个新能源汽车产业的安全性和可靠性的担忧。而对比亚迪新能源持积极态度的群体则认为作为目前中国领先的新能源汽车制造商之一,比亚迪在技术、产品和市场等方面都有着不错的表现和前景。

(3)对“环保”的总体情感分析

“环保”的情感值变化如图3所示。

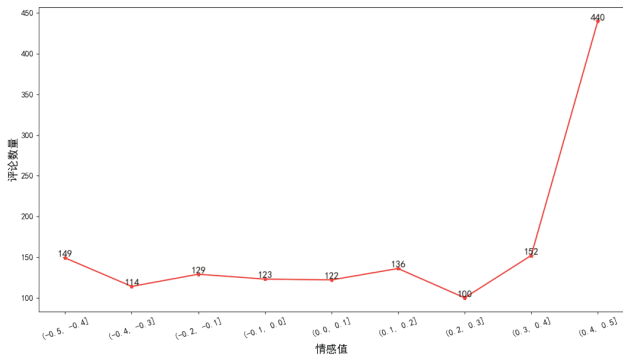


图1 “GDP”情感值变化图

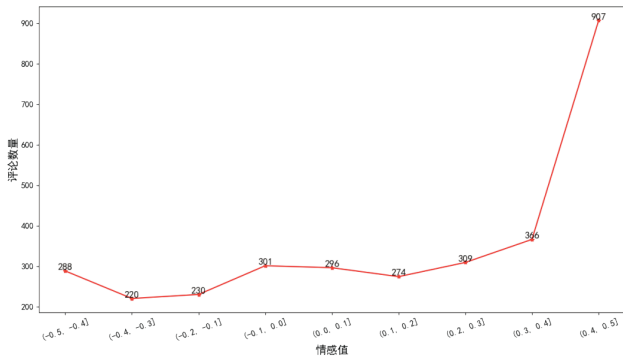


图2 “比亚迪新能源”情感值变化图

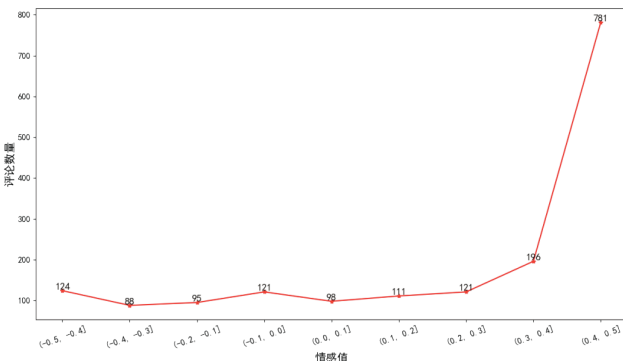


图3 “环保”情感值变化图

从图3可以得出,环保已成为社会责任和可持续发展的核心要素之一,其重要性也日益凸显,从个人到国家都对环保问题高度重视。政府在推动环保事业方面给予了极大的政策支持,这为提高环保意识奠定了基础。因此,在政府政策的推动和社会环境的影响下,情感水平的平均值处于较高水平上。

(4)对“社会发展”的总体情感分析

“社会发展”的情感值变化如图4所示。从图4可以观察到公众对“社会发展”话题的情感值在[-0.5,0.4]分布较为均匀,而在[0.4,0.5]区间骤增,所以情感均值是偏向积极的。主要是因为公众对“脱贫攻坚任务胜利完成”、“生态环境明显改善”、“基础设施更加完善”等社

会发展成就的认同。

(5)对“新能源消费”的总体情感分析

“新能源消费”的情感值变化如图5所示。从图5明显可以观察到公众在“新能源消费”方面的情感值分布存在较大差异,说明公众对此话题的态度存在着褒贬不一的情况,各有侧重。其中,持正向情绪的微博发言用户表达了对新能源消费补贴的支持以及对新能源产业快速发展的称赞。而持有负面情绪的微博发言用户则表达了对新能源发展前景的担忧,这显示了公众对新能源消费及其相关领域的发展依然存在着一定的疑虑与忧虑。

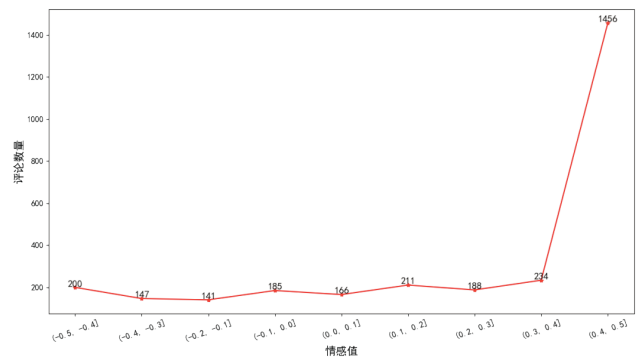


图4 “社会发展”情感值变化图

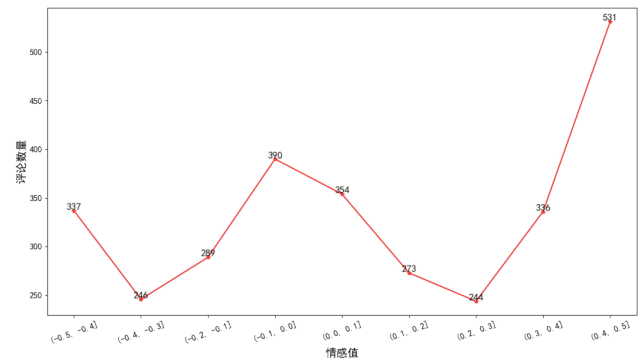


图5 “新能源消费”情感值变化图

4 基于词云图对微文评论的结果分析

4.1 研究思路

词云图是一种以词语频率或权重为基础,在分词基础上将文本中出现频率高、权重大的关键词突出显示,将文本数据可视化展示的图形化方法。词云图通过文字出现的大小、色彩直观地展示文本中不同词语的分布情况,为文本数据提供了简洁、直观的展示方式。词云图中字体的大小与该词组的词频密切相关,反映了公众对于相关热点和主题的关注程度,具有重要的数据展示价值。

司的新能源产业、GDP以及新能源产品消费等方面的评价时,用户群体的满意程度呈现出逐步递减的态势;(2)新能源、GDP、经济发展与环境保护等议题均为全社会关注焦点。公众对环保型新能源产品保持高度认可和支持,但仍对区域发展水平、政策保障以及财政拨款方面存在疑虑。

以上研究结果对我国新能源市场发展提供以下建议:(1)优化现有新能源消费环境,贯彻实施各类保障补贴政策,加大基础设施建设力度,推动新能源的全方位发展。(2)重视网络交汇处热点事件,关注民生问题,响应国内外双循环,拉动消费投资,实现经济高质量发展。

参考文献

- [1] 赵艳莉.中国新能源消费对经济增长的影响[J].安庆师范大学学报(社会科学版),2021,40(03):96-99.
- [2] 辛金国,童佳耀.浙江省绿色发展对能源消费结构的影响研究[J].统计科学与实践,2022,(08):30-33+56.
- [3] 葛鑫.绿色金融对我国能源消费结构优化的影响研究[D].长春:吉林大学,2022.
- [4] 张人文.盐亭县农民生活能源消费选择的影响因素研究[D].乌鲁木齐:新疆农业大学,2022.
- [5] 刘乐.绿色金融对能源消费结构的影响[D].武汉:华中科技大学,2021.
- [6] 熊邦娟.绿色金融发展对我国能源消费结构的影响研究[D].重庆:重庆工商大学,2021.
- [7] 王盼.产业结构调整对能源消费结构的影响研究[D].长沙:湖南大学,2020.
- [8] 倪强.中国居民能源消费城乡差异分解及影响因素研究[D].杭州:浙江工商大学,2018.
- [9] 邢万里.2030年我国新能源发展优先序列研究[D].北京:中国地质大学(北京),2015.
- [10] 朱家明,张蒙.基于空间杜宾模型的绿色金融经济增长效应研究[J].中国传媒大学学报(自然科学版),2023,30(1):75-82.
- [11] 崔连标,刘泽晨,王佳雪.数字经济对长江经济带实体经济发展的影响研究-基于环境规制和经济发展的调节效应[J/OL].华北水利水电大学学报(社会科学版),2023-12-27.

编辑:王谦