

分类号: C931.6

密 级: 公开

学校代码: 10363

学 号: 2220720107



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目 考虑用户需求细粒度情感分析的
新能源汽车属性重要度排序研究

论 文 作 者 盛志远

指 导 教 师 沈超

学 科 (专 业) 管理科学与工程

研 究 方 向 大数据与数字化管理

论文提交日期: 2025 年 5 月 14 日

分类号：C931.6 单位代码：10363

密级：公开学号：2220720107

题 目：考虑用户需求细粒度情感分析的新能源汽车属性重要度排序研究

英文并列

题 目：A Study on the Importance Ranking of New Energy Vehicle
Attributes Based on Fine Grained Sentiment Analysis of User Needs

学生姓名：盛志远

指导教师：沈超

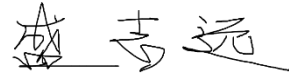
专 业：管理科学与工程

研究方向：大数据与数字化管理

论文答辩日期：2025 年 5 月 14 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律结果由本人承担。



学位论文作者签名：

日期：2025 年 5 月 14 日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

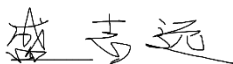
保密 ☐，在 ____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：

日期： 2025 年 5 月 14 日



指导教师签名：

日期： 2025 年 5 月 14 日



考虑用户需求细粒度情感分析的新能源汽车属性 重要度排序研究

摘 要

当前，新能源汽车市场竞争日趋激烈，设计出不断满足客户个性化、时代性需求的汽车产品能够助力车企迅速占领市场，这是汽车厂商需要考虑的关键问题。新能源汽车作为功能多样、技术密集的一类复杂产品，由于多年的发展，其产品生产流程虽趋于标准化，但在面向客户进行创新型开发时需要考虑诸多不确定性因素。最关键的，面对消费者对新能源汽车质美价优的普遍需求，设计师们在产品开发时不能面面俱到，粗略给出大而不全且不计成本的产品，这时，了解客户对新能源汽车产品功能属性的需求重要度就显得尤为重要。

传统情境下的问卷调查、焦点小组访谈等调查方式不足以准确、客观、全面地了解客户需求，尤其是隐形需求。社交媒体情境下，海量的消费者在线评论数据是用户对产品体验的真实表达，能较好的反映出客户对产品深层次内心需求。借助文本挖掘技术，可以捕捉到在线评论中消费者对产品特性及其偏好态度的准确描述，体现了消费者对产品功能的重要度排序，而这恰恰映射了产品开发时设计师们所要考虑的产品属性优先级。

机器学习特别是深度学习技术的发展，使得从在线评论文本数据

中获取的知识更加丰富且精细，尤其在对自然语言中客户深层次情感的挖掘上，可以获取更加细粒度的客户情感需求。为此，本研究以新能源汽车为研究对象，借助于消费者在线评论数据，基于文本挖掘与深度学习方法，对在线评论中新能源汽车用户的需求和细粒度情感进行识别和分析，并在此基础上，采用多准则决策方法对新能源汽车产品属性的重要度进行排序，从而为设计师开发产品提供决策支持。具体的研究内容包括：

（1）运用文本挖掘方法识别用户的产品需求。获取了 2014-2019，2020-2024 两个连续五年的某汽车论坛新能源汽车评论数据。在文本预处理的基础上，考虑不同阶段用户关注重点、表达方式和需求特征的差异性，结合 LDA 模型的显性主题提取能力和 LSA 模型的隐性语义识别能力，从多个层面提取客户需求信息，以获得更为全面的客户需求，同时，分析了汽车用户需求特征的变化趋势。

（2）通过深度学习与机器学习方法对评论文本进一步处理以获取细粒度的情感需求。结合 Bert-Bi-LSTM 和支持向量机，充分利用深度学习与传统机器学习模型的优势，实现更加精细的情感分类。在用户情感分类的基础上，进一步计算不同情感状态下用户对新能源汽车各需求属性的关注度分布，以量化情感因素对需求排序的影响。引入贝叶斯推断方法对需求权重进行动态调整，从而精准刻画情感波动如何影响用户偏好，确保需求排序的准确性。结合 Kano-IPA 模型，发现情感极性直接影响需求属性的分类。

（3）基于多准则决策方法，考虑用户的情感偏好，实现对产品

需求重要度排序。结合 VIKOR 模型与 LightGBM 模型，构建了一种兼具静态权衡与动态特征选择的排序框架。通过多模型融合的情感分析方法，精准识别用户评论中的情感倾向，并结合贝叶斯推断动态调整情感权重，以量化情感因素对需求排序的影响。结果表明，不同情感状态下用户对新能源汽车属性的关注度存在显著差异，情感极性直接影响属性分类。通过构建属性-情感-决策框架，为新能源汽车厂商和设计师面向用户需求的研发提供决策支持，有助于企业精准定位市场需求并优化产品设计策略。

关键词：属性重要度排序，新能源汽车，细粒度情感分析，深度学习，在线评论

A Study on the Importance Ranking of New Energy Vehicle Attributes Based on Fine Grained Sentiment Analysis of User Needs

ABSTRACT

Currently, the competition in the new energy vehicle (NEV) market is becoming increasingly intense. Designing automotive products that continuously meet customers' personalized and contemporary needs is crucial for manufacturers aiming to rapidly capture market share. As NEVs are complex products characterized by diverse functions and intensive technologies, their production processes have become relatively standardized through years of development. However, uncertainties must still be addressed during customer-oriented innovative development. Faced with consumers' general expectations for high-quality and cost-effective NEVs, designers cannot cater to all aspects simultaneously. It is therefore essential to identify the importance of different product attributes from the customer's perspective to inform product development decisions.

Traditional methods such as surveys and focus group interviews often fall short in accurately, objectively, and comprehensively capturing customer needs, particularly implicit needs. In the context of social media, massive volumes of online consumer reviews offer authentic reflections of user experiences and serve as valuable data sources for uncovering deeper

customer requirements. Through text mining techniques, it is possible to extract precise descriptions of product features and user preferences from these reviews, thereby revealing the relative importance of different product attributes. This, in turn, directly informs the prioritization of product development.

The advancement of machine learning, particularly deep learning, has enabled the extraction of more fine-grained and richer knowledge from online reviews. Especially in mining consumers' deep emotional needs expressed in natural language, these techniques enhance the understanding of sentiment-driven preferences. This study focuses on NEVs and utilizes online review data, employing text mining and deep learning methods to identify and analyze user needs and fine-grained sentiments. Based on this, a multi-criteria decision-making approach is adopted to prioritize NEV product attributes, providing design decision support. The main contributions of this study include:

(1) Identifying user product needs through text mining. Review data from an automotive forum covering two continuous five-year periods (2014–2019 and 2020–2024) were collected. After text preprocessing, considering variations in user concerns, expressions, and needs across different time periods, both Latent Dirichlet Allocation and Latent Semantic Analysis models were employed to extract explicit topics and latent semantics, respectively, to comprehensively identify customer needs.

The study also analyzes the evolution of user demand characteristics over time.

(2) Extracting fine-grained sentiment needs from review texts using deep learning and machine learning approaches. By integrating BERT-BiLSTM and Support Vector Machine models, the strengths of deep learning and traditional machine learning are leveraged for refined sentiment classification. Based on sentiment classification results, the distribution of user attention toward various NEV attributes under different emotional states is quantified. Bayesian inference is introduced to dynamically adjust attribute weights, accurately reflecting how sentiment fluctuations affect user preferences. The Kano-IPA model is further employed to demonstrate the impact of emotional polarity on attribute categorization.

(3) Prioritizing product attribute importance using a multi-criteria decision-making framework that incorporates emotional preferences. A hybrid ranking framework combining the VIKOR model and LightGBM is constructed, integrating static trade-offs and dynamic feature selection. Sentiment analysis results are enhanced through model fusion and sentiment weight adjustment via Bayesian inference, enabling quantification of the impact of emotions on demand prioritization. The results reveal significant differences in attribute attention under different emotional states, and emotional polarity is found to directly influence

attribute categorization. By constructing an attribute–emotion–decision framework, this study provides design decision support for NEV manufacturers and designers, aiding enterprises in accurately targeting market needs and optimizing product development strategies.

Keywords: New Energy Vehicle, Attribute Importance Ranking, Fine-Grained Sentiment Analysis, Deep Learning, Online Reviews

目 录

摘 要.....	I
ABSTRACT.....	IV
目 录.....	VIII
第 1 章 绪论.....	1
1.1 研究背景.....	1
1.2 研究目的与意义.....	2
1.2.1 研究目的.....	2
1.2.2 研究意义.....	2
1.3 研究内容与方法.....	3
1.3.1 研究内容.....	3
1.3.2 研究方法.....	4
1.4 研究框架与技术路线.....	5
1.5 论文创新性.....	7
第 2 章 相关理论方法与文献综述.....	8
2.1 相关理论与方法.....	8
2.1.1 自然语言处理.....	8
2.1.2 分类、决策与排序.....	16
2.2 文献综述.....	19
2.2.1 文献检索情况.....	19
2.2.2 产品开发需求识别与分析.....	21
2.2.3 基于在线评论的情感分析.....	23
2.2.4 产品属性重要度排序方法.....	24
2.2.5 文献评述.....	26
第 3 章 基于在线评论的 NEV 用户需求识别与获取	27
3.1 数据采集与清洗.....	27
3.2 文本标准化处理.....	28

3.3 主题提取建模分析.....	29
3.4 实验结果与讨论.....	32
3.5 本章小结.....	33
第 4 章 融合多模型的 NEV 用户细粒度情感分析	35
4.1 SVM 情感分析模型构建.....	35
4.2 Kano 模型与 IPA 模型构建.....	36
4.3 Bert-Bi-LSTM 模型构建.....	38
4.4 贝叶斯推断的用户情感量化分析模型构建.....	40
4.5 实验结果与讨论.....	40
4.6 本章小结.....	47
第 5 章 基于 MCDM 的 NEV 属性重要度排序.....	48
5.1 主观权重设定排序方法.....	48
5.2 客观权重设定排序方法.....	49
5.3 实验结果与讨论.....	50
5.4 本章小结.....	55
第 6 章 结论与展望.....	57
6.1 研究结论.....	57
6.2 研究局限性与展望.....	57
参考文献.....	59
附录.....	69
附录 1 新能源汽车用户在线评论示例.....	69
附录 2 在线评论数据预处理示例.....	69
附录 3 在线评论属性划分方法示例.....	70
附录 4 情感分析结果部分示例.....	71
攻读硕士学位期间发表或录用的学术论文.....	72
发表论文.....	72
申请发明专利.....	72
参与科研项目.....	72
致谢.....	73

插图清单

图 1-1 研究框架	5
图 1-2 技术路线图	6
图 2-1 LDA 模型的贝叶斯网络结构	8
图 2-2 支持向量机模型图	11
图 2-3 Transformer 模型的编码器	12
图 2-4 Bert 模型的网络结构	13
图 2-5 LSTM 神经网络结构	14
图 2-6 Bert-Bi-LSTM 模型结构	15
图 2-7 Leaf-wise 的决策树生长策略	17
图 2-8 直方图算法示意图	18
图 2-9 关键词共现图 (CNKI)	20
图 2-10 关键词共现图 (WOS)	21
图 3-1 部分新能源汽车用户在线评论数据	27
图 3-2 部分预处理数据	29
图 4-1 新能源汽车 IPA 分析图	41
图 4-2 积极情感下贝叶斯推断	45
图 4-3 中性情感下贝叶斯推断	46
图 4-4 消极情感下贝叶斯推断	46
图 5-1 特征值选择	50

表格清单

表 3-1 LDA 主题模型核心参数	30
表 3-2 LSA 主题模型核心参数	31
表 3-3 潜在狄利克雷分配模型分析结果	31
表 3-4 潜在语义分析模型分析结果	32
表 4-1 LSA 核心参数值	35
表 4-2 得分区间及对应标签	36
表 4-3 Kano 类别分类规则	36
表 4-4 Kano-IPA 模型	38
表 4-5 Bert-Bi-LSTM 模型核心参数	39
表 4-6 新能源汽车 Kano 模型分析	41
表 4-7 Kano-IPA 模型分类结果	42
表 4-8 BERT-BiLSTM 分析结果	43
表 5-1 LightGBM 模型核心参数设置	49
表 5-2 VIKOR 模型分析结果	51
表 5-3 LightGBM 分析结果	52

第 1 章 绪论

1.1 研究背景

随着全球能源结构调整与碳中和目标的推进，新能源汽车（New Energy Vehicle, NEV）产业迎来爆发式增长，已成为未来汽车行业转型的重要方向。据中国汽车工业协会统计，2024 年中国新能源汽车产销量已突破 1288 万辆，预计 2030 年保有量将超过 1.5 亿辆。在政策支持、技术进步和市场需求的共同推动下，新能源汽车逐步进入普及阶段，行业竞争日趋激烈，用户需求也呈现出更加多元化、动态化的特征。

在市场快速扩张的同时，消费者的购车决策模式正在发生深刻变化。传统燃油车时代，用户关注的核心属性主要集中在续航、价格、动力和安全性等基本功能上，而在新能源汽车时代，智能化、外观设计、驾驶体验和充电便捷性等新兴属性逐渐成为影响购车决策的重要因素。此外，伴随新能源汽车技术的迭代升级，消费者对于品牌、自动驾驶、车机交互和能耗管理等方面的期待不断提高，不同用户群体的需求差异也日益显著。因此，精准捕捉用户需求变化，理解其情感倾向，并据此优化产品设计和市场策略，已成为新能源汽车行业发展的关键问题。

然而，用户在不同情境下的需求表达往往带有强烈的主观情感色彩，如何量化情感因素对需求偏好的影响，是当前需求分析研究中的一大挑战。现有研究大多关注正向评价对产品优化的作用，而忽略了中性和负面评价所蕴含的深层次需求信号。实际情况是，用户的负面评论可能反映产品的关键痛点，而中性评价则可能揭示尚未被满足的潜在期望。为了更全面地识别用户需求，可以借助真实情境中细粒度情感进行深入挖掘。

在此背景下，在线评论作为用户需求反馈的重要途径，正成为研究情感与需求关系的关键数据来源。伴随电子商务、社交媒体等信息技术的快速发展^[1]，越来越多的消费者愿意在网上表达自己的消费感受。同时，在线评论也直接影响着消费者的购物选择，这也使得其成为制造企业了解消费者需求和满意度的重要方式之一^[2]。如何从大量在线评论数据中挖掘出影响消费者满意度的因素，以及对这些因素进行排序都是值得研究的问题^[3]。

近年来，学术界对在线评论的各类研究也在不断深入。现有研究在技术和应用两个方面广泛利用文本挖掘和情感分析方法，提取在线评论中的关键信息，以支持产品优化和市场分析^[4]。本研究在利用文本挖掘、情感分析及在线评论数据进行信息提取和应用的基础上，构建新能源汽车属性重要度排序模型。

1.2 研究目的与意义

1.2.1 研究目的

本研究计划达到以下目的：

第一，基于在线评论数据识别用户需求，采用潜在狄利克雷分配(Latent Dirichlet Allocation,LDA)和潜在语义分析(Latent Semantic Analysis,LAS)相结合的方式，从不同层面提取用户需求信息。构建更全面的用户需求分析模型。

第二，量化情感因素的作用机制，构建细粒度情感分析模型，突破传统二元划分的局限性，解析不同情感状态下用户对属性的关注差异。

第三，实现动态优先级排序：结合贝叶斯推断与多属性决策模型，动态调整情感权重，捕捉中性情感下的理性评估与消极情感下的痛点反馈。

1.2.2 研究意义

(1) 理论意义

首先，相关研究丰富了情感分析理论，提出的细粒度情感分析模型，不再局限于传统的二元划分，弥补了传统方法在处理用户情感多样性上的不足。不仅为情感驱动的需求分析提供了全新的视角，也为后续探讨情感与用户行为之间的内在联系提供支持。其次，完善多模型融合方法，通过构建贝叶斯推断与 LightGBM 模型，调整权重机制，实现了情感因素与多准则决策方法的融合。既兼顾了静态属性评估的稳定性，又具备了动态响应用户情感波动的能力，为复杂系统下的用户行为分析提供了新的方法参考。最后，完善排序方法和理论。所构建的属性、情感和排序的综合框架，将传统静态模型与动态模型相结合，从而更全面地反映用户需求的细粒度特性。这一框架不仅能够实时捕捉和调整用户需求的优先级，还为产品设计、市场策略和用户体验优化提供了科学的理论支持。

（2）实践意义

本研究构建的新能源汽车属性重要度排序模型，通过整合多维情感分析与多准则决策方法，为企业提供了一种科学、实时的用户需求洞察工具。借助在线评论数据，该模型能够精准捕捉消费者在关键属性上的情感波动和关注度变化，为产品设计、精准营销及售后服务的优化提供可靠的数据支撑。通过实时动态调整情感权重与需求排序，企业可更快响应市场变化，降低决策风险，实现资源的合理配置，进而提升产品竞争力和用户满意度。此外，研究提出的模型不仅适用于新能源汽车行业，还具有推广到其他快速发展领域的潜力，为企业实现精细化管理和科学决策提供了实践范例，具有较好的现实应用价值。

1.3 研究内容与方法

1.3.1 研究内容

本论文围绕新能源汽车属性重要度排序展开研究，旨在构建一个集文本挖掘、情感分析与多准则决策于一体的综合分析框架。具体研究内容包括以下三个方面：

首先，为了识别与获取用户需求，通过模型融合方法，整合潜在狄利克雷分配与潜在语义分析两种主题提取技术。潜在狄利克雷分配基于概率模型，适用于长文本分析，能够从用户评论中提取出显性主题；而潜在语义分析则通过奇异值分解发掘词语之间的潜在语义关系，有效识别出用户隐性需求。具体的，首先对数据进行分词、去停用词和噪声清洗，然后分别训练两种模型，再利用词向量计算，最终构建出一个系统化的用户需求识别模型。

其次，为量化情感因素对用户属性关注度的影响，采用多模型融合的方法对用户评论进行情感分析。具体来说，结合了 Bert-Bi-LSTM 模型和支持向量机模型（Support Vector Machine, SVM），充分利用深度学习与传统机器学习各自的优势。Bert-Bi-LSTM 用于提取评论中的深层语义特征，提升长文本情感捕捉能力；支持向量机补充了对短文本及边界样本的分类；经过模型融合，将用户评论准确划分为不同情感类别，并计算在不同情感状态下，用户对新能源汽车各核心属性的关注度分布。为进一步优化情感对需求排序的影响，采用贝叶斯推断方法对需求权重进行动态调整，同时结合 Kano-IPA 模型，揭示情感极性如何直接影响属性分类。

最后，结合多准则决策方法构建了一个兼具静态权衡与动态特征选择的新能源汽车属性重要度排序模型。通过融合文本挖掘、细粒度情感分析和多准则决策方法，构建了一个需求识别、情感权重动态调整和属性重要度排序于一体的综合分析框架，为新能源汽车产品设计、市场定位和精准营销提供了较好的数据支撑和决策参考。

1.3.2 研究方法

本研究在总结国内外最新研究成果的基础上，运用文本挖掘、机器学习、深度学习、自然语言处理等理论与方法，系统地研究细粒度情感因素下新能源汽车属性重要度排序，其具体研究方法如下：

（1）文献综述法

对国内外学者关于在线评论、自然语言处理、以及情感分析等相关研究进行系统的收集整理，总结归纳出已有论文的研究重点和趋势，寻找不足或可拓展之处作为深入研究的创新之处，搜寻解决问题的方法以判断其可行性，提炼出研究问题，最后确定论文主题。

（2）文本挖掘

对新能源汽车用户在线评论数据进行预处理，包括评论文本的去噪、分词与词性标注等，以提升后续分析的准确性。通过主题建模和关键词提取方法，对评论文本进行内容识别与主题归类，从而挖掘出用户关注的产品属性。

（3）深度学习

采用传统机器学习与深度学习相结合的方法对用户评论进行细粒度情感分类，通过多模型融合技术提高情感分类的精度和鲁棒性，捕捉文本中的细微情感变化。利用贝叶斯推断方法，对不同情感极性的权重进行动态更新；量化不同情感主导下，消费者对各属性评价的相对权重，为后续排序提供数据支撑。

（4）多准则决策与排序

使用 VIKOR 方法，模拟不同情感主导下的排序情况，直观展示消费者评价的变化；使用 LightGBM 排序学习方法，基于贝叶斯推断得到的权重自动学习属性排序，降低人为主观因素影响。

1.4 研究框架与技术路线

为系统阐述本文的整体逻辑脉络与研究流程，构建了如图 1-1 所示的研究框架。

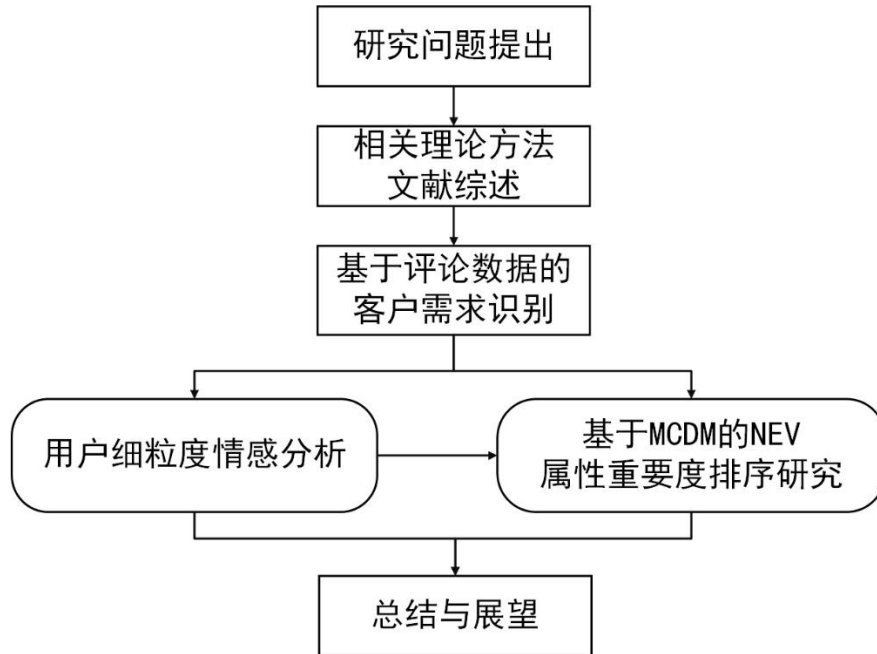


图 1-1 研究框架

第一章，引言。首先阐明研究背景。接着阐述本研究的研究意义，以理论意义和现实意义分别展开，紧接着介绍研究内容，最后阐述本研究的创新性。

第二章，相关理论基础与文献综述。首先介绍了本研究所使用的方法，详细阐述了相关理论；接着通过 VOSviewer 软件对文献进行计量分析揭示自然语言处理未来研究趋势；最后对自然语言处理相关文献进行综述，为后续研究奠定基础。

第三章，基于在线评论的 NEV 用户需求识别与获取。在这一章中，通过文本挖掘技术，从大量新能源汽车用户的在线评论中提取出关键主题和属性，并将这些属性系统地分类。具体来说，先对评论数据进行清洗、分词和停用词过滤，然后利用 LDA 和 LSA 两种方法，对两组数据提取用户需求，这一过程可以全面了解用户关注的重点，为后续的情感分析和决策提供数据支持。

第四章，融合多模型的 NEV 用户细粒度情感分析。本章使用机器学习与深度学习相结合的方法对两组数据进行情感分析，提炼出情感因素对用户需求的重要影响。第一组数据构建 SVM 情感分类器，结合 Kano 和 IPA 模型构建了一个稳定的产品属性分类模型；第二组数据，使用 Bert-Bi-LSTM 模型进行情感分析，

并结合贝叶斯推断方法动态调整情感权重，为后续分析提供有效的数据支持。这种方法互补的分析能够更精准地了解不同情感状态下消费者对汽车各属性的关注变化，从而为产品优化和市场策略制定提供有力的数据支撑。

第五章，基于 MCDM 的 NEV 属性重要度排序。本章在上一章贝叶斯推断获得情感权重的基础上，构建了新能源汽车属性重要度排序模型，并对比了两种排序方法。首先使用 VIKOR 方法，通过人工设定权重来模拟不同情感主导下消费者对各属性的关注优先级；随后，采用数据驱动的 LightGBM 方法，将贝叶斯推断得到的情感权重作为输入，自动学习排序。两种方法的比较分析揭示了在积极、中性和消极情绪下，消费者对汽车属性的关注重点存在明显差异，为产品优化和市场策略制定提供了有力的数据支撑。

第六章，结论与展望。对本研究的主要研究内容的结论进行总结，提出现有研究不足之处，对未来的研究展望。

本研究的技术路线如图 1-2 所示，基于在线评论数据，探讨如何准确识别用户需求以及细粒度情感因素对新能源汽车属性重要度的影响。

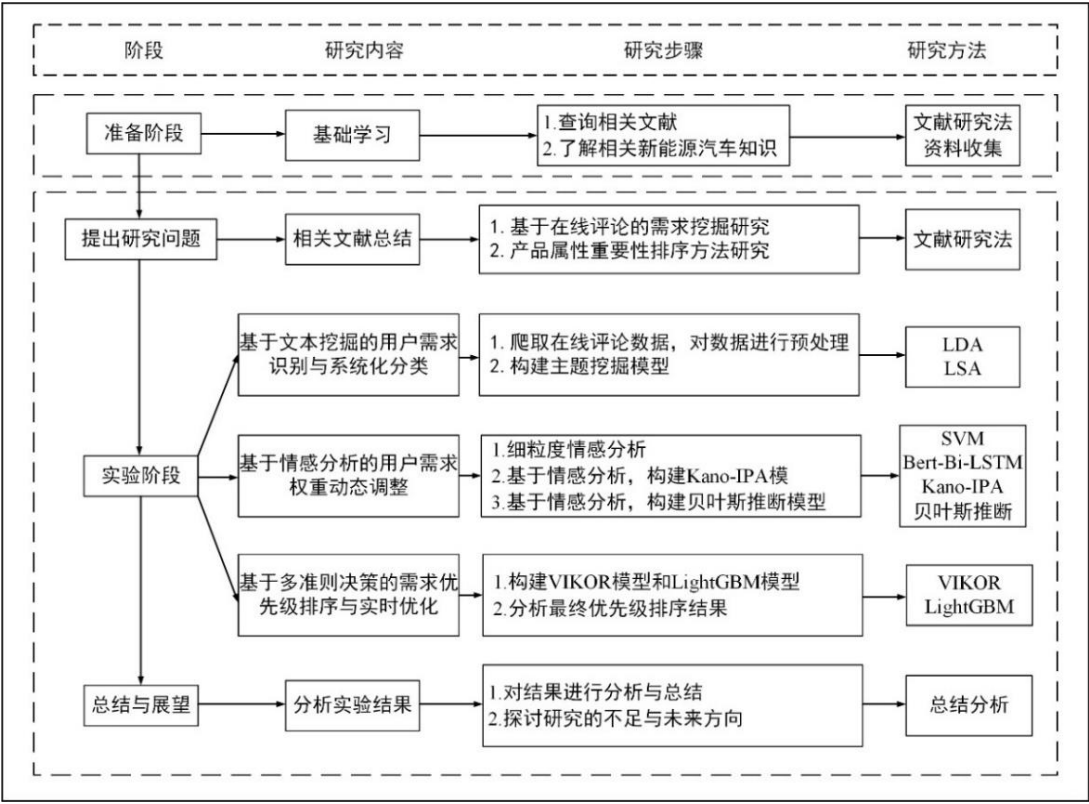


图 1-2 技术路线图

第一，通过广泛阅读新能源汽车用户需求分析相关文献，发现情感因素在用

户需求权重分配中的作用尚未被充分研究。因此，确定以新能源汽车用户评论数据为研究对象，探讨情感因素如何影响用户对不同产品属性的关注权重，并构建情感驱动的需求排序模型。

第二，基于文本挖掘的用户需求识别与系统化分类，采用主题模型对用户评论数据进行主题挖掘，识别新能源汽车用户关注的核心需求属性。通过机器学习与深度学习模型进行情感分析。结合 Kano 模型与 IPA 分析方法，对提取的需求属性进行分类。针对用户评论数据进行情感分析，并结合 Kano 模型和 IPA 方法分析用户需求权重。引入贝叶斯推断方法，分析三种情感主导下的权重占比。

第三，基于上一阶段得到的情感权重，分别采用 VIKOR 和 LightGBM 方法对新能源汽车需求进行排序。VIKOR 方法通过设定不同权重方案计算属性重要度，而 LightGBM 方法基于数据驱动学习情感权重与重要度排序的关系，以探讨两种方法在不同情感主导场景下的适用性。最终，对比两种方法的排序结果，并分析其在产品优化和市场策略制定中的实际价值。

1.5 论文创新性

本文基于在线评论数据，对新能源汽车用户需求情感分析模型进行了深入探索。相较于过往研究，本研究不仅关注用户需求的静态特征，还通过多阶段演化分析揭示需求变化趋势，贡献体现在以下几个方面：

第一，运用自然语言处理方法，建立了基于在线评论数据的用户需求识别模型，并结合机器学习与深度学习方法进行细粒度情感分析。

第二，引入贝叶斯推断方法，定量分析不同情感主导情况下用户对新能源汽车各属性的关注权重。弥补了传统基于词频统计或简单权重设定的不足，使情感分析更具解释性和适应性。

第 2 章 相关理论与方法与文献综述

2.1 相关理论与方法

2.1.1 自然语言处理

自然语言处理（Natural Language Processing, NLP）是人工智能领域中的一个重要研究方向，致力于实现计算机对人类语言的理解、分析和生成。随着互联网数据的爆炸式增长，尤其是文本数据在社交媒体、电商评论和在线论坛等平台中的广泛应用，自然语言处理技术在信息抽取、主题识别、情感分析等方面发挥着越来越重要的作用。

在自然语言处理任务中，主题建模是一种常用的文本分析方法，用于识别文本集合中潜在的主题结构。其中，潜在狄利克雷分配模型和潜在语义分析是经典的主题建模算法，具有良好的可解释性和应用效果。

LDA 是基于概率模型的主题模型算法^[5]。作为一种无监督的机器学习方法，LDA 主题模型可以用来挖掘潜在的主题和对应的主题词，不需要提前进行人工标注。具体模型如图 2-1 所示。

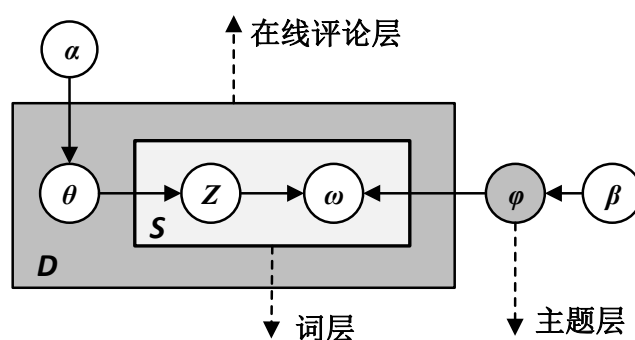


图 2-1 LDA 模型的贝叶斯网络结构

图中 α 和 β 分别为主题分布 θ 和主题词分布 φ 的先验分布参数， Z 和 ω 表示模型生成的主题和最终的主题词， D 表示文档数， S 表示文档中的单词数。主题模型的生成过程如下。

首先从全局中选择长度为 N 的文档，然后使用来自参数 α 的先验分布和参

数 β 的先验分布的样本，来生成文档在主题上的分布 θ 和主题词上的主题分布 φ 。

最后 Z 和 ω 分别使用 θ 和 φ 从多项式分布中采样，模型的联合分布^[6]如方程式所示。

$$p(\omega, Z, \theta, \varphi | \alpha, \beta) = \prod_{n=1}^N p(\theta_m | \alpha) p(Z_{m,n} | \theta_m) p(\varphi | \beta) p(\omega_{m,n} | \varphi_{Z_{m,n}}) \quad (2-1)$$

通常情况下，LDA 模型的超参数 α 和 β 选择默认值，而 z 和 ω 可以通过变分推断算法或吉布斯采样算法获得。LSA^[7]是一种基于统计的计算方法，通过对大规模文本数据的处理来提取和表达单词的上下文。LSA 包括五个步骤，构建词文档矩阵、加权运算、奇异值分解、约简和关联度计算。

文本矩阵构建，LSA 使用文本矩阵来描述出现在多个文本中的单词的数量。单词文档矩阵是一个稀疏矩阵。矩阵的列向量表示文本中的所有单词。矩阵的列向量表示文档的数量，矩阵中的元素表示单词在文档中出现的频率。

权属函数定义，术语频率-逆文档频率^[8]（TF-IDF）是一种广泛应用于信息检索和数据挖掘领域的加权方法。在文本矩阵中，矩阵中的每个元素都表示单词在文档中出现的频率。一个单词出现的频率越高，这个单词对文档的重要性就越大。因此，单词的频率可以用来衡量单词对文档的重要性。对重要性的判断可以从以下两个方面来体现。一方面，一个单词在文本中的频率越高，单词与文本之间的相关性就越高，即单词的权重就越大，索引可以用 TF 表示。另一方面，从所有文本的语料库来看，一个单词的频率越高，该单词与文本的关联程度就越低。因此，可以使用逆文本频率指数来表示一个单词与整个语料库的相关性程度，并记为 IDF。如果一个单词与整个语料库更相关，则 IDF 值越低。如果一个单词相对于整个语料库的相对程度较低，则 IDF 值越高。IDF 的公式可表示如下：

$$IDF = \log \frac{N}{d} \quad (2-2)$$

其中 IDF 为逆文本频率索引，N 为语料库中包含的文本总数， d 表示包含该单词的文本数。综上所述，在考虑一个单词与单词文本的相关性时，有必要考虑 TF 索引和 IDF 索引之间的相互作用。TF-IDF 的规定如下^[9]：

$$TF-IDF = TF \times IDF \quad (2-3)$$

奇异值分解与降维，奇异值分解（SVD）是一种矩阵分解方法。可降低矩阵的维数，奇异值分解公式如下：

$$X = T \times S \times D \quad (2-4)$$

其中 X 矩阵是初始 $m \times n$ 个单词文档矩阵， T 是一个 $m \times r$ 的矩阵， D 是一个 $r \times n$ 阶矩阵， S 是一个 $r \times r$ 的对角矩阵。

相关度计算，奇异值分解可以映射到低维的语义词和文本向量空间，语料库中的所有词和文本都可以用潜在语义空间的坐标来表示，从而两个词之间的关联可以用潜在语义空间的坐标来表示余弦值。余弦值越大，词之间的关联度越大。伪文档由新的输入句子 q 组成，该句子由其单词组成并形成索引向量 X_q 。伪文档的坐标可以表示为

$$D_q = X_q^T \times T \times S^{-1} \quad (2-5)$$

其中 T 和 S 分别表示奇异分解得到的矩阵。 S^{-1} 表示为 S 的逆矩阵。计算检索语句 q 的伪文档坐标 D_q ，迭代的将 D_q 中的所有评论文档与评论集 R 进行比较，并计算两个余弦之间的角度，余弦值越大，词之间的相似度越大^[10]。

情感分析是自然语言处理领域中的一个重要研究方向，旨在识别和提取文本中所包含的主观情感信息。SVM 是一种基于统计学习理论的监督学习方法，广泛应用于分类和回归任务，尤其适用于小样本、高维度、非线性问题的处理。其基本思想是通过构建一个最优超平面，如图 2-2 所示将不同类别的数据点进行最大间隔划分，从而实现分类目标。

在二维空间中，这个超平面就是一条直线，而在三维或更高维空间中则是一个平面或超平面。SVM 模型图中，可以看到数据点被分为两个类别，通常用不同的符号表示。SVM 的目标是找到一个分类超平面，使得正负两个类别的样本被分隔在超平面两侧，而且这个超平面与最近的正负样本点之间的距离最大。这个最大间隔称为最优间隔，而那些正好处在边界附近、与超平面距离最小的点被称为支持向量，它们直接决定了分类边界的位置^[11]。

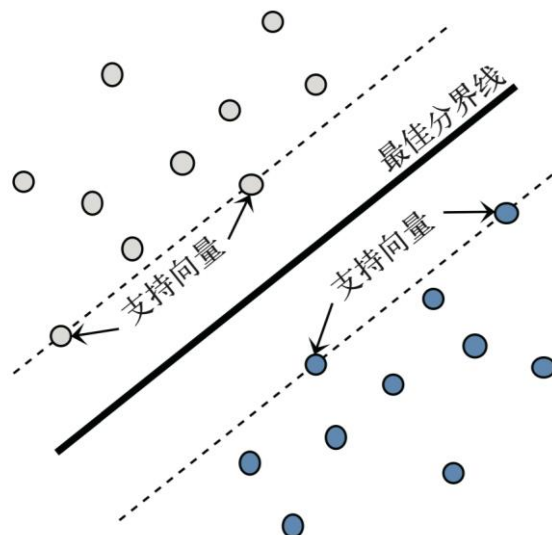


图 2-2 支持向量机模型图

在 SVM 模型图中，通常可以看到三条平行的线：中间那条是最优超平面（记作 H_0 ），它用于对样本进行分类；另外两条线分别代表正类和负类的边界线，分别称为 H_+ 和 H_- 。这两条线距离超平面相等，且正负类样本应该处在各自对应边界线的外侧。模型的优化目标就是在满足所有样本正确分类的前提下，使得这两条边界线之间的间隔最大。SVM 的一个显著特点是，它在分类时只依赖于支持向量，而与远离边界的样本无关，从而提升了模型的泛化能力。

当数据不能被一个线性超平面完全分开时，SVM 还引入了软间隔机制，允许部分样本可以越过边界线，但会通过惩罚项来控制错误的程度。此外，SVM 还可以借助核函数技巧将原本线性不可分的数据映射到高维空间，在高维空间中找到一个可以线性分隔的超平面，再映射回原始空间完成非线性分类。常用的核函数包括线性核、多项式核、高斯径向基函数（RBF）等。

总的来说，SVM 模型图形直观地展示了样本的分类过程、超平面的位置、支持向量的作用以及最大间隔的目标，这种几何直观性也是 SVM 被广泛应用于图像识别、文本分类、生物信息学等领域的重要原因之一^[12]。

Bert 模型^[13]是基于 Transformer 架构开发的预训练模型，广泛应用于自然语言处理任务。其主要流程是首先进行语言模型的预训练，随后根据具体的下游任务（如分类或标注）进行适当调整。

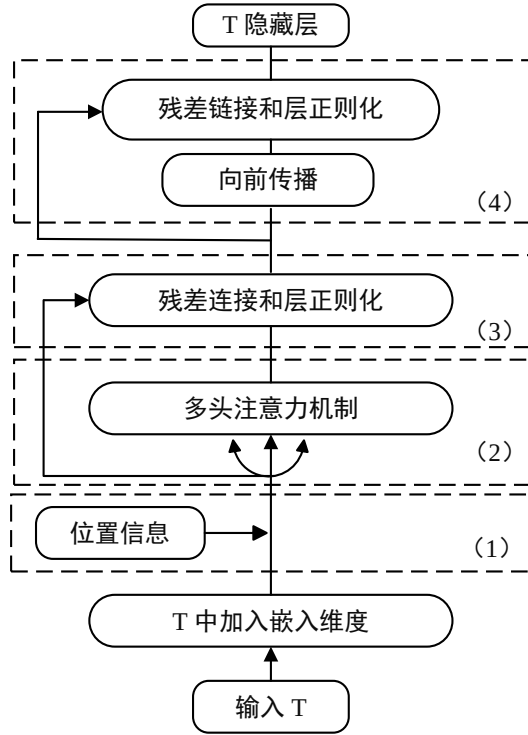


图 2-3 Transformer 模型的编码器

如图 2-3 所示，为 Transformer 模型的编码器部分。第 1 部分 Transformer 模型中没有使用序列的顺序信息，位置编码提供每个字的位置信息给 Transformer，才能识别出语言中的顺序关系^[14]。第 2 部分多头自注意力机制实际上是由 h 个 *self-attention* 并行组成。输入为 $T_{embedding}$ 经过线性变换得到 Q 、 K 、 V ($Q = \text{linear_q}(T_{embedding})$ 、 $K = \text{linear_k}(T_{embedding})$ 、 $V = \text{linear_v}(T_{embedding})$) 后同时经过 h 次的放缩点积 *attention*，得到多头信息。将公式 2-6 结果进行拼接，再通过线性变换得到的值作为多头 *attention* 的输出^[15]。

$$head_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (2-6)$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(head_1, \dots, head_h)W^O \quad (2-7)$$

第 3 部分残差连接，即在上步中注意力机制进行转置使其维度与 $T_{embedding}$ 一致，将两者相加的作用是隐藏层可归一为标准正态分布。在反向传播时，为避免构建 Bert 模型层次太深或出现梯度消失等问题，训练时常用的方法为残差连接。

$$T_{hidden} = \text{Activate}(\text{Linear}(\text{Linear}(T_{attention}))) \quad (2-8)$$

第 4 部分前向传播，即为两层线性映射并用激活函数。

Bert 预训练语言模型因采用双向 Transformer 从而具备获取更长的上下文信息的特征抽取能力^[16]，基于多层 Transformer 网络结构中的 Encoder 搭建的模型结构框架，模型结构如图 2-4 所示^[17]。

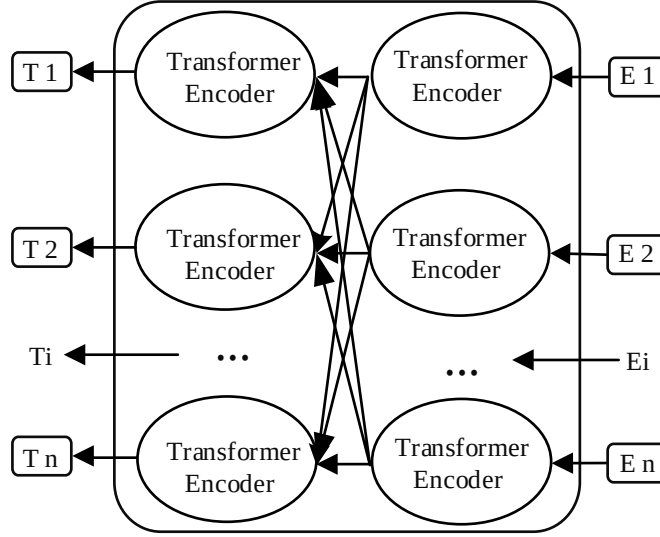


图 2-4 Bert 模型的网络结构

模型由输入标记 $E1, E2, \dots, En$ 开始，这些标记通过嵌入层转换为向量表示，依次输入多个变换器编码器层。每个编码器层通过自注意力机制捕捉输入标记之间的关系，并经过前馈神经网络处理，最终输出上下文嵌入 $T1, T2, \dots, Tn$ 。这些输出表示不仅考虑了各个标记的自身信息，还结合了周围标记的上下文，使得模型能够进行精确的自然语言理解和生成任务。

LSTM 的网络结构^[18]由 t 时刻的输入词 x_t ，状态信息 c_t 临时状态信息 $\tilde{c}_t$ ，隐层状态 h_t ，遗忘门 f_t ，输入门 i_t ，输出门 o_t 组成，模型输入门 i_t 根据 x_t 和 $t-1$ 时刻的隐藏层向量 h_{t-1} 控制当前状态是否可以进入记忆单元，经过计算后产生输出 i_t 、 $\tilde{c}_t$ 和状态信息 c_t ；遗忘门 f_t 根据 x_t 和 h_{t-1} 控制上一时刻的细胞状态有多少内容可以进入当前状态，经过计算后产生输出 f_t ；输出门 o_t 根据 h_{t-1} 、 x_t 和 c_t 控制记忆单元的输，经过计算后产生输出 o_t 当前的隐层状态 h_t 。当文中所有的词向量拼接映射成词向量矩阵 $X = \{x_1, x_2, \dots, x_n\}$ 时，LSTM^[19]按照公式 (2-9) ~ (2-14) 更新迭代运算，神经网络结构如图 2-5 所示。

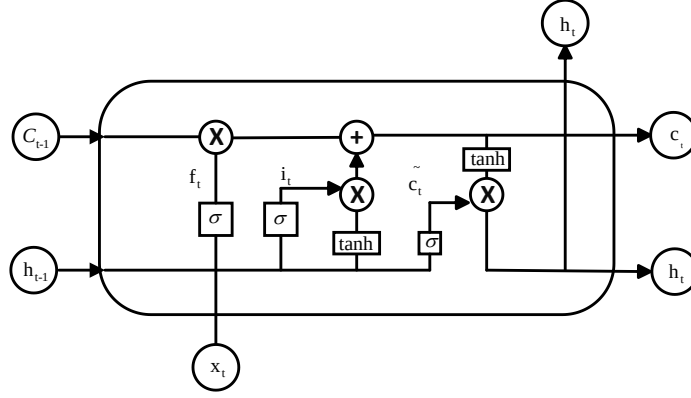


图 2-5 LSTM 神经网络结构

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \quad (2-9)$$

$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o) \quad (2-10)$$

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \quad (2-11)$$

$$\tilde{c}_t = \tanh(w_g[h_{t-1}, x_t] + b_g) \quad (2-12)$$

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \quad (2-13)$$

$$h_t = o_t \times \tanh(c_t) \quad (2-14)$$

述公式中： σ 为 sigmoid 函数； w_f 、 w_i 、 w_g 、 w_o 为模型的参数； b_f 、 b_i 、 b_g 、 b_o 为偏置项；所定义门向量 i_t 、 f_t 、 o_t 取值范围为(0,1)。BiLSTM 通过^[20]结合正向 LSTM 和反向 LSTM 所获得双向的语义信息， t 时刻 BiLSTM 输出的隐藏状态 h_t 可以表示为：

$$h_t = \vec{h}_t \oplus \overleftarrow{h}_t \quad (2-15)$$

上述公式中， $\vec{h}_t$ 为正向 LSTM 输出的隐藏状态； $\overleftarrow{h}_t$ 为反向 LSTM 输出的隐藏状态。

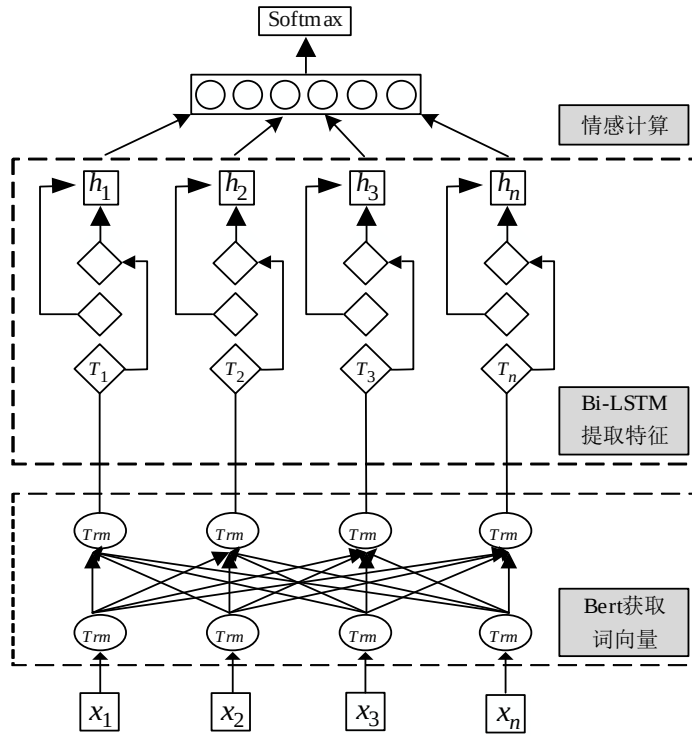


图 2-6 Bert-Bi-LSTM 模型结构

基于 Bert-Bi-LSTM 的新能源汽车用户在线评论的情感分析模型共包含四层^[21]，模型结构如图 2-6 所示。首先，使用 Bert 模型将文本数据转化为向量。其次，使用 Bi-LSTM 提取特征。最后，将 Bi-LSTM 输出的特征向量作为全连接层的输入，通过该层的输出，并应用 Softmax 函数，得到一个向量表示该文本的情感倾向值^[22]。

评估自然语言处理中情感分析模型的表现^[23]，通常会使用准确率 (Accuracy)、精确率 (Precision)、召回率 (Recall)、F1 值等指标。

准确率是最直观的评价指标，表示模型正确预测的样本占总样本的比例。准确率越高，表明模型的总体预测效果越好。计算公式如下。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2-16)$$

精确率是指模型预测为正类的样本中，真正正类的比例。精确率越高，表明模型在预测为正类时的正确性越高。计算公式如下。

$$Precision = \frac{TP}{TP + FP} \quad (2-17)$$

召回率是指实际为正类的样本中，模型正确预测为正类的比例。召回率越高，表示模型能够识别更多的正类样本。计算公式如下。

$$Recall = \frac{TP}{TP + FN} \quad (2-18)$$

F1 值是精确率和召回率的调和平均数，用于综合考虑精确率和召回率。F1 值较高表明模型在准确预测正类和识别正类方面都做得较好。F1 值的范围是 0 到 1，值越高，模型的性能越好。计算公式如下。

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2-19)$$

其中， TP 为真正例，模型正确预测为正类的样本。 TN 为真负例，模型正确预测为负类的样本。 FP 为假正例，模型错误预测为正类的负类样本。 FN 为：假负例，模型错误预测为负类的正类样本。

2.1.2 分类、决策与排序

在产品属性评价与用户需求分析过程中，仅依赖情感分析或主题挖掘还不足以全面刻画用户的关注重点与满意度感知。为了进一步实现对产品属性的系统分类、精准决策与排序，有必要引入更具解释力的理论模型进行分析。

Kano 模型是由日本贾南教授和狩野纪在 1980 年代提出的，它阐明了产品属性与顾客满意度之间的非对称和非线性关系^[24]。传统的 Kano 模型存在一些弊端，如 Kano 模型在进行属性分类时并没有考虑到不同属性排序，某些属性可能对用户满意度的影响更大，而其他属性可能只是次要因素^[25]。

IPA 是一种结构性方法，它通过考虑用户感知到的重要性和性能来产生特定的决策想法^[26]。该方法将每个属性的重要性和性能值放入二维坐标平面，并将其分为四个象限，以帮助定量评估产品或服务的性能和重要性。

在面对用户需求分类与属性排序问题时，除了借助 Kano 模型和 IPA 模型等多维评价工具外，还需要在决策过程中引入更具逻辑推理能力和不确定性处理能力的理论方法。

贝叶斯推断起源于 18 世纪由托马斯·贝叶斯提出的贝叶斯定理，是统计学中的一种重要方法。它强调在统计推断中利用先验分布与观测数据相结合，通过不断更新后验分布来反映对未知参数的最新认识^[27]。

贝叶斯推断的核心是贝叶斯定理，其数学表达式为：

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \quad (2-20)$$

$P(\theta)$ 为参数 θ 的先验分布，代表在获得新数据之前对参数的初步认识； $P(D|\theta)$ 为似然函数，表示在参数 θ 条件下，观察到数据 D 的概率； $P(\theta|D)$ 为后验分布，即在观测数据 D 后对参数 θ 的更新认识； $P(D)$ 为数据的边际似然，是归一化常数。

在实际应用中，首先为每个需求属性设定一个合理的先验分布^[28]，然后根据用户在线评论中的情感数据构建似然函数，利用贝叶斯定理计算出后验分布。通过不断更新后验分布，可以动态调整各属性的权重，从而精准刻画情感波动对用户偏好的影响。这种方法具有很强的适应性和实时性，能够让新数据的到来逐步完善对系统状态的估计^[29]。

在用户需求属性的多维排序分析中，除了基于感知与认知维度的 Kano 模型、IPA 模型与贝叶斯推断方法外，进一步引入具有更强排序性能与多属性综合评价能力的模型，可以提升排序结果的科学性与实用性。

LightGBM 是由微软开发的一种高效梯度提升决策树框架。它在传统梯度提升方法的基础上，通过直方图算法和叶子生长策略的改进，实现了对大规模数据的快速处理和高效训练，已广泛应用于分类、回归以及排序等任务^[30]。

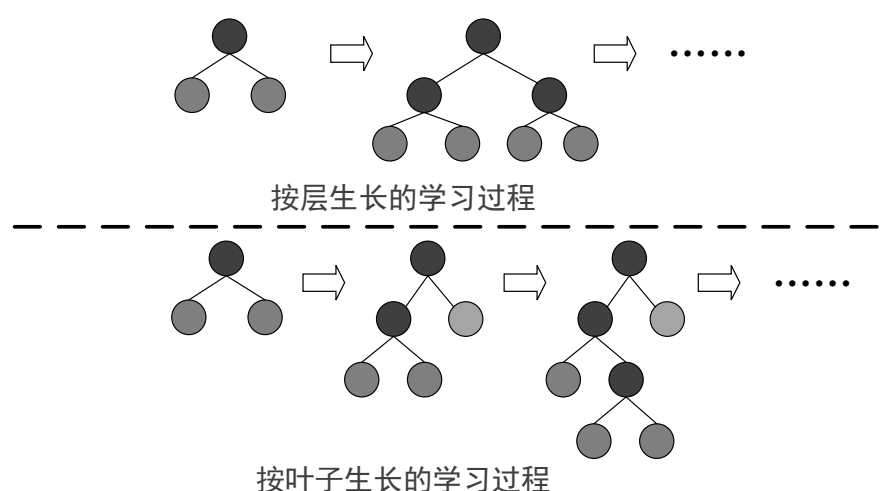


图 2-7 Leaf-wise 的决策树生长策略

LightGBM 的核心思想是构建一组决策树，如图 2-7 所示，每棵树通过对前一棵树的残差进行拟合来不断降低预测误差。其关键技术包括^[31]直方图算法，将连续特征离散化为有限的区间（直方图），大幅减少计算量和内存占用，同时加速数据分割过程如图 2-8 所示。基于叶子的生长策略，不同于传统按层生长的方

法，LightGBM 采用基于叶子的最佳优先生长策略，从而更好地捕捉数据的非线性特征。排序目标函数：在排序任务中，LightGBM 可结合 LambdaRank 等排序目标函数，优化排序性能。通过构建对数似然损失或其他排序损失函数，模型能够直接优化用户需求属性排序的准确性^[32]。

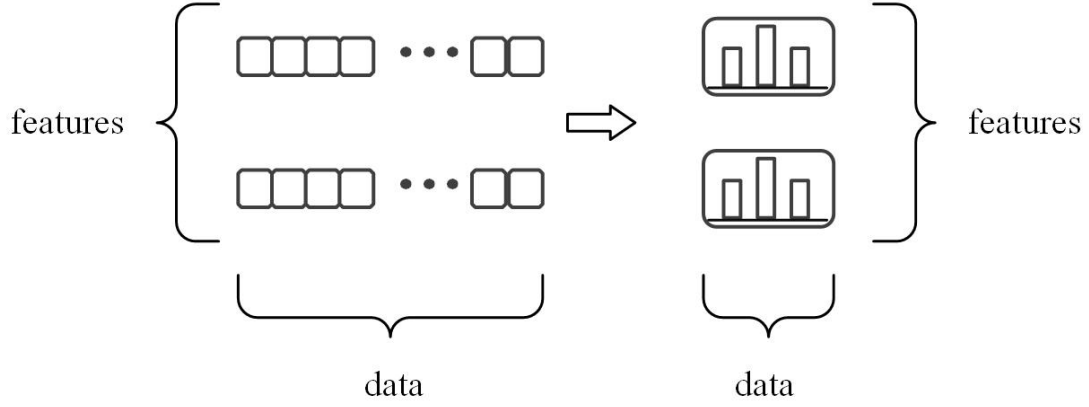


图 2-8 直方图算法示意图

VIKOR^[33]模型是一种多属性决策方法，旨在通过确定正理想解和负理想解来帮助决策者在存在多个相互冲突的标准时做出选择，其优势在于^[34]，它能够有效处理多个冲突的评价标准，通过提供一个综合评分来帮助决策者在不同方案之间进行比较。并且，VIKOR 方法能够识别和优先考虑接近理想解的方案，这对决策者在选择最佳方案时尤为重要。此外，该方法在不确定性和模糊性环境中表现出色，能够通过权重调整和排名方法灵活适应不同情境。VIKOR 方法主要包括以下四个步骤^[35]。

第一步，确定各评估值中的正理想解 f_j^* 和负理想解 f_j^- 其中，正理想解指的是第 j 个属性在各评估属性中的最大值，负理想解指的是第 j 个属性在各评估商品中的最小值，计算公式为：

$$f_j^* = \max_i f_{ij} \quad (2-21)$$

$$f_j^- = \min_i f_{ij} \quad (2-22)$$

第二步，计算每个评估属性的 S_i 值和 R_i 值， S_i 值代表群体最大效用， R_i 值代表个体最小遗憾。其中 $i=1,2,3,\dots,m$ ，计算公式为：

$$S_i = \sum_{j=1}^n \frac{w_j(f_j^* - f_{ij})}{f_j^* - f_j^-} \quad (2-23)$$

$$R_i = \max \left[\frac{w_j(f_j^* - f_{ij})}{f_j^* - f_{ij}} \right] \quad (2-24)$$

其中， w_j 为第 j 个属性的权重， $j=1,2,3,...,n$ ， n 为属性数量； $\frac{w_j(f_j^* - f_{ij})}{f_j^* - f_{ij}}$

为评估值到理想解的加权距离； S_i 值代表第 i 个评估属性的评估值到正理想解的加权距离， R_i 值代表第 i 个评估属性的评估值到负理想解的加权距离^[36]。

第三步，计算 Q_i 值，其中 $i=1,2,3,...,m$ ，计算公式为：

$$Q_i = v \left[\frac{S_j - S^*}{S^- - S^*} \right] + (1-v) \left[\frac{R_j - R^*}{R^- - R^*} \right] \quad (2-25)$$

其中， Q_i 表示第 i 个评估产品的 VIKOR 值，以代表产品的用户满意度， $S^* = \min_i S_i, S^- = \max_i S_i, R^* = \min_i R_i, R^- = \max_i R_i$ 。 S^* 为群体最大效用， R^* 为最小遗憾， v 为群体最大效用权重， $(1-v)$ 是个体遗憾权重。 v 可以在 0 到 1 范围内调整，这表示当 v 大于 0.5 时， Q_i 指数代表了从众原则。（在消费品满意度研究中一般取值 0.5）

第四步，在不同取值 v 下计算 Q_i 值，从而对分析属性进行排序。

2.2 文献综述

2.2.1 文献检索情况

2.2.1.1 相关文献来源

数据来源主要是 Web of Science (WOS) 和中国知网 (CNKI) 两大数据库。选取 WOS 的核心数据集并设置，以“Natural Language Processing”OR“Online Reviews”为话题检索，文件类型设定为论文、综述论文、会议论文。通过检索筛选获得符合主题的文章 39825 篇。选取 CNKI 数据库并设置，以自然语言处理为

主题进行检索筛选符合主题的文章 3161 篇。本节将选择 VOSviewer 软件对国内外文献进行关键词共现网络分析，分析自然语言处理的研究现状。

2.2.1.2 文献检索情况分析

关键词作为论文的高度概括，是掌握文章研究方向及研究热点的关键。本节通过 VOSviewer 软件对知网、Web of Science 的相关文献进行关键词聚类绘制了关键词分布的知识图谱，如图 2-9 和 2-10 所示。

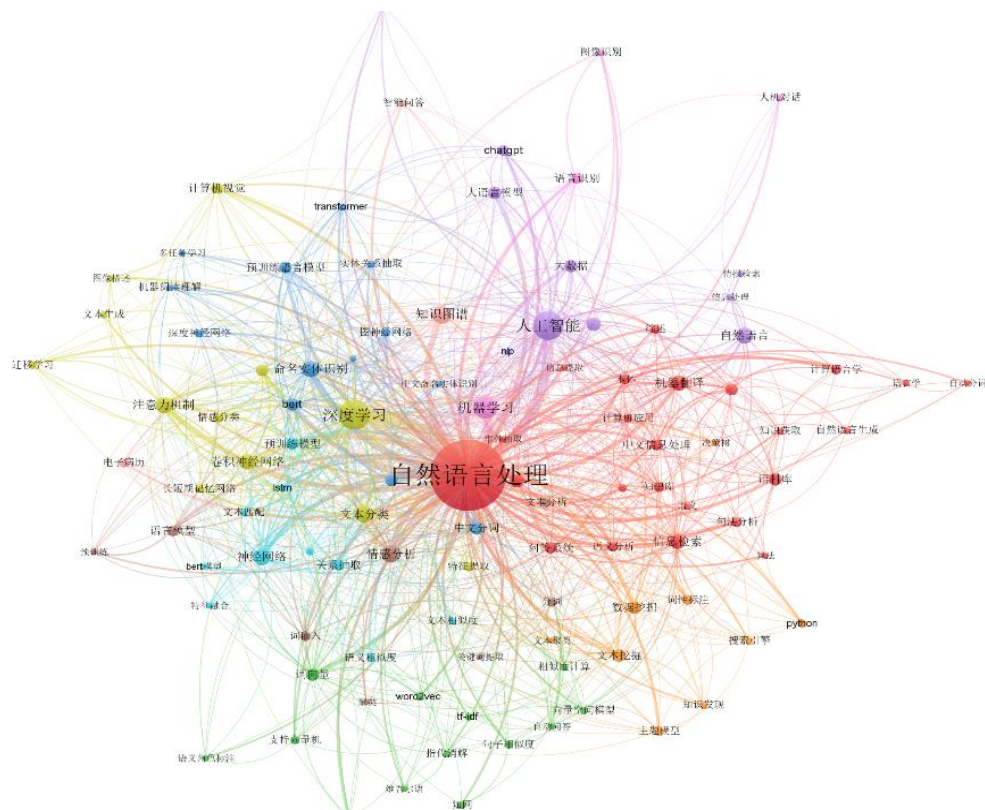


图 2-9 关键词共现图（CNKI）

当前 NLP 领域的研究主要集中在深度学习与大模型（LLM）研究、NLP 在文本分析和知识获取中的应用、NLP 与跨学科应用三个核心方向。深度学习与大模型研究方向主要围绕 Transformer 架构、BERT、GPT、大语言模型（LLM）及其微调方法展开，该方向的关键词节点较大、连线密集，说明学术界对此关注度极高，同时不同概念之间的相互引用和研究联系紧密。近年来，NLP 研究正由传统机器学习向深度学习主导的 LLM 方向转变，特别是大规模预训练模型的应用与优化成为研究重点。文本分析和知识获取方向涉及文本分类、情感分析、知识图谱构建、命名实体识别和关系抽取等，该方向的关键词形成多个紧密连接的小

型集群，显示研究较为细分，存在多个具体应用场景，如舆情分析、智能客服、自动摘要、推荐系统等。

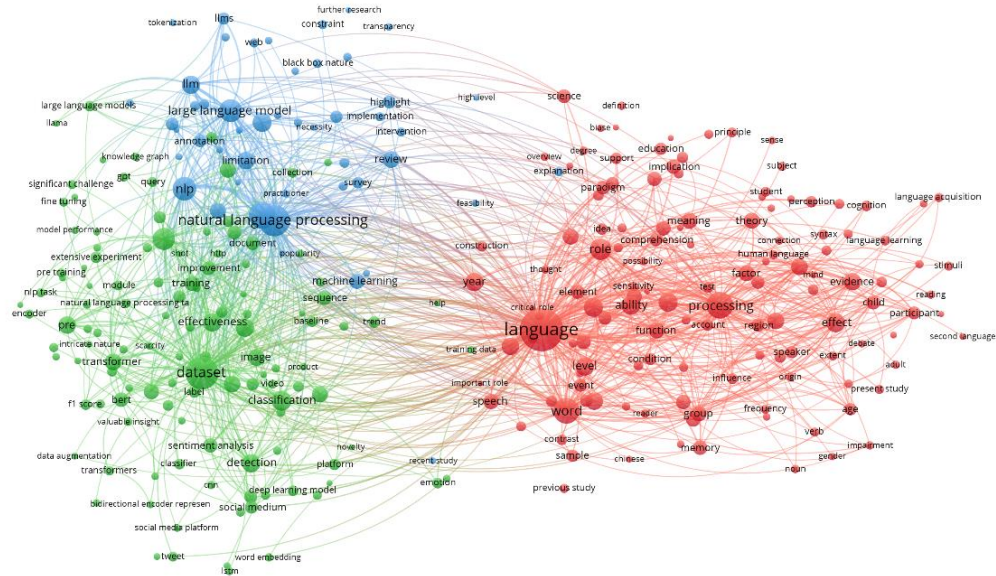


图 2-10 关键词共现图（WOS）

NLP 与跨学科应用则涉及语言学、认知科学、社会科学等交叉研究，如语言学习、语义理解、人机交互等，该方向的关键词分布较为分散，但与核心 NLP 研究领域仍保持较强的关联性，近年来 NLP 在智能教育、翻译、心理学建模和社交媒体分析等方面的应用不断增加，说明研究正朝着多学科融合方向发展。从关键词的节点大小和连接密度来看，Transformer 及 LLM、知识图谱与信息抽取、情感分析与文本挖掘是当前学术界最为关注的核心研究领域，其中 Transformer 及 LLM 相关研究形成最大的节点和最密集的连接，表明当前 NLP 研究的核心仍然是基于 Transformer 架构的大模型训练与优化。

2.2.2 产品开发需求识别与分析

近年来，电子商务、社交媒体等信息技术的快速发展，越来越多的消费者愿意通过在线评论表达自己的消费感受^[37]。这些在线评论不仅直接影响消费者的购物选择，也为制造企业提供了宝贵的洞察，帮助其了解消费者的需求和满意度^[38]。因此，如何从海量在线评论中挖掘出影响消费者满意度的因素，并对这些因素进行合理排序，成为了一个值得深入研究的问题^[39]。

学术界对自然语言处理的各类研究也在不断深入。现有的以在线评论为对象的研究,主要包括^[40]基于评论内容的文本挖掘技术和文本挖掘应用两个方面。技术方面, Blei 等^[41]介绍了潜在狄利克雷分配模型,验证了 LDA 在文档建模、文本分类和协同过滤任务中的表现优于单词模型混合和概率性潜在语义分析模型。Wu^[42]等基于 TF-IDF、Word2Vec、神经网络等方法,提出了一个挖掘电影评论典型意见的通用框架,从挖掘出的信息中发现了问题,为电影的制作提供了一定的参考。刘伟和徐鹏涛^[43]构建了基于 IAM 模型的双路径分析理论模型,发现 O2O 电商平台在线点评有用性的影响因素,包括信息丰富性、信息可读性及点评负面性。这类研究主要聚焦于情感分析技术、信息提取和文本分类技术。其中涉及到的方法有文档频率方法、情感词典法、朴素贝叶斯、支持向量机等。评论内容的文本挖掘技术在科研中具有广泛的应用,为研究者提供了从海量文本数据中获取有价值信息和洞察的有效手段,有助于科学研究的深入和实证分析的开展。MaasAL^[44]等人结合无监督和有监督方法的模型,利用文档级的情感极性标注(如在线评论的星级评分)来增强词向量的情感表达能力。实验结果表明,该模型在情感分类任务上优于之前提出的多种方法。此外,研究还引入了一个大规模的电影评论数据集,以作为该领域更稳健的基准测试集。

随着大语言模型的发展^[45],其在自然语言处理(NLP)任务中展现的能力引发了广泛关注。朱丹浩等^[46]研究了面向古文自然语言处理生成任务的大语言模型评测,构建了基于语言和知识两个维度的评测任务,并对 13 个主流大语言模型进行评测。结果显示,ERNIE-Bot 在古籍知识方面表现突出,GPT-4 在语言能力上表现最佳,而 ChatGLM 系列开源模型表现优异。文章为古籍领域的大语言模型评测提供了标准,并为模型训练过程中的基座模型选择提供了依据。陈炫婷等^[47]研究了 GPT-3 和 GPT-3.5 模型在多种自然语言处理任务中的鲁棒性,发现其在分类任务中表现良好,但在信息抽取任务中存在严重混淆和幻觉现象。尽管模型性能在迭代中有所提升,但鲁棒性仍有待加强,特别是在处理复杂任务和文本变换时。Qin 等人^[48]评估了 ChatGPT 在 20 个流行的自然语言处理数据集上的零-shot 学习能力,涵盖了 7 个代表性的任务类别。研究表明,ChatGPT 在需要推理能力的任务(如算术推理)上表现良好,但在处理特定任务(如序列标注)时仍然面临挑战。

文本挖掘应用方面，大多数研究集中在消费者细分、产品推荐和产品改进这三个方面^[49]。例如，Wang^[50]等提出基于产品评论数据挖掘的个性化推荐算法，可以根据生成的历史数据对每个用户进行推荐，优化用户体验，提高消费者意愿。Fan^[51]等结合了朴素贝叶斯算法和 Bass/Norton 模型，利用历史销售数据和在线评论数据进行产品销售预测。这部分研究主要聚焦于^[52]实现产品的个性化推荐；对具有竞争关系的产品进行比较观点挖掘，得到产品的竞争优势和不足；市场分析；基于在线评论实现销量的预测。还有学者致力于研究在线评论数据的相关因素对消费者购买行为的影响，如杜学美^[53]等以在线评论为研究对象，构建了纳入接受者专业能力这一调节因素的在线评论对消费者购买意愿的影响模型，发现在线评论的数量、质量、效价等，以及接收者心里表征类因素，都正向影响在线评论对消费者购买意愿的影响。

2.2.3 基于在线评论的情感分析

情感分析是自然语言处理领域的重要研究方向，旨在从文本中识别和提取主观情感信息。当前已广泛应用于产品评价、市场调研、股票预测、医疗反馈、旅游体验等多个领域，其研究热点不断向多层次（如文档级、句子级、短语级和方面级）分析方向发展，逐步实现情感信息的细粒度挖掘^[54]。与此同时，研究者们也在不断探索更加精细的建模方法，如基于词嵌入的深度学习技术和多模态情感识别，以提高情感分析的准确性与泛化能力。

在情感分析方法方面，当前主流技术可分为词典方法、机器学习方法和混合方法。词典方法基于情感词汇库，通过对文本中词汇的情感倾向进行打分与聚合，来判断整体情感极性^[55]。此类方法无需训练数据，适用于标注资源匮乏的应用场景，常见词典包括 SentiWordNet、Bing Liu 情感词典等。如 Volkova 等^[56]构建适用于社交媒体语境的 Twitter 特定情感词典，基于高精度的通用情感词典，通过自举方法结合大量未标注 Twitter 数据和少量有标注数据，逐步扩展出多语言情感词汇。Weichselbraun^[57]等提出了一种创建情感词典的半自动方法，该方法通过将情感值分配给情感术语，引入了一个对未标记的域文档进行作的引导过程，以扩展创建的词典，并根据特定用例对其进行自定义。李寿山等^[58]利用英文种子词典，借助机器翻译把原评论和对应的翻译评论作为一篇文档，计算其他词与种子

词的 PMI；然后利用词之间的 PMI 值，构建连接矩阵，借助标签传播算法将英文的情感词极性传播到中文词上，克服中文在现有资源上的一些劣势，从而构建情感词典。

然而，词典方法常面临领域依赖性强、情感词语多义性和上下文适应性差等问题。机器学习方法则利用统计学习模型对文本进行建模，常用算法包括朴素贝叶斯、支持向量机、最大熵模型等。随着深度学习技术的快速发展，卷积神经网络、Bi-LSTM 以及 Transformer 结构下的 BertT 等模型被广泛用于情感分析任务，在特征学习和上下文建模方面表现出更高的精度和鲁棒性。如梁军等^[59]提出了一种基于 RAE 的情感极性转移模型，在不增加神经网络复杂度的前提下，提高了模型对中文情感分析的准确性，算法的表现已经很接近当前采用许多手工标注特征的传统算法的性能。Mrabti^[60]等提出一种基于几何思想的可解释性情感分析模型，结合最大间隔分类和两种改进算法，在提升模型透明度的同时，在评论数据集上取得优异性能，优于十种主流分类算法。吴伟等^[61]通过构建情绪指数并应用多种机器学习方法，系统分析了投资者情绪对中国原油期货价格的影响，发现 LSTM 模型结合情绪指数在价格预测中表现最优，揭示了情绪与市场波动之间的显著联系。

尽管技术手段不断丰富和优化，情感分析依然面临多方面的挑战^[62]。首先，非结构化文本中存在大量语言模糊现象，如俚语、缩写、拼写错误等，增加了文本预处理难度。其次，反语、讽刺与幽默等隐性语言表达极易误导情感判断，是当前模型识别能力的瓶颈^[63]。此外，词语的情感极性在不同语境下会发生变化，多义词识别和上下文情感归因成为重要问题^[64]。跨领域和跨语言迁移也是研究的难点，模型在训练后难以在目标域中保持性能稳定。与此同时，情感强度的量化、多情绪共现的分析、以及中立情感的识别也逐渐成为研究热点。为应对上述挑战，研究者们正在探索更精细化的情感分析技术^[65]，如方面级情感分析、情感表示学习、情感极性细化分级、模糊集理论、多模型融合、注意力机制与多模态情感分析等方法，进一步推动情感计算技术的发展与落地应用。

2.2.4 产品属性重要度排序方法

在现代市场中，由于消费者需求的多样化和个性化，如何有效地识别和排序

不同需求属性，成为产品设计和营销策略制定中的关键问题。近年来，数据分析技术的进步，尤其是自然语言处理和情感分析的应用，研究者开始探索基于用户评论、行为数据等信息进行需求属性的排序。这些方法能够从大量非结构化数据中提取有价值的信息，为企业提供科学的需求排序和决策支持。本部分将综述当前在需求排序领域的主要研究成果，重点讨论基于文本数据分析、情感分析及其他辅助技术的需求排序模型及其应用。

需求排序是指^[66]，通过相应的方法确保用户关注度高的需求优先实现。目前相关研究主要有^[67]采用经典模型或自行构建评价体系来洞察用户需求以及针对已有方法进行改进，构建新的研究模型这两类研究。如吴江^[68]等构建了游客满意度分析框架，采用无监督方法抽取游客对于景区的细粒度属性观点，并基于 SnowNLP 和 XGBoost 分别评估游客对不同属性的感知情感及感知重要性，进而采用 IPA 模型对景区属性的满意度进行分析。Mohammed^[69]等基于 ANN 和特征融合技术，根据回收需求对生成的垃圾进行自动分类，改善了智慧城市中的垃圾分类服务和回收决策。Yaseen^[70]等采用迭代模型，将需求按照优先级进行分阶段的排序，减少项目的估计时间，并确保项目按时交付。Liang 等^[71]提出了一种基于在线评论的酒店选择定量方法，通过分析评论中的情感词汇，将游客的情感偏好转化为分布语言格式，表示为情感等级，确定分布语言评价中的理想解和最差解，基于酒店评价中词汇的频率和分布语言评价，确定了评价特征的权重向量。进一步，提出了一种新的 DL-VIKOR 方法，用于对酒店进行排序和选择。李少波等^[72]提出一种产品在线评论数据驱动的方法。将词频与评估、强度和活动 3 个维度相结合的 TFEPA 方法，并使用该方法提取在线评论中的感性词，采用面向在线评论的词聚类结合程度副词的方法计算感性评价值，再从认知心理学的角度，结合产品属性参数与用户感性意象，构建了基于 BP 神经网络的非线性映射模型，用于模拟用户心理评估机制。最后，评估了模型的泛化能力，验证了所提方法的可行性与有效性。冯坤和杨强等^[73]提出了一种基于在线评论和随机占优准则的顾客满意度测评方法，利用 LDA 模型提取主题、情感分析计算满意度得分，并通过随机占优准则和 PROMETHEE-II 方法对影响因素排序，最终通过商品评论数据，验证了方法的有效性。史丽丽^[74]提出了一种混合神经网络（HNN）模型用于从中文在线评论中提取产品特征，并结合 Kano 模型对特征进行分类和排序，

提升了特征提取的精度，并为产品优化提供支持。

2.2.5 文献评述

在当前的研究中，自然语言处理和情感分析技术的不断发展，学者们已经开始探讨基于在线评论数据挖掘用户需求及情感倾向的各类方法^[75]。这些方法不仅有助于从大量的非结构化数据中提取有价值的信息，还为产品设计和营销策略的优化提供了重要支持。通过文本挖掘和情感分析，能够有效识别消费者对产品特性或服务的关注点，并基于此对需求属性进行排序，从而实现市场定位和产品优化。

然而，尽管现有研究已经取得了一些成果，但在消费者需求识别和排序方面，仍然存在不少挑战。特别是在新能源汽车领域，如何基于消费者的在线评论、情感分析结果及多准则决策方法构建一个合理的需求排序模型，依然是一个有待解决的问题。现有研究多集中在单一情感维度的分析，忽视了情感的细粒度特性对消费者需求的潜在影响。因此，如何结合情感极性与用户偏好的关系，特别是在细粒度情感驱动下对需求进行排序，将为新能源汽车产品优化提供重要的理论依据和实践指导。

在此背景下，通过分析新能源汽车用户评论数据，结合情感分析与多准则决策方法，旨在探讨不同情感极性对需求偏好的影响，并构建一个情感驱动的需求排序模型。该模型将帮助更好地理解用户的多元需求，推动新能源汽车产品的持续优化，进而为企业制定精准的营销策略和产品改进方案提供依据。

第 3 章 基于在线评论的 NEV 用户需求识别与获取

在新能源汽车市场，消费者的购车决策与使用体验往往伴随着丰富的主观感受和 demand 倾向，而在线评论作为用户自主表达的平台，能够真实反映他们的使用体验、关注点及潜在需求。与传统问卷调查相比，在线评论数据具有广泛的来源、实时更新的特点，并能够涵盖不同地域、消费层次和个性化偏好的用户反馈，从而减少样本偏倚问题。如何从这些海量且多样化的评论数据中精准挖掘用户需求，识别其对新能源汽车不同属性的关注重点，是本章研究的核心。

3.1 数据采集与清洗

通过 Python 编写定向爬虫程序，从汽车之家平台抓取新能源汽车用户在线评论数据，覆盖多种主流品牌及其热门车型，最终获取有效评论共计三十多万条。在数据爬取过程中，首先通过浏览器开发者工具分析汽车之家评论页面的结构，发现其评论内容采用 Ajax 动态加载的方式，传统的静态爬取方法无法直接获取完整数据^[76]。

编号	裸车购买价	购车地点	动力类型	油耗	电耗	评论时间
1	7.98万	定西	插电式混动	3.4	13.2	2024/4/24
2	11.58万	日照	插电式混动	0	16.2	2024/4/23
3	7.98万	重庆	插电式混动	8	15	2024/4/23
4	7.98万	德州	插电式混动	0	16	2024/4/23
5	12.58万	东莞	插电式混动	5	14.8	2024/4/23

编号	评论内容
1	首先我需要说明一点，我呢不是很懂车，买这台车呢主要是因为我父亲比较偏向于新能源，我哥呢也比较关注比亚迪。
2	我个人觉得混动挺实用的，又能省钱，跑长途也没问题。
3	车漆有点薄，遇到一点不好的路况，小石子溅起来很容易刮花。
4	车里的隔音效果有点差，在高速开车的情况下有明显的风噪、胎噪音。
5	秦plus的内饰让我很满意，因为我很喜欢这种设计感。而且车内的颜色搭配也很时尚大气，用料上也比较高级，整体看起来档次感十足。

图 3-1 部分新能源汽车用户在线评论数据

采用动态爬取策略，利用 Selenium 模拟用户操作，使页面加载完整后再提取评论内容。同时，为提高爬取效率和数据质量，结合 Scrapy 框架进行大规模数据抓取，使得爬虫能够并行处理多个评论页面。此外，由于汽车之家设有一定

的反爬机制，例如频繁请求可能导致 IP 封禁，采取了多种策略来规避，如模拟正常用户浏览行为、使用随机 User-Agent 头部信息、设置合理的请求间隔以及采用代理 IP 池进行轮换访问。这些措施有效提高了数据爬取的稳定性和成功率。爬取到的评论数据包含用户评论内容、发布时间、用户评分、车型信息等关键字段，并保存到 Excel 表格中，以便后续分析处理，部分评论数据如图 3-1 所示。在数据爬取完成后，对原始新能源汽车评论数据进行了系统化的清洗，以提高数据质量，减少噪声对分析结果的影响。数据清洗主要包括。

第一，由于在线评论数据可能存在用户重复发布、爬取过程中数据重复存储等情况，首先对数据进行了去重处理。具体而言，基于评论内容计算哈希值，并对哈希值相同的评论进行去重，以剔除完全重复的文本。此外，为避免因用户表达习惯导致的轻微文本变动（如标点、空格）而未能识别重复评论，还计算了评论文本的相似度，并对相似度高于设定阈值的评论进行合并或筛除，从而有效减少冗余数据对分析结果的干扰^[77]。

第二，在线评论数据通常包含大量非用户真实需求的内容，如广告、营销话术、无意义字符、表情符号等^[78]。采用基于正则表达式的方法清除常见的广告文本，同时过滤无意义字符。此外，针对评论中出现的特殊符号，基于表情库进行替换或剔除，以保证文本格式的统一性。

第三，在数据采集过程中，部分评论可能缺失关键信息，如车型、用户评分等。对于这些缺失值，采取不同策略进行处理。对于缺失车型信息的评论，基于评论所属页面的 URL 路径提取相关车型信息，并进行填充。对于缺失评分的数据，如果评论包含明确的情感倾向，可以结合文本情感分析模型推测其可能的评分范围；如果无法推测，则采用相邻数据插值法，根据相同用户或相似评论的评分值进行填补，确保数据的完整性。对于短文本或空评论，如果评论字数少于设定阈值且无关键信息，则直接删除，以减少无效数据对分析的影响。

3.2 文本标准化处理

在数据清洗完成后，需要对中文文本数据进行标准化预处理，以适配后续的主题建模分析，部分预处理数据如图 3-2 所示。主要包括分词、去停用词、词干化与标准化等步骤^[79]。

（1）分词与词性标注

由于中文文本中单词之间没有空格分隔，直接输入主题建模算法可能会导致分析结果失真，因此需要进行分词处理^[80]。本研究采用 Jieba 分词工具，并结合新能源汽车领域的专业术语词典，以提高分词的准确性。其默认词典可能无法识别新能源汽车领域的专有名词，因此自定义添加了行业术语，如续航里程、智能辅助驾驶、快充慢充和能量回收等，以确保这些词不会被错误切分。为了提升主题模型的分析效果，进一步对分词结果进行词性标注，并剔除非核心词汇。

（2）停用词过滤

采用哈工大停用词表进行停用词过滤，以剔除对文本主题分析无贡献的虚词、助词等。针对新能源汽车用户评论的特点，扩展了自定义停用词库。例如，车型和购买优惠等高频通用词往往不会影响对用户需求的识别，因此也被剔除，以减少干扰^[81]。

（3）词干化与标准化

由于用户在评论中可能使用不同的词语来表达相同的概念，为避免语义重复，进行了以下标准化处理^[82]。同义词合并，例如电耗与能耗具有相同含义，因此在分析过程中将其统一替换为能耗。智能驾驶与自动驾驶也被归为同一类别，以减少冗余主题的产生。简繁体转换，由于部分用户可能使用繁体字输入（如駕駛和電池），统一采用 OpenCC 进行简繁转换，以保证数据的一致性。

评论文本预处理
说明、懂车、买车、父亲、新能源、关注、比亚迪
混动车、实用、省钱、长途、无问题。
车漆、薄、路况、小石子、刮花
隔音、效果差、高速、风噪、胎噪
秦plus、内饰、满意、设计感、颜色搭配、时尚、大气、用料、高级、档次感

图 3-2 部分预处理数据

3.3 主题提取建模分析

为了提高数据分析的准确性，将爬取的新能源汽车用户评论数据进行分组处

理^[83]。分组不仅有助于对比不同时期用户需求的变化趋势，还能更清晰地揭示新能源汽车市场发展的演变，以及用户关注重点的动态调整^[84]。

在方法选择上，采用了潜在狄利克雷分配与潜在语义分析来识别隐含需求。为科学评估潜在狄利克雷分配模型对新能源汽车用户评论数据的表征能力，采用主题困惑度作为核心评价指标^[85]。结合 LDA 与 LSA 两种主题提取方法，旨在发挥二者在主题建模中的互补优势。LSA 依托奇异值分解对文本语义空间进行压缩，能够有效捕捉词项间的隐含关系与全局语义结构，而 LDA 则通过概率生成机制对文档主题进行建模，具有良好的可解释性与主题稀疏性控制能力。二者结合不仅提升了模型对主题的表达力和鲁棒性，还增强了主题语义的一致性与区分度，从而在多维情感与用户关注点分析中，实现了更具深度与广度的文本主题挖掘效果。

首先使用 CountVectorizer 创建一个向量化器，将之前准备好的语料库转换为文档-词频矩阵，并指定提取的特征词数量。初始化潜在狄利克雷分配模型并拟合文档-词频矩阵，接下来循环计算不同主题数量下的主题困惑度，循环范围从 1 到所设定的主题数量 16。对每个主题数量，初始化并训练一个潜在狄利克雷分配模型，计算其困惑度和得分。基于主题困惑度的计算方法得到第一组在线评论数据的最佳主题数量为 8 个，第二组最佳主题数量为 9 个。将这些主题看作新能源汽车的主属性并作为客户满意度的影响因素。

表 3-1 LDA 主题模型核心参数

核心参数设置	设定值
n_components	8,9
max_iter	50
learning_method	batch
CountVectorizer.max_features	500
CountVectorizer.max_df	0.5
CountVectorizer.min_df	1

在本研究中，LDA 模型的核心参数如表 3-1 所示。选择将主题数（n_components）设定为 8 和 9，以便从文本数据中提取出 8 或 9 个具有代表性

的主题，这样既能平衡主题的多样性，又能避免过于复杂的模型。最大迭代次数（`max_iter`）设定为 50，以确保模型在训练过程中能够充分收敛，同时避免因迭代次数过多导致的训练时间过长。学习方法（`learning_method`）采用批量学习（`batch`），这种方法有助于在处理大规模数据集时更加稳定和高效。为了控制特征空间的维度，将 `CountVectorizer` 的 `max_features` 参数设定为 500，意味着模型将仅保留出现频率最高的 500 个词，从而减少计算复杂度并降低过拟合的风险。此外，`max_df` 被设定为 0.5，意味着如果某个词在超过 50% 的文档中出现，该词将被忽略，这样可以去除一些在所有文档中普遍存在但对区分不同主题贡献较小的高频词。最后，`min_df` 设定为 1，确保了在至少 1 个文档中出现的词汇被保留，从而避免了丢失稀有但可能重要的词^[86]。通过这些设置，LDA 模型能够在保证计算效率的同时，更好地捕捉文本数据中的潜在主题和结构，为后续分析提供有价值的洞察。

在完成第潜在狄利克雷分配模型分析后，进行了潜在语义分析模型^[87]，具体分析结果如表 3-3 和 3-4 所示。LSA 模型的核心参数设置如表 3-2 所示。

表 3-2 LSA 主题模型核心参数

核心参数设置	设定值
<code>n_components</code>	8,9
<code>top_keywords_idx</code>	6,7
<code>random_state</code>	6
<code>max_features</code>	1000

将主题数（`n_components`）设置为 8 和 9，目的是从文本数据中提取 8 或 9 个具有代表性的主题^[88]，以便准确捕捉到文本中的潜在语义结构，同时避免过于复杂的模型。

表 3-3 潜在狄利克雷分配模型分析结果

舒适性	外观	空间	功能	性价比	动力	政策售后	能耗
皮面	外型	内部空间	多媒体	价格	起速	补助	耗电量
座椅	英俊	乘用	智能	亲民	提速	送绿牌	省电
音响设备	酷帅	前后排	遥控	物有所值	爆发性	免税	平均用电

避震	中意	纵向	人机交互	预算内	加速度	回访	磨合期
包裹	款式	储物间	智行	销售价格	起步	PDI	比较稳定
体感	大气	后备箱	人脸识别	物美价廉	储备	分期	维持
隔音性	造型	空间布局	辅助	钟意	持久性	不限牌	节电模式

为了确保结果的可重复性，随机种子（random_state）设定为 6，这样每次运行模型时，训练过程中的随机性会被控制，保证了结果的一致性。在特征选择方面，max_features 设定为 1000，表示仅保留出现频率最高的 1000 个特征，从而有效减少特征空间的维度，提高计算效率，并避免过拟合。top_keywords_idx 设定为 6 和 7，意味着每个主题将提取出最具代表性的 6 至 7 个关键词，用于更好地理解 and 解释每个主题的含义。这些参数的选择能够帮助 LSA 模型在保证计算效率的同时，准确提取文本数据中的潜在语义信息，从而为后续的分析 and 应用提供有价值的支持。

表 3-4 潜在语义分析模型分析结果

舒适性	外观	空间	驾驶感受	性价比	续航	内饰	智能化	操控
座椅	设计	后排	方向盘	价格	电耗	座椅	功能	丝滑
隔音	车身、	后备箱	刹车	品牌	空调	用料	智能	踏板
空调	线条	座椅	精准	优惠	混动	质感	影像	反应
玻璃	大灯	储物	手感	成本	成本	屏幕	辅助驾驶	踩油门
静音	尾灯	轴距	稳定性	颜值	模式	手感	自动泊车	技术
减速带	颜色	尺寸	坑洼	综合	电池	耐脏	无线充电	调校

3.4 实验结果与讨论

通过潜在狄利克雷分配和潜在语义分析两种方法对新能源汽车用户评论进行主题提取，提取出了多个用户主要关注的维度。在这些维度下，进一步挖掘了用户具体讨论的核心关键词，揭示了消费者对新能源汽车产品的详细关注点。从两组数据的分析结果来看，新能源汽车用户的需求重点和关注点在近十年间发生了明显变化，反映出行业技术发展、市场环境变化以及消费者认知升级的影响。

（1）用户关注点的整体变化

从第一组数据中可以看出,用户更加关注基本功能性需求,如舒适性、外观、空间、性价比、动力、政策售后及能耗。其中,补助、送绿牌和免税等关键词表明,政策优惠是影响用户购车决策的重要因素。同时,智能化相关的词较少,说明当时新能源汽车的智能技术尚未成为主流卖点。到了第二组数据,用户关注点逐渐向驾驶体验、科技智能化以及操控性倾斜。除了续航、智能化、操控等新维度的加入,用户对功能、刹车影像和自动泊车等智能驾驶辅助系统的关注度大幅提升,说明消费者对新能源汽车的期待已从基本交通工具向智能出行体验转变。此外,电耗、混动和电池成为关注焦点,表明用户对能耗管理和新能源技术的认知更加深入。

(2) 需求演变的驱动因素

新能源汽车用户需求的变化,主要受到以下因素的驱动。技术进步,电池技术、智能驾驶技术的提升,使用户需求从基本性能向高级驾驶体验、智能化拓展。市场成熟度,新能源汽车市场竞争加剧,消费者对产品的要求更加细化,并关注长期使用体验。政策环境变化,早期政府补贴对购买决策影响较大,而近年来,政策重点从购车补贴转向充电基础设施建设,用户关注点也转移至续航、电池寿命等实际使用体验。消费认知升级,用户对新能源汽车的理解从早期的新鲜感,逐步升级为更全面、理性的评价体系,推动行业不断优化产品体验。

早些年间,新能源汽车用户主要关注车辆的基础功能、补贴政策及动力性能,伴随科技的发展,用户需求向智能化、续航优化、驾驶体验及品牌价值等方向升级。这一趋势为新能源汽车企业提供了明确的优化方向,即从传统性能竞争转向智能科技和全生命周期使用体验优化,以更好地满足用户需求。

3.5 本章小结

本章围绕新能源汽车用户需求分析,采用两种主题建模方法,对用户评论数据进行了深入挖掘。通过数据分析,提取出了用户关注的多个核心维度,包括舒适性、外观、空间和驾驶感受等。

通过对新能源汽车用户评论数据的分析,探讨了用户关注点在不同时间段的变化趋势。研究发现,用户需求经历了从基础功能关注向智能化、驾驶体验和续航优化方向的转变。总体来看,新能源汽车用户需求的演变受到技术进步、市场成熟度、政策调整以及消费认知升级等多方面因素的影响。未来,新能源车企应

进一步优化智能化驾驶体验、提升续航表现，并强化品牌建设，以满足不断升级的用户需求。

第 4 章 融合多模型的 NEV 用户细粒度情感分析

在新能源汽车市场竞争日趋激烈的背景下，用户需求已不再局限于基础性能，而是受到多维度因素的影响，尤其是情感因素在需求表达中的作用愈发显著。不同情感极性下，用户对同一属性的关注点及其重要性可能存在显著差异。因此，结合情感分析技术来深入挖掘用户需求变化趋势显得尤为重要^[89]。

为系统性揭示情感因素对用户需求的影响，本章基于新能源汽车用户评论数据，分两组进行分析。具体而言，采用传统机器学习与深度学习相结合的方法进行情感分析，并结合 Kano 模型和 IPA 方法^[90]，分析第一组数据中产品属性的重要性及用户满意度。为进一步揭示不同情感主导下用户关注权重的分布，引入贝叶斯推断方法，分析用户在正向、中性和负向情感主导下对第二组数据中各属性的关注比例^[91]。为下一阶段的属性重要度排序做准备。

4.1 SVM 情感分析模型构建

本章构建 SVM 情感分类器，分析第一组数据中每个主属性下用户在线评论的情感极性^[92]。

表 4-1 LSA 核心参数值

核心参数设置	设定值
kernel	linear
test_size	0.2
analyzer	word
ngram_range	1,2

支持向量机（SVM）模型的核心参数设置如表 4-1 所示，以确保模型的有效性和准确性^[93]。选择将核函数（kernel）设定为线性（linear），因为线性核在高维空间中具有较好的计算效率，且能够有效处理线性可分的文本数据。测试集比例（test_size）设定为 0.2，意味着将数据的 20%用于模型评估，确保评估结果的代表性和有效性。为了提取文本的更丰富特征，将分析单元（analyzer）设定为“word”，即模型在训练过程中将基于单词级别的特征进行处理。同时，

ngram_range 设定为 1,2，表示模型将考虑 1-gram 和 2-gram 特征，这有助于捕捉文本中单个词汇以及连续词组的上下文信息，从而提升模型的表达能力。通过这些参数的配置，SVM 模型能够有效地进行文本分类任务，并提供较为准确的预测结果。

根据数据集中的得分值将连续的分值映射到离散的情感类别中，如表 4-2 所示。将情感分类为异常消极、消极、平淡、积极、异常积极这五类情感，用于确定每个主题下的每条用户评论数据的情感极性^[94]。

表 4-2 得分区间及对应标签

情感得分区间	标签
-1, -0.6	Very_negative（异常消极）
-0.6, -0.2	Negative（消极）
-0.2, 0.2	Neutral（平淡）
0.2, 0.6	Positive（积极）
0.6, 1	Very_positive（异常积极）

4.2 Kano 模型与 IPA 模型构建

本研究在情感分析的基础上，提出了将客户满意度划分为 Kano 类别的分类规则，如表 4-3 所示。对于分类规则解释如下^[95]。

表 4-3 Kano 类别分类规则

N_a^{vp}	N_a^{pos}	N_a^{neu}	N_a^{neg}	N_a^{vn}	Kano 类别
+++	+	-	+	++	诱因属性
+++	+	-	+	+	期望属性
-	-	++++++	-	-	淡漠属性
+	++	-	+	+++	反向属性
+	++	-	+++	+	基本属性

首先 N_a^i 表示用户对产品属性 a 的在线评论数据情感分析结果统计。当 N_a^{pos} 或者 N_a^{neg} 在 N_a^i 中所占比例很小, 或者 N_a^{neu} 在 N_a^i 中占比最大, 则推测用户对这一属性并不太在意, 也就是说中性情感对其影响最大, 因此将此属性归为淡漠要素。其次本研究使用了 N_a^{vp} 和 N_a^{pos} 以及 N_a^{vn} 和 N_a^{neg} 之间的差值, 将产品属性分为 Kano 模型的其他四类^[96]。若 $N_a^{vp} \geq N_a^{pos}$ 并且 $N_a^{neg} \leq N_a^{vn}$, 则产品属性 a 超出了用户的期望, 会出现满意情绪, 若没有达到用户期望并不会使用户出现不满情绪, 根据 Kano 模型中诱因要素的定义, 将属性 a 归入诱因要素。若 $N_a^{vp} \geq N_a^{pos}$ 并且 $N_a^{neg} \geq N_a^{vn}$, 用户对属性 a 越满意, 对该产品评价越高, 相反越不满意评价越低。根据 Kano 模型中期望要素的定义, 将属性 a 归纳为期望要素。若 $N_a^{vp} \leq N_a^{pos}$ 并且 $N_a^{neg} \leq N_a^{vn}$, 属性 a 对用户来说是没有必要的, 如果产品具备这个属性, 用户反而会感到不满意, 这些属性可能是负面的, 或者是不适合特定的用户群体。根据 Kano 模型中反向要素的定义, 将产品属性 a 划分为反向要素。若 $N_a^{vp} \leq N_a^{pos}$ 并且 $N_a^{neg} \geq N_a^{vn}$, 当用户对产品属性 a 满意时, 对其的评价不会提高, 而当对产品属性 a 不满意时, 用户对产品的评价会降低, 就是说这种产品属性是用户认为必须存在的基本属性。

IPA 模型从重要性和满意度两个方面来对新能源汽车用户满意度属性进行分类^[97]。在重要性方面, 本研究通过构建重要性变量来实现属性重要性的测量。

$$importance = \left(\frac{Number_w}{Number_T} \right) \quad (4-1)$$

$Number_w$ 指的是主题出现次数, $Number_T$ 为总评论数。在满意度评估方面, 采用了情感分析的结果作为基础。设定了一个分数规则, 其中异常积极对应获得两分, 积极获得一分, 平淡不得分, 消极扣一分, 异常消极扣两分。随后, 对每个被评估属性的分数进行统计, 并将这些分数规范化到一个介于一到三之间的数值范围, 以此作为绩效值的表现。这种方法有助于更准确地衡量和比较不同属性的绩效水平, 从而为进一步的改进和决策提供了可靠的数据支持。

表 4-4 Kano-IPA 模型

IPA 矩阵所在象限	属性改进	继续保持	Kano 类别
第一象限	—	++++	诱因属性
第四象限	++	+++	
第二象限	-	+++	
第三象限	+	+++	
第一象限	-	+++++	期望属性
第四象限	+++++	-	
第二象限	-	+++	
第三象限	++	++	
第一象限	-	+	淡漠属性
第四象限	+	-	
第二象限	-	++++	
第三象限	-	+++	
第一象限	-	+++++	基本属性
第四象限	+++++	-	
第二象限	-	+++	
第三象限	+	+++	
第一象限	矛盾属性	矛盾属性	反向属性
第二象限	矛盾属性	矛盾属性	
第三象限	-	+++++	
第四象限	矛盾属性	矛盾属性	

为了减少 Kano 模型以及 IPA 模型的局限性，结合两者以更全面和综合的方式来分析产品属性和客户需求^[98]，如表 4-4 所示。首先，对产品不同属性，依据本研究所提出的 Kano 类别分类规则划分成诱因属性、期望属性、淡漠属性、基本属性和反向属性五类。其次利用 IPA 模型对产品属性的重要性和满意度进行计算，绘制 IPA 图，最后依据 Kano-IPA 模型，得出属性改进的优先级。

4.3 Bert-Bi-LSTM 模型构建

本章采用 Bert-Bi-LSTM 模型对第二组数据进行情感分析，以识别用户在不同情感极性下对车辆属性的关注差异^[99]。该方法结合了 Bert 的强大文本表示能力以及 Bi-LSTM 对序列数据的深度学习能力，能够精准捕捉用户评论中的情感倾向。首先，使用 Bert 中文预训练模型对用户评论进行处理，包括分词、填充和截断等操作，并将其输入 Bert 编码器，获取 last_hidden_state 作为文本的语义特征表示。Bert 的双向 Transformer 结构能够充分考虑上下文信息，为后续的 Bi-LSTM 处理提供高质量输入^[100]。在 Bert 编码的基础上，本章加入了 Bi-LSTM 层，

以进一步挖掘文本的时序特征。双向 LSTM 能够同时捕捉评论的前向和后向依赖关系，增强对情感语义的理解，同时通过全局平均池化和最大池化提取不同层次的情感特征，并进行拼接以供分类任务使用。

表 4-5 Bert-Bi-LSTM 模型核心参数

核心参数设置	设定值
model_name	bert-base-chinese
padding strategy	True
labels	0,1,2
Bidirectional Long Short-Term Memory	LSTM(64, return_sequences=True)
epochs	3

Bert-Bi-LSTM 模型的核心参数设置如表 4-5 所示。选择使用 bert-base-chinese 作为预训练模型，它基于 Bert 架构，专为中文文本设计，能够有效地捕捉中文文本中的语义信息。padding strategy 设置为 True，意味着输入的文本会进行填充处理，使其长度一致，以便在神经网络中进行批量训练，避免因文本长度不一致导致的训练问题^[101]。labels 设定为 0, 1, 2，表示模型将进行三分类任务，分别对应不同的标签类别。这一设置使得模型能够区分并处理多种情感或话题的分类问题。

在模型架构方面，采用了 Bi-LSTM，LSTM 的单元数设置为 64，并且设置了 return_sequences=True，意味着 LSTM 将输出每个时间步的隐藏状态，而不仅仅是序列的最终隐藏状态，这有助于捕捉文本中的上下文信息和序列间的复杂关系。最后，epochs 设定为 3，表示模型将训练 3 轮，以确保模型在训练数据上充分拟合，避免过拟合并提高泛化能力。通过这些参数设置，Bert-Bi-LSTM 模型能够有效地结合 Bert 的强大语言表示能力和 Bi-LSTM 的序列建模能力，提供高效且精确的文本分类结果。

在 Bi-LSTM 提取的文本特征基础上，本研究采用 Softmax 分类层进行情感极性分类^[102]。全连接层将池化特征映射到三维情感类别（负面、中性、正面），并通过 Softmax 激活函数计算每个类别的概率分布，训练过程中采用交叉熵损失函数进行优化，以提升模型对情感分类的准确性。模型训练使用 Adam 优化器，

设置训练轮数为 3，以确保训练过程的稳定性和高效性。在训练完成后，本研究使用训练好的 BERT-BiLSTM 模型对新数据进行情感预测。新输入的用户评论经过 BERT 语义编码和 BiLSTM 处理后，获得最终的情感分类结果，通过 Softmax 层计算负面、中性、正面三种情感类别的概率，将预测结果与情感概率添加到数据集中，根据最大概率值映射至相应的情感标签。通过对不同情感类别的分布进行统计分析，可以评估用户评论的情感倾向，探讨情感极性与新能源汽车产品属性之间的关系。

4.4 贝叶斯推断的用户情感量化分析模型构建

在本章中，采用基于贝叶斯推断的方法来计算不同情感极性的相对权重，以量化不同情感倾向对新能源汽车属性评价的影响。具体地，利用 Dirichlet 分布作为先验分布，结合观测数据进行贝叶斯更新，从而推导出更符合实际情况的情感权重分布。首先，根据情感主导情境（积极主导、中性主导、消极主导）设置先验分布参数，分别对应 Dirichlet 分布中的 α 值，以反映不同情感状态下的初始假设。随后，收集并统计第二组数据，获取各情境下的情感极性分布情况，并将其作为观测数据。通过贝叶斯更新，将先验信息与观测数据结合，计算后验分布参数，即将原始 Dirichlet 分布参数与对应的观测数据相加，得到新的参数值。基于后验分布进一步计算情感极性的相对权重，即归一化后的后验分布参数比例，从而得出在不同情感主导情境下，积极、中性和消极情感的最终占比^[103]。此外，为了直观展示后验分布，绘制了 Dirichlet 分布概率密度图，并在图中标注了最终的权重点，以便观察不同情感极性在各情境下的相对变化趋势。通过这一过程^[104]，能够量化不同情感极性对新能源汽车评价的影响，并为后续的用户需求属性分析提供数据支撑。

4.5 实验结果与讨论

（1）Kano-IPA 模型分析结果

为了进一步探讨用户对新能源汽车不同属性的关注重点及其影响，在量化情感极性权重的基础上，引入 Kano 模型和 IPA 模型进行深入分析。分析结果如表 4-6 所示。

表 4-6 新能源汽车 Kano 模型分析

属性	数量 Very_Neg	数量 Neg	数量 Neu	数量 Pos	数量 Very_Pos	属性类型
舒适性	525	3047	1386	3052	256	基本属性
外观	2445	2364	792	1269	1395	诱因属性
功能	39	1007	289	1335	1361	期望属性
性价比	100	963	756	3910	2537	基本属性
空间	47	411	58	3520	4229	期望属性
动力	76	239	214	6639	1100	基本属性
售后政策	128	106	136	398	463	诱因属性
能耗续航	375	1487	1488	3504	1390	基本属性

而 IPA 模型则结合用户感知的重要性和表现水平，识别关键改进方向，分析结果如图 4-1 所示。

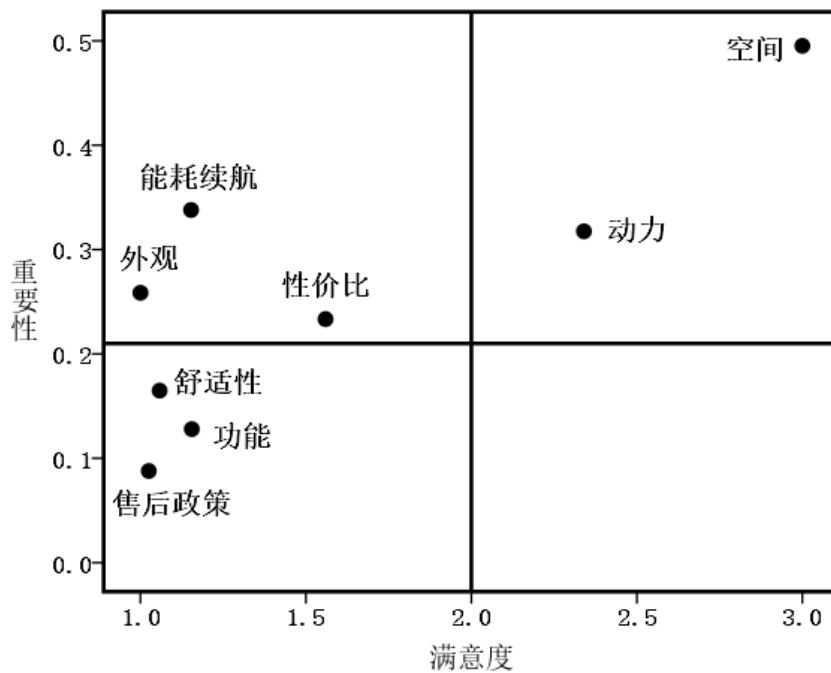


图 4-1 新能源汽车 IPA 分析图

虽然 Kano 模型识别了不同属性的需求类型，IPA 模型分析了用户感知中的重要性和表现水平，但单独使用任一模型都可能存在局限性^[105]。因此，为了更精准地识别需要优化的关键属性，并结合用户需求类型和满意度表现，进一步采用 Kano-IPA 综合分析方法。具体分析结果如表 4-7 所示。

这些结果为新能源汽车制造商提供了重要的历史参考，揭示了过去市场环境 下消费者对产品属性的偏好和关注点。这种分析不仅有助于理解新能源汽车市场

的发展趋势，还能为未来的产品优化和市场策略提供借鉴。尽管市场环境在不断变化，但消费者需求的演变往往是渐进的，而非突变的，因此，对历史数据的深入分析依然具有重要价值。

表 4-7 Kano-IPA 模型分类结果

IPA 矩阵所在象限	Kano 类别	属性	Kano-IPA 模型	
			属性改进	继续保持
第一象限	基本属性	动力	—	+++++
	期望属性	空间		+++++
第三象限	基本属性	舒适性	+	+++
	期望属性	功能	++	++
	诱因属性	售后政策	+	+++
第四象限	基本属性	性价比	+++++	
	诱因属性	外观	++	+++
	基本属性	能耗续航	+++++	

根据 Kano-IPA 模型分析结果，新能源汽车制造商应继续保持在空间和动力属性上的竞争优势。持续投资在这些领域，确保产品在关键性能方面达到或超越用户预期，包括技术、品质和性能的持续优化。与此同时，企业应充分利用这些属性作为市场宣传的核心亮点，以吸引更多潜在客户。在过去几年里，这些关键属性一直是用户高度关注的重点，未来仍可能对市场表现起到重要影响，因此，不仅要巩固已有优势，还需持续探索改进和创新的机会，以确保产品在市场竞争中保持领先地位。

在资源允许的情况下，对舒适性和功能属性进行优化也是提升市场竞争力的重要方向。这些属性的改进能够增强用户的整体体验，提高产品吸引力。在售后政策方面，虽然调整可能受限于行业规定和市场环境，但提供高质量的售后服务仍然至关重要。历史数据分析表明，优质的售后服务有助于提升客户满意度、增强品牌忠诚度，并推动用户口碑传播。因此，即使政策调整空间有限，制造商仍可通过提高服务响应速度、优化客户支持体系以及加强用户教育等方式提升整体售后体验，从而间接提高用户对品牌的认可度。

对于能耗与续航属性，分析结果表明，消费者对新能源汽车的续航能力和能耗效率始终保持高度关注。这一趋势在过去几年中较为稳定，说明续航里程仍然是影响用户购买决策的重要因素。因此，制造商应继续致力于提升电池技术和动力系统的效率，包括优化电池容量、提高电动机效率、改进空气动力学设计等。

同时，充电基础设施的完善对于提升用户体验至关重要，企业可以与充电设施提供商合作，扩大充电网络覆盖范围，并推广快速充电技术。此外，由于新能源汽车的使用模式与传统燃油车存在差异，用户教育也是提升续航体验的关键环节，例如提供驾驶技巧培训、优化充电习惯建议等，以帮助用户更高效地使用新能源汽车。

在性价比方面，分析结果表明，消费者对高性价比产品的需求较为稳定且持续存在。因此，制造商应在保证产品核心性能的同时，注重成本控制和价值优化，以提供更具竞争力的产品。高性价比不仅能够提升市场竞争力，还能扩大用户群体，特别是吸引价格敏感型消费者。通过持续优化生产成本、提高供应链效率，以及提供多元化的价格策略，企业可以进一步强化品牌在市场中的竞争优势。

外观设计作为吸引潜在客户的重要因素，分析结果表明其优化策略可根据资源状况进行调整。在资源充足的情况下，外观设计的创新和升级可以增强市场吸引力，而在资源有限的情况下，保持现有设计并优化其他核心属性可能是更为稳妥的选择。尽管市场趋势可能有所变化，但外观设计始终是影响消费者购买决策的关键因素之一，因此，制造商应结合市场动态，在设计上保持一定的创新力度，以适应不断变化的用户审美需求。

需要强调的是，本研究基于第一组的数据分析，这一时间段的市场环境、技术发展水平和消费者偏好为新能源汽车产业的演进提供了重要的参考。尽管当前市场可能已发生一定变化，但过去的数据仍然能够揭示消费者需求的演变路径，为企业制定长期发展战略提供依据。此外，许多影响用户满意度的核心因素具有较强的稳定性，历史数据的分析仍然对当前市场具有一定的指导意义。因此，新能源汽车制造商应结合历史趋势与最新市场动态，综合分析用户需求的变化，以确保在竞争激烈的市场环境中保持持续领先。

（2）Bert-Bi-LSTM 模型分析结果

采用 Bert-Bi-LSTM 模型对新能源汽车用户评论进行情感分析，从正面、负面和中性三种情感极性出发，探讨用户对各个属性的不同态度。通过分析各属性下不同情感类别的评论数量，能够进一步理解用户关注的重点，以及他们对某些属性的满意度和改进期望。如表 4-8 所示，展示了 Bert-Bi-LSTM 模型分析得到的情感分布情况。

表 4-8 BERT-BiLSTM 分析结果

属性	正面评论	负面评论	中性评论
空间	29365	1628	1274
驾驶感受	3038	8980	20340
续航	3253	202	28967
外观	18027	418	13998
内饰	21067	3392	7449
性价比	312	42	12
智能化	1937	4244	7257
操控	5702	1042	1042
舒适性	4692	4194	2111

新能源汽车用户对不同属性的评价呈现显著差异。空间表现受到高度认可，表明宽敞空间是新能源汽车的重要竞争优势。驾驶感受评价则呈现两极分化，反映出用户对动力输出、操控性及驾乘体验的分歧。续航里程作为用户关注的核心指标，中性评价占据主导，而正面评价明显多于负面，表明用户对续航能力存在一定不确定性或更高期望。外观设计获得广泛认可，表明当前新能源车型的造型符合市场审美。内饰方面的正面评价占主导地位，应继续保持。性价比的评价数量相对较少，但正面评价仍是最高，表明用户对新能源汽车的价格与性能匹配度总体持积极态度。智能化功能的用户反馈较为分化，负面评论高于正面，且中性评论占比较大，说明智能驾驶辅助系统、车载娱乐系统的稳定性和易用性仍有优化空间。操控性能整体获得认可，电动车的低重心设计和平顺加速或为优势因素，但仍需针对用户反馈优化转向精度、悬挂调校等细节，以进一步提升驾驶操控体验。舒适性评价呈现均衡分布，反映出座椅舒适度、悬挂系统调校、车内噪音控制等方面存在改进空间。

Bert-Bi-LSTM 模型的情感分析结果表明，空间、外观、内饰、操控等属性获得较高评价，续航与性价比尽管正面评价占优，但仍存在优化空间，而驾驶感受、智能化与舒适性则是用户意见分歧较大的重点改进领域。该分析为新能源汽车制造商提供了关键用户反馈，可用于指导产品迭代、市场策略优化及品牌定位调整，以精准满足消费者需求并提升市场竞争力。

（3）贝叶斯推断结果

在新能源汽车用户评价分析中，不同情感极性的影响程度并不均衡，而是受特定情境的调节。因此，在本部分研究中，采用贝叶斯推断方法，对不同情感主导情境（积极主导、中性主导、消极主导）下的情感极性权重进行量化计算。

分析结果如图 4-2、4-3 和 4-4 所示。通过引入 Dirichlet 先验分布并结合观测数据进行贝叶斯更新，推导出不同情境下更符合实际的情感极性权重分布。积极主导情境下，用户的情感极性分布表现为：积极情感占比 60%，中性情感占比 30%，消极情感占比 10%。这表明当用户整体情绪较为正面时，评价倾向于强调新能源汽车的优点。在中性主导情境下，情感极性分布相对均衡，其中中性情感占比 50%，积极情感占比 30%，消极情感占比 20%。这一结果显示，当用户情绪未呈现明显偏向时，评价内容更趋于客观，可能涉及对新能源汽车性能的全面分析。在消极主导情境下，消极情感占比 50%，中性情感占比 30%，积极情感占比 20%。这一结果表明，当用户情绪倾向负面时，评价更集中于新能源汽车的不足。这一分析不仅有助于理解用户在不同情境下的情感偏向，还为后续新能源汽车属性重要度排序的研究提供了关键数据支持。

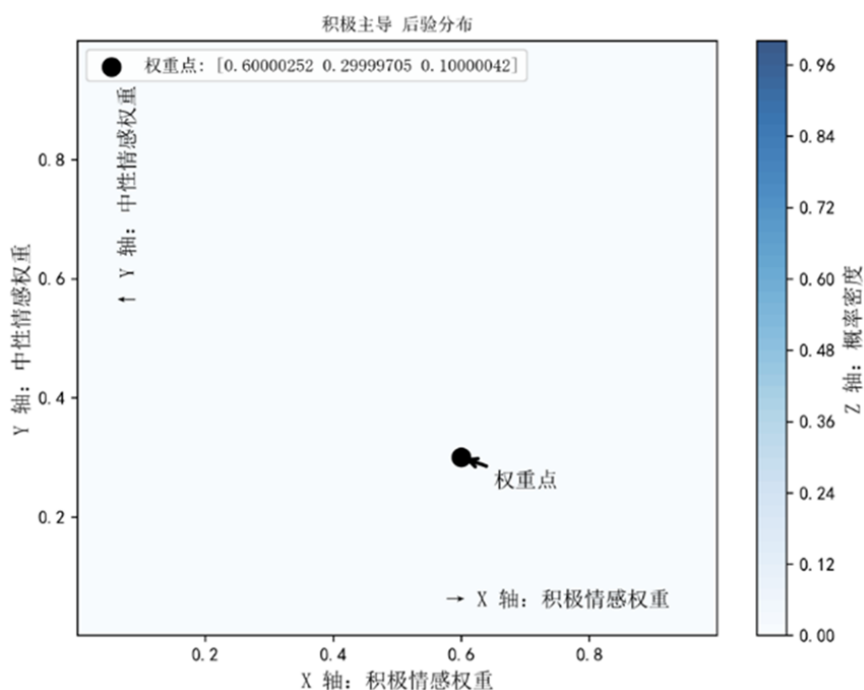


图 4-2 积极情感下贝叶斯推断

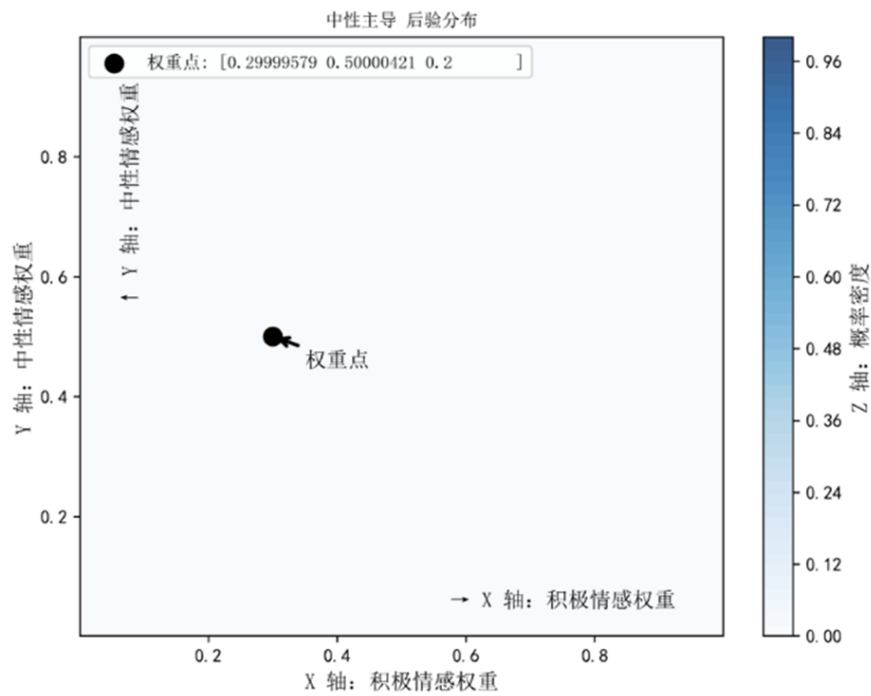


图 4-3 中性情感下贝叶斯推断

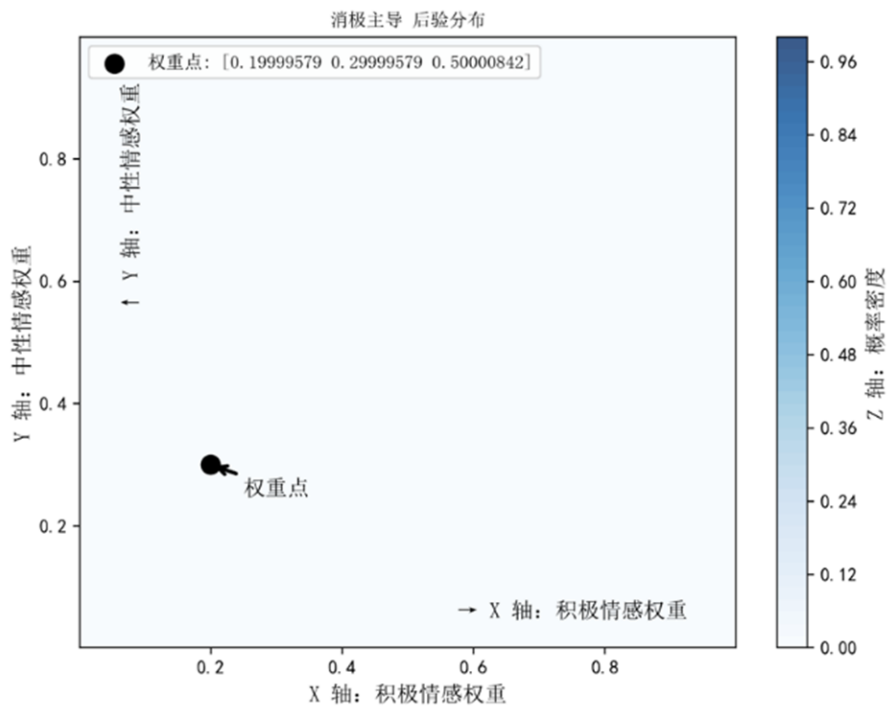


图 4-4 消极情感下贝叶斯推断

本章对新能源汽车用户评论数据，采用不同阶段的数据分析方法，探讨了情感因素对用户需求表达的影响。研究表明，新能源汽车市场的成熟，用户需求已从基础功能关注转向更加多维度的考量，情感极性在不同情境下显著影响用户对各属性的关注度及其权重。

在第一阶段,用户评论内容较为简洁,关注点主要集中在新能源汽车的续航、价格、充电便利性等核心性能。通过 SVM 进行情感分类,并结合 Kano 模型和 IPA 方法分析用户满意度,研究发现,用户对于新能源汽车的核心性能存在较为稳定的需求排序,其中续航能力在多个情感极性下均占据重要位置。然而,负面情感主导的评论主要集中在续航衰减、充电设施不足、价格过高等问题上,反映出该阶段用户对新能源汽车的主要顾虑仍围绕基础功能展开。相比之下,在第二阶段,用户评论表达方式更为自由,关注点更加多元化。通过 Bert-Bi-LSTM 结合贝叶斯推断的分析,为后续的属性重要度排序做准备。

4.6 本章小结

本章通过分析新能源汽车用户评论数据,系统地探讨了情感极性对用户需求表达的影响,并在不同阶段采用了不同的数据分析方法,以精准刻画用户需求变化。本研究结合情感分析技术,基于 SVM、Bert-Bi-LSTM 模型和贝叶斯推断的方法,深入挖掘了情感极性对新能源汽车属性关注点的影响。

对于第一组数据,通过支持向量机进行情感分类,结合 Kano 模型和 IPA 方法,分析了用户对各产品属性的重要性及满意度,进一步揭示了用户对续航、价格、充电设施等核心功能的关注。在这一阶段,用户的关注点更多集中于续航衰减和充电便利性等核心性能问题,表明基础功能仍然是这一阶段用户需求的核​​心所在。对于第二组数据,采用了 Bert-Bi-LSTM 模型进行情感分析,结合贝叶斯推断方法,分析了不同情感极性下,用户对各属性的关注权重,为后续研究提供支持。

通过本章的分析,能够清晰看出,随着时间的推移和市场的成熟,用户需求的表达方式和关注点发生了显著变化,情感极性对需求排序的影响越来越明显。基于情感分析的需求挖掘方法,尤其是 SVM、Bert-Bi-LSTM 模型与贝叶斯推断的结合,提供了一种更加精准、动态的用户需求分析模型。下一章将基于本章的分析结果,进一步探讨如何利用情感权重信息优化新能源汽车的产品设计和市场策略,并提出具有针对性的改进建议。

第 5 章 基于 MCDM 的 NEV 属性重要度排序

新能源汽车用户的需求具有多维度特征,不同用户对不同属性的关注度存在差异,且这种关注度会受到情感因素的影响。因此,如何科学地对新能源汽车属性进行重要度排序,进而优化产品设计和市场策略,是本研究的核心问题之一^[106]。

在上一章的研究中,通过贝叶斯推断方法分析了用户情感极性与新能源汽车不同属性关注度之间的关系,获得了不同情感主导情况下的属性权重。在此基础上,本章旨在构建新能源汽车需求排序模型,并对排序方法进行对比分析,以探讨不同方法在情感主导场景下的适用性。具体而言,本章采用了两种方法对新能源汽车属性进行排序。

5.1 主观权重设定排序方法

在本章研究中,采用 VIKOR 方法对新能源汽车属性重要度排序,以量化不同情感主导情况下用户对各属性的关注度。VIKOR 方法的核心思想是基于加权综合评估不同指标的重要度,因此,权重的设定会直接影响最终的排序结果。由于研究的目的是分析不同情感主导情况下的属性排序,而非让模型自动学习权重,因此采用人为设定的权重,以确保排序逻辑符合研究目的。

VIKOR 方法是一种多准则决策方法,适用于处理具有相互冲突属性的决策问题^[107]。本研究基于三种情感主导场景构建决策矩阵,接下来对决策矩阵进行了归一化处理,以确保数据的可比性。归一化后,计算了全局效用度量,通过加权求和的方法得出了各属性的综合得分。这一度量反映了各评价指标在整体决策中的贡献程度。

随后,确定了理想解和反理想解,即在所有评分数据中,每个指标的最大值和最小值。理想解代表最优情况,而反理想解代表最差情况。基于这一标准,计算了每个情感类别与理想解及反理想解之间的离差度量,其中离差度量用于衡量各情感类别相对最优解和最差解的距离。在计算离差度量后,进一步求得相对接近度,该指标反映了各选项(即各情感类别的评分)相对接近理想解的程度。相对接近度越高,说明该选项在整体评价中越接近理想情况。

最后,计算了 VIKOR 值,以综合衡量各情感类别在不同评价指标上的表现。

该值可用于排序和决策，帮助判断在新能源汽车评论数据中，用户关注的情感维度如何影响整体评价趋势。通过该分析，能够更清晰地识别用户情感倾向，进而为产品优化和市场策略提供数据支持^[108]。

5.2 客观权重设定排序方法

在本章研究中，采用 LightGBM 排序学习方法，对新能源汽车用户评论数据进行排序分析，以探讨用户关注的属性在不同情感类别下的影响。

表 5-1 LightGBM 模型核心参数设置

核心参数设置	设定值
objective	lambdarank
metric	ndcg
learning_rate	0.05
num_leaves	31
min_data_in_leaf	1
verbose	-1

在本研究中，LightGBM 排序模型参数设置如表 5-1 所示。采用了 LambdaRank 作为目标函数（objective='lambdarank'），该函数是一种基于排序学习的优化方法，能够更好地提升模型在排序任务中的表现。模型评估指标选用 NDCG（Normalized Discounted Cumulative Gain，metric='ndcg'），用于衡量模型在排序结果中的准确性，尤其关注前几位的重要排序效果。为提升模型的稳定性和泛化能力，设置学习率（learning_rate=0.05）为较小值，使模型在训练过程中缓慢而稳定地收敛；同时，为控制模型复杂度、避免过拟合，设置最大叶子节点数为 31（num_leaves=31），这是 LightGBM 的默认推荐值。由于本研究样本数据量较小，因此将每个叶子节点的最小样本数设置为 1（min_data_in_leaf=1），以确保每个属性信息都能被模型充分学习。此外，为避免训练过程中产生冗余输出信息，设置日志等级为静默模式（verbose=-1）。综合来看，这些参数设置具有较好的逻辑性和适配性，既考虑了模型性能，又兼顾了数据规模与实际应用需求。

该方法基于贝叶斯推断，能够自动学习数据中的权重，避免人为设定带来的主观性，使得排序结果更加符合实际数据分布^[109]。首先，构建了包含九个核心属性（如空间、驾驶感受、续航和外观等）的数据集，同时包括用户对这些属性的提及频率、评分均值、以及正面、中性、负面情感的数量，以全面反映用户对不同属性的关注度和情感倾向。其次基于上一章贝叶斯推断的结果，给三种情感主导下设置了权重。最后，计算了加权得分，作为各属性在不同情感类别下的重要度衡量标准，将加权得分离散化为三个等级，以便于 LightGBM 进行排序学习。

属性	正面评论	负面评论	中性评论	提及频率	用户评分	
空间	29365	1628	1274	30815	4.760475	权重设定
驾驶感受	3038	8980	20340	4509	4.749255	
续航	3253	202	28967	18901	4.744052	
外观	18027	418	13998	25204	4.845593	
内饰	21067	3392	7449	25431	4.597092	
性价比	312	42	12	19947	4.786173	
智能化	1937	4244	7257	5443	4.782214	
操控	5702	1042	1042	264	4.708969	
舒适性	4692	4194	2111	6493	4.539969	

图 5-1 特征值选择

在特征工程方面，选取了提及频率、用户评分、正面评论数量、负面评论数量和中性评论数量作为模型输入特征，目标变量(标签)则为属性的重要度等级，如图 5-1 所示。在模型训练阶段，采用了 LightGBM 的 LambdaMART(LambdaRank) 算法，该算法基于梯度提升树，能够高效处理大规模排序问题，并在训练过程中优化排序效果^[110]。训练完成后，利用模型对各属性进行预测得分计算，并按照预测得分降序排列，最终得到新能源汽车用户关注属性的排序结果。

5.3 实验结果与讨论

(1) VIKOR 模型结果

在 VIKOR 模型的分析中，分别设置了三种不同的情感权重组合，以探讨消费者在不同情绪主导下对新能源汽车属性的关注差异，具体分析结果如表 5-2 所示。当积极情感作为主导时，排序结果为外观、舒适性、性价比、智能化、操控、内饰、空间、续航、驾驶感受。这一结果表明，在积极情绪占主导时，消费者更加关注能够带来愉悦体验和视觉吸引力的汽车属性。外观的时尚感和吸引力能够激发消费者的积极情绪，而舒适性和性价比则满足了他们对驾乘享受和经济价值

的追求。因此，在消费者情绪高涨时，提升汽车的外观设计、乘坐舒适性和性价比将有助于增强品牌吸引力。

表 5-2 VIKOR 模型分析结果

汽车属性	S_i	R_i	情感权重		情感权重		情感权重	
			0.6: 0.2: 0.2		0.2: 0.6: 0.2		0.2: 0.2: 0.6	
			Q_i	排名	Q_i	排名	Q_i	排名
空间	0.088	0.265	0.610	7	0.230	9	0.210	9
驾驶感受	0.013	0.033	0.320	9	0.720	2	0.561	3
续航	0.178	0.080	0.356	8	0.652	3	0.252	7
外观	0.056	0.169	0.783	1	0.751	1	0.383	4
内饰	0.056	0.169	0.651	6	0.355	7	0.251	8
性价比	0.057	0.171	0.737	3	0.611	4	0.337	5
智能化	0.040	0.120	0.734	4	0.604	5	0.334	6
操控	0.027	0.081	0.727	5	0.327	8	0.583	2
舒适性	0.014	0.040	0.741	2	0.380	6	0.780	1

当中性情感为主导时，排序结果变为外观、驾驶感受、续航、性价比、智能化、舒适性、内饰、操控、空间。这表明，在中性情绪主导的情况下，消费者更倾向于以理性视角评估车辆性能和实用性，尤其是驾驶体验和续航能力。相比情绪化决策，消费者在此状态下更注重汽车的实际使用价值。因此，提升续航能力和驾驶体验将成为影响他们购车决策的重要因素。

当消极情感作为主导时，排序结果变为舒适性、操控、驾驶感受、外观、性价比、智能化、续航、内饰、空间。在消极情绪主导时，消费者对舒适性和操控性的关注度显著提升。这可能源于他们在消极情绪状态下对驾驶体验的敏感性增强，更容易察觉不适感。因此，汽车制造商在优化用户体验时，应着重提升座椅舒适度、悬架调校以及驾驶操控性，以降低负面体验带来的影响。

基于 VIKOR 模型分析结果，外观设计始终占据重要位置，无论情感权重如何变化，外观设计都是消费者关注的核心属性。这表明，在新能源汽车市场中，产品的视觉吸引力和时尚感依然是影响消费者决策的重要因素。因此，汽车制造

商应持续加强外观设计创新，打造具有个性化和辨识度的车型，以提升市场竞争力。

(2) LightGBM 模型结果分析

表 5-3 LightGBM 分析结果

汽车属性	情感权重		情感权重		情感权重	
	0.6: 0.3: 0.1		0.3: 0.5: 0.2		0.2: 0.3: 0.5	
	得分	排名	得分	排名	得分	排名
空间	4.70	0	-0.94	3	-1.20	4
驾驶感受	-1.18	4	4.79	0	4.74	0
续航	-0.66	3	4.75	1	4.38	1
外观	4.64	1	4.74	2	-0.41	3
内饰	4.58	2	-1.10	5	3.57	2
性价比	-5.04	8	-4.70	7	-4.86	7
智能化	-4.25	6	-1.08	4	-1.53	5
操控	-4.48	7	-4.97	8	-4.95	8
舒适性	-1.22	5	-4.27	6	-4.11	6

在 LightGBM 方法的分析中，使用了三组不同的情感权重组合来研究消费者在不同情绪主导下对新能源汽车属性的关注程度，具体分析结果如表 5-3 所示。通过得分和排名的变化，深入探讨消费者在不同情感状态下对各项属性的重要度排序。

在积极情感主导的情况下，用户的关注更多偏向于与驾驶体验和外观相关的属性，空间得分为 4.70，排序第一。这表明在积极情感影响下，用户高度关注车辆的空间表现，可能是因为空间直接关系到舒适性和使用便捷性。驾驶感受排名靠后，得分为-1.18。负面情感主导下，驾驶感受的评分较低，表明这方面可能受到批评或反馈较差，用户关注的焦点可能更多地集中在其他实用性或视觉上的因素。外观得分为 4.64，排名第二，显示了外观作为车辆的颜值因素，能够有效吸引用户的注意力。性价比和智能化分别排名倒数，得分较低，说明在积极情感下，用户不太关心性价比和智能化水平。

在中性情感主导的情况下，用户对属性的评价更趋于中立，并且不同属性之间的得分差异较小，排名也较为平衡，驾驶感受和续航的得分高，分别为 4.79 和 4.75，排名前两位，显示出用户在中性情感主导下对驾驶体验和续航能力的关注较为均衡。空间、外观、内饰等属性得分相对较高，这表明用户的关注度相对分散，未出现明显的高低差异。性价比和舒适性的得分较低，且排名后位，反映了用户对于这些属性的评价较为平稳，可能没有突出优缺点。

在消极情感主导的情况下，用户的负面评价显著影响了属性的得分，尤其是那些涉及使用体验和质量的属性，驾驶感受、续航、外观等属性得分较高，表明这些属性相较于其他部分，仍然较为受欢迎。驾驶感受得分为 4.74，排名第一，表示即使在消极情感下，驾驶体验仍是用户的重点关注对象。内饰得分为 3.57，排名第二，显示出虽然内饰在消极情感下存在一定问题，但用户对其的评价依然较为正面。智能化、操控、性价比等属性得分较低，且排名较低，尤其在负面情感影响下，这些属性的负面评价显著影响了它们的排名。

在积极情感主导下，消费者对车辆的空间、外观等舒适性和视觉属性表现出了较高的关注。研究发现，积极情感使得消费者更加注重体验感和感官享受，这一方面体现在对车辆内饰、驾驶感受以及外观设计的偏好上。消费者在情绪愉悦时，更倾向于选择那些能带来更大空间感、更好视觉体验的车辆。空间作为排名第一的属性，表现出消费者在积极情感下对舒适度的重视，尤其是在家庭出行或长途驾驶时，宽敞的空间能带来更加愉悦的驾乘体验。此外，外观设计也往往是消费者在购车时的第一印象，时尚、独特的外观设计能够吸引消费者的眼球，进而影响他们的购买决策。因此，汽车企业在此情感情境下，应加大在空间布局和外外观设计方面的投入，推出符合消费者审美和舒适需求的车型，以吸引更多情感积极的消费者。

在中性情感主导时，消费者的评价趋向理性、客观，关注点更加广泛和均衡。此时，消费者的决策不再由强烈的情绪驱动，而是基于对车辆综合性能的综合考量。从得分来看，驾驶感受、续航等核心性能指标的排名较高，显示出在中性情绪下，消费者更注重车辆的实用性和可靠性。尤其是在续航这一属性上，消费者对车辆的使用体验、长途驾驶的可持续性和稳定性有更高的要求。这一情境下，车辆的技术性能、舒适性、续航能力和智能化水平等多项属性都显得同样重要。

企业在此情境下需要保持各项属性之间的平衡，特别是在驾驶感受和续航能力方面，确保车辆的基本性能能够满足广泛用户的需求。此外，中性情感主导下，消费者更加理性，不容易受外部营销影响，因此，提供稳定的产品质量和高性价比也变得更加重要。

当消费者处于消极情感主导的状态时，他们对车辆的评价往往受到负面情绪的强烈影响。与积极情感状态下的偏好相比，消极情绪会使消费者更加注重实际的使用体验，尤其是在驾驶感受、操控性和舒适性等方面。研究发现，驾驶感受、续航等性能指标在消极情感下仍然表现较好，显示出消费者在面对负面情绪时，仍然需要车辆提供安全感和舒适感。尤其是在负面情绪主导下，驾驶的安全性和稳定性成为了消费者购买决策的关键因素，操控性、驾驶安全性和舒适度成为了优先考虑的要素。

然而，消极情感下，性价比、智能化等方面的属性排名明显下降，这表明消费者在情绪低落时，对一些附加功能和技术要求并不会给予过多关注，反而更倾向于关注那些能直接改善驾驶体验、提供情感舒适的核心功能。这意味着，消费者在消极情绪下更关注车辆的基础性能和舒适性，而对于智能化功能、性价比等附加价值的关注度较低。汽车企业在这种情绪情境下，应加强产品的舒适性和操控性能，尤其是提升车辆的安全性能和操作稳定性，以帮助消费者缓解负面情绪并提升他们的满意度。此外，企业还应关注如何避免负面情绪的积累，避免消费者因过多负面情绪的影响而对品牌形象产生不利影响。

在本章研究中，对数据集进行了 VIKOR 模型和 LightGBM 模型分析，以探讨消费者在不同情绪状态下对新能源汽车属性关注的差异。两种方法在揭示总体趋势上具有较高的一致性，但由于其权重确定机制不同，结果在细节上存在一定差异，为消费者偏好提供了多角度的解释。

具体而言，在积极情绪主导的情境下，VIKOR 模型的分析结果显示，消费者对外观、舒适性和性价比等属性尤为关注，表明正向情绪能够激发消费者对愉悦体验和视觉吸引的偏好。而 LightGBM 方法则发现，在积极情绪状态下，车内空间得分最高，其次为外观，而驾驶感受的评分较低。这一差异可能源于 VIKOR 模型中预设的权重更倾向于反映研究者主观设定的评价标准，而 LightGBM 则通过贝叶斯推断自动生成权重，从而捕捉到数据中更为细微的偏好特征。

在中性情绪条件下，两种方法均指出驾驶体验和续航能力为消费者较为重视的属性。VIKOR 模型侧重于这两项指标的排序，而 LightGBM 方法则显示出各属性得分较为平衡，说明在理性状态下，消费者对车辆整体性能作出较为均衡的评价，这与中性情绪下决策过程趋向理性、全面的特征相一致。

在消极情绪主导时，VIKOR 模型表明消费者对舒适性与操控性的要求显著提升，反映出在负面情绪下，消费者更加关注能够减轻驾驶不适和提升车辆稳定性的关键属性。与此同时，LightGBM 方法显示，即使在消极情绪下，驾驶体验依然维持较高的得分，内饰设计也占有一定重要性，而一些附加功能如智能化和性价比的评分相对较低。这提示在消极情绪中，消费者主要关注基础驾驶性能与舒适性，而对附加属性的敏感性降低。

综上所述，两种方法均确认外观设计在消费者评价中始终处于核心地位，但在不同情绪状态下，消费者关注的侧重点存在明显差异：在积极情绪下，消费者更重视感官体验和车内空间的舒适性；在中性情绪下，则理性关注驾驶体验和续航能力；而在消极情绪下，则主要看重能够直接改善驾驶体验、提升舒适度和操控性的基础性能。此外，VIKOR 模型中权重的主观预设与 LightGBM 方法中通过贝叶斯推断自动获得权重的机制不同，导致了二者在细节排序上的微妙差异，但总体结论具有较高的一致性。这一结果为新能源汽车制造商提供了理论依据，表明在产品设计和市场策略制定中，应在保持外观吸引力的同时，根据不同情绪状态下消费者需求的变化，优化车内空间布局及核心驾驶性能，以更好地满足市场不断演进的多层次需求。

5.4 本章小结

本章通过运用两种不同的分析方法，VIKOR 模型和 LightGBM 模型对消费者在不同情感主导下，对新能源汽车属性的关注度进行了深入分析。选择这两种方法的主要原因在于它们能够从不同角度和方式对消费者偏好进行有效建模和分析，从而提供全面且多维的分析结果。

首先，VIKOR 模型通过人为设定情感权重，直观地模拟了不同情感状态下消费者对新能源汽车属性的关注差异。通过调整情感权重的比例，VIKOR 模型能够展示消费者在不同情绪情境下的决策优先级。这一模型的优势在于其简洁、

透明且能够清晰展示每个属性的优先排序，同时为新能源汽车制造商提供了关于如何根据消费者情绪调整产品设计的直观指导。

其次，LightGBM 方法通过数据驱动的方式，使用自动化的学习机制根据训练数据得出情感权重，并根据这些权重对消费者对各项属性的关注进行排名。这种方法避免了人为设定权重带来的偏差，并通过大量数据训练得到更加精准且可靠的结果。该方法的优势在于其高效性、准确性以及能够处理复杂的非线性关系，为新能源汽车市场中的实际用户行为建模提供更具操作性的见解。

通过这两种方法的对比分析，得出了一些关键结论。VIKOR 模型提供了较为直观的情感权重设定和排序结果，有助于在实际市场中快速进行情感场景的模拟和决策。而 LightGBM 方法则能够在处理复杂数据和大规模样本时展现其数据驱动的优势，提供更加精确且科学的消费者偏好分析。两者的结合，不仅弥补了单一方法的局限性，还使得分析结果更加全面且可靠。VIKOR 和 LightGBM 方法的互补性使得本研究能够从不同的视角对消费者情感和需求进行深度挖掘。VIKOR 模型的直观性和 LightGBM 方法的精准性相结合，不仅为新能源汽车属性优化提供了理论依据，也为实际产品设计和市场策略的调整提供了数据支持。

通过构建的新能源汽车属性排序模型为企业实现产品设计的动态调整提供了方法支持。通过 VIKOR 模型可实现基于专家经验或战略导向的静态权重设定，用于指导产品初始设计阶段的决策；而引入 LightGBM 模型，利用其基于贝叶斯优化的权重学习机制，能够在持续收集用户反馈数据的基础上动态调整各属性的重要性排序。新能源汽车企业可据此建立“数据—模型—反馈”闭环机制，定期采集用户评论与行为数据，借助 LightGBM 实时更新用户偏好结构，从而动态优化产品配置与功能设计。该方法不仅提高了产品设计对市场需求变化的响应速度，也有助于实现从以企业为中心向以用户为中心的转变，提升产品竞争力与用户满意度。

第 6 章 结论与展望

6.1 研究结论

本研究通过主题挖掘，融合情感分析与多准则决策方法，构建了属性-情感-决策三维分析框架，系统揭示了新能源汽车属性重要度的动态变化规律，研究结果表明，不同情感状态下消费者对汽车属性的关注存在显著差异，在积极情绪主导时，外观设计与空间布局的排名最高，反映出消费者在情绪高涨时对视觉吸引力和实用性有较高的需求；而在中性情绪主导时，驾驶感受与续航性能成为核心关注点，显示出消费者在理性评价时更注重车辆的实际性能和使用价值；在消极情绪主导下，虽然智能化与操控性受到显著负面评价，但驾驶感受仍保持较高排名，表明即使在负面情绪中，核心功能属性依然不可替代。

此外，LSA 与 LDA 主题挖掘模型的互补应用，提高了识别与获取用户需求的全面性。结合 Bert-Bi-LSTM 模型和支持向量机模型，充分利用深度学习与传统机器学习模型的优势，实现更加精细的情感分类，从而精准刻画情感波动如何影响用户偏好，确保需求排序的准确性。采用了 Kano 模型与 IPA 模型相结合的方式，将属性划分为多种类型，为企业提供保持-改进-优化三级策略。使用 VIKOR 模型与 LightGBM 模型，结合贝叶斯推断对属性重要度实现动态排序。为决策支持提供了有力的数据支撑和理论依据。

6.2 研究局限性与展望

本研究数据来源局限于单一平台，因此可能遗漏部分用户群体的需求表达。研究主要依赖文本分析，未能结合语音语调、面部表情等多模态数据，从而对隐晦情感的捕捉能力有限。

针对上述局限性，未来研究可从多个方向进一步深化与拓展。首先，多源数据融合与群体细分将是关键，未来可整合来自社交媒体、客服工单、传感器日志等多个渠道的数据，构建全域用户画像，并利用分层主题模型对用户需求进行地域、年龄及使用场景等多维度的细分，以实现更精准的属性重要度排序。其次，多模态情感分析与实时响应也是重要方向，通过融合语音、图像以及生理信号的数据，开发更为全面的多模态情感分类模型，同时结合边缘计算与流数据处理技

术，实现情感权重与排序结果的秒级动态更新，以适应快速变化的市场环境。再者，动态模型优化与跨领域验证方面，未来可以探索非参数贝叶斯模型，以自动适应情感分布的变化，并将构建的框架扩展至智能家居、消费电子等其他领域，验证其通用性和迁移能力。最后，用户参与式决策机制的构建也十分必要，通过建立用户反馈闭环系统，允许消费者直接参与属性重要度投票，并探索联邦学习技术，在保护用户隐私的前提下实现多企业数据协同分析，从而进一步提升整个决策支持体系的精准性和适应性。

参考文献

- [1] 张紫琼,叶强,李一军. 互联网商品评论情感分析研究综述 [J]. 管理科学学报, 2010, 13 (06): 84-96.
- [2] 沈超, 王安宁, 方钊, 等. 基于在线评论数据的产品需求趋势挖掘[J]. 中国管理科学, 2021, 29(5): 211-220.
- [3] Albayrak T, Caber M. Prioritisation of the hotel attributes according to their influence on satisfaction: A comparison of two techniques[J]. Tourism Management, 2015, 46: 43-50.
- [4] 薛晨杰,黄义兵,王召义.在线评论产品属性提取文献综述[J].现代商业,2023,(01):168-172.
- [5] Jelodar H, Wang Y, Yuan C, et al. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey[J]. Multimedia Tools and Applications, 2019, 78: 15169-15211.
- [6] Jing X Y, Zhang D, Tang Y Y. An improved LDA approach[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2004, 34(5): 1942-1951.
- [7] Kim S G, Kang J. Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews[J]. Information Processing & Management, 2018, 54(6): 938-957.
- [8] Yang S, Yao J, Qazi A. Does the review deserve more helpfulness when its title resembles the content? Locating helpful reviews by text mining[J]. Information Processing & Management, 2020, 57(2): 102179.
- [9] Qaiser S, Ali R. Text mining: use of TF-IDF to examine the relevance of words to documents[J]. International Journal of Computer Applications, 2018, 181(1): 25-29.
- [10] Aizawa A. An information-theoretic perspective of tf-idf measures[J]. Information Processing & Management, 2003, 39(1): 45-65.
- [11] Abdullah D M, Abdulazeez A M. Machine learning applications based on SVM classification a review[J]. Qubahan Academic Journal, 2021, 1(2): 81-90.

- [12] Kurani A, Doshi P, Vakharia A, et al. A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting[J]. *Annals of Data Science*, 2023, 10(1): 183-208.
- [13] Thuy D T T, Thuy L T M, Bach N C, et al. Designing a deep learning-based application for detecting fake online reviews[J]. *Engineering Applications of Artificial Intelligence*, 2024, 134: 108708.
- [14] Chen H, Wang Y, Guo T, et al. Pre-trained image processing transformer[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 12299-12310.
- [15] 李赵春, 周永照, 冯卫奔, 等. 基于 Transformer 模型的手势脑电信号分类识别[J]. *科学技术与工程*, 2023, 23(5): 2044-2050.
- [16] Areshey A, Mathkour H. Transfer learning for sentiment classification using bidirectional encoder representations from transformers (BERT) model[J]. *Sensors*, 2023, 23(11): 5232.
- [17] 刘继, 顾凤云. 基于 BERT 与 BiLSTM 混合方法的网络舆情非平衡文本情感分析[J]. *情报杂志*, 2022, 41(4): 104-110.
- [18] Yu Y, Si X, Hu C, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. *Neural Computation*, 2019, 31(7): 1235-1270.
- [19] Gers F A, Schmidhuber J, Cummins F. Learning to forget: Continual prediction with LSTM[J]. *Neural Computation*, 2000, 12(10): 2451-2471.
- [20] Siarni-Namini S, Tavakoli N, Namin A S. The performance of LSTM and BiLSTM in forecasting time series[C]//*2019 IEEE International Conference on Big Data (Big Data)*. IEEE, 2019: 3285-3292.
- [21] Lin X M, Ho C H, Xia L T, et al. Sentiment analysis of low-carbon travel APP user comments based on deep learning[J]. *Sustainable Energy Technologies and Assessments*, 2021, 44: 101014.
- [22] Jain P K, Srivastava G, Lin J C W, et al. Unscrambling customer recommendations: a novel lstm ensemble approach in airline recommendation prediction using online reviews[J]. *IEEE Transactions on Computational Social Systems*, 2022, 9(6): 1777-1784.

- [23] 赵妍妍, 秦兵, 刘挺. 文本情感分析[J]. 软件学报, 2010, 21(8): 1834-1848.
- [24] Xiao Q, Li S, Zhang X, et al. Deconstructing online hospitality review systems: User quality experience toward design features[J]. Journal of Organizational and End User Computing (JOEUC), 2022, 34(2): 1-17.
- [25] Xu Q, Jiao R J, Yang X, et al. An analytical Kano model for customer need analysis[J]. Design Studies, 2009, 30(1): 87-110.
- [26] 郑治伟, 罗兹柏, 张晓燕. 基于 LSTM-IPA 模型的文化遗产人员解说服务质量评估研究——以北京故宫为例[J]. 旅游科学, 2023, 37(4): 129-143.
- [27] 殷成团, 孙忠滨, 熊梦婕, 等. 基于贝叶斯推断的长江口复合洪水潜在风险[J/OL]. 水科学进展, 2025, 03(06): 1-11.
- [28] Yuille A, Kersten D. Vision as Bayesian inference: analysis by synthesis?[J]. Trends in Cognitive Sciences, 2006, 10(7): 301-308.
- [29] 刘乐平, 袁卫. 现代贝叶斯分析与现代统计推断[J]. 经济理论与经济管理, 2004 (6): 64-69.
- [30] Ke G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision tree[J]. Advances in neural information processing systems, 2017, 30.1-9.
- [31] Yan J, Xu Y, Cheng Q, et al. LightGBM: accelerated genomically designed crop breeding through ensemble learning[J]. Genome Biology, 2021, 22: 1-24.
- [32] Zhang C, Zhao M, Cai M, et al. Multi-stage multi-attribute decision making method based on online reviews for hotel selection considering the aspirations with different development speeds[J]. Computers & Industrial Engineering, 2020, 143: 106421.
- [33] Kaushik V, Khare A, Boardman R, et al. Why do online retailers succeed? The identification and prioritization of success factors for Indian fashion retailers[J]. Electronic Commerce Research and Applications, 2020, 39: 100906.
- [34] You X Y, You J X, Liu H C, et al. Group multi-criteria supplier selection using an extended VIKOR method with interval 2-tuple linguistic information[J]. Expert Systems with Applications, 2015, 42(4): 1906-1916.

- [35] 由丽萍, 王嘉敏. 基于情感分析和 VIKOR 多属性决策法的电子商务顾客满意感测度[J]. 情报学报, 2015, 34(10): 1098-1110.
- [36] Opricovic S, Tzeng G H. Extended VIKOR method in comparison with outranking methods[J]. European Journal of Operational Research, 2007, 178(2): 514-529.
- [37] He S, Wang Y. Evaluating new energy vehicles by picture fuzzy sets based on sentiment analysis from online reviews[J]. Artificial Intelligence Review, 2023, 56(3): 2171-2192.
- [38] 王安宁, 张强, 彭张林, 等. 在线评论的行为影响与价值应用研究综述 [J]. 中国管理科学, 2021, 29 (12): 191-202.
- [39] 邓之宏, 邵兵家. 中国网络团购消费者满意度影响因素研究[J]. 商业研究, 2013 (4): 34-39.
- [40] 张振刚, 罗泰晔. 基于在线评论数据挖掘和 Kano 模型的产品需求分析[J]. 管理评论, 2022, 34(11): 109-117.
- [41] Jelodar H, Wang Y, Yuan C, et al. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey[J]. Multimedia Tools and Applications, 2019, 78: 15169-15211.
- [42] Wu T, Hao F, Kim M. Typical opinions mining based on Douban film comments in animated movies[J]. Entertainment Computing, 2021, 36: 100391.
- [43] 刘伟, 徐鹏涛. O2O 电商平台在线点评有用性影响因素的识别研究——以餐饮行业 O2O 模式为例[J]. 中国管理科学, 2016, 24(5): 168-176.
- [44] Maas A, Daly R E, Pham P T, et al. Learning word vectors for sentiment analysis[C]//Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. 2011: 142-150.
- [45] 杨善林, 李霄剑, 张强, 等. AIGC 的科学基础 [J]. 工程管理科技前沿, 2023, 42 (06): 1-14.
- [46] 朱丹浩, 赵志梹, 张一平, 等. 面向古文自然语言处理生成任务的大语言模型评测研究[J]. 信息资源管理学报, 2024, 14(5): 45-58.
- [47] 陈炫婷, 叶俊杰, 祖璨, 等. GPT 系列大语言模型在自然语言处理任务中的鲁棒性[J]. 计算机研究与发展, 2024, 61(05): 1128-1142.

- [48] Qin C, Zhang A, Zhang Z, et al. Is ChatGPT a general-purpose natural language processing task solver?[J]. arXiv preprint arXiv:2302.06476, 2023.
- [49] 蒲中敏, 张晨曦, 徐泽水. 基于在线评论的消费者偏好挖掘研究综述[J]. 中国管理科学, 2025, 33(1): 209-220.
- [50] Wang Z, Wan M, Cui X, et al. Personalized recommendation algorithm based on product reviews[J]. Journal of Electronic Commerce in Organizations (JECO), 2018, 16(3): 22-38.
- [51] Fan Z P, Che Y J, Chen Z Y. Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis[J]. Journal of Business Research, 2017, 74: 90-100.
- [52] 孙鲁平, 张丽君, 汪平. 网上个性化推荐研究述评与展望[J]. 外国经济与管理, 2016, 38(6): 82-99.
- [53] 杜学美, 丁璟好, 谢志鸿, 等. 在线评论对消费者购买意愿的影响研究[J]. 管理评论, 2016, 28(03): 173-183.
- [54] Ligthart A, Catal C, Tekinerdogan B. Systematic reviews in sentiment analysis: a tertiary study[J]. Artificial Antelligence Review, 2021: 1-57.
- [55] Bhatia P, Ji Y, Eisenstein J. Better document-level sentiment analysis from rst discourse parsing[J]. arXiv preprint arXiv:1509.01599, 2015.
- [56] Volkova S, Wilson T, Yarowsky D. Exploring sentiment in social media: Bootstrapping subjectivity clues from multilingual twitter streams[C]//Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2013: 505-510.
- [57] Weichselbraun A, Gindl S, Scharl A. Using games with a purpose and bootstrapping to create domain-specific sentiment lexicons[C]//Proceedings of the 20th ACM International Conference on Information and Knowledge Management. 2011: 1053-1060.
- [58] 李寿山, 李逸薇, 黄居仁, 等. 基于双语信息和标签传播算法的中文情感词典构建方法[J]. 中文信息学报, 2013, 27(6): 75-81.

- [59] 梁军, 柴玉梅, 原慧斌, 等. 基于深度学习的微博情感分析[J]. 中文信息学报, 2014, 28(5): 155-161.
- [60] El Mrabti S, Jaouad E L M, Hachmoud A, et al. An explainable machine learning model for sentiment analysis of online reviews[J]. Knowledge-Based Systems, 2024, 302: 112348.
- [61] Wu W, Xu M, Su R, et al. Modeling crude oil volatility using economic sentiment analysis and opinion mining of investors via deep learning and machine learning models[J]. Energy, 2024, 289: 130017.
- [62] Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: A survey[J]. Ain Shams Engineering Journal, 2014, 5(4): 1093-1113.
- [63] Pandey A, Vishwakarma D K. Progress, achievements, and challenges in multimodal sentiment analysis using deep learning: A survey[J]. Applied Soft Computing, 2024, 152: 111206.
- [64] Islam M S, Kabir M N, Ghani N A, et al. Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach[J]. Artificial Intelligence Review, 2024, 57(3): 62.
- [65] Singh U, Abhishek K, Azad H K. A survey of cutting-edge multimodal sentiment analysis[J]. ACM Computing Surveys, 2024, 56(9): 1-38.
- [66] 贾甜, 梁一鸣, 黄威. 基于 Kano 模型的原型模式系统需求排序方法研究[J]. 工业工程, 2018, 21(2): 87-93.
- [67] 吴剑锋, 汪圆圆, 矫东芳, 等. 基于 Kano 和 DEMATEL 方法的产品需求排序策略[J]. 机械设计, 2022, 39(02): 138-145.
- [68] 吴江, 李秋贝, 胡忠义, 等. 基于 IPA 模型的乡村旅游景区游客满意度分析[J]. 数据分析与知识发现, 2023, 7(07): 89-99.
- [69] Mohammed M A, Abdulhasan M J, Kumar N M, et al. Automated waste-sorting and recycling classification using artificial neural network and features fusion: A digital-enabled circular economy vision for smart cities[J]. Multimedia Tools and Applications, 2023, 82(25): 39617-39632.

- [70] Yaseen M, Ibrahim N, Mustapha A. Requirements prioritization and using iteration model for successful implementation of requirements[J]. International Journal of Advanced Computer Science and Applications, 2019, 10(1):121-127.
- [71] Liang X, Liu P, Wang Z. Hotel selection utilizing online reviews: A novel decision support model based on sentiment analysis and DL-VIKOR method[J]. Technological and Economic Development of Economy, 2019, 25(6): 1139-1161.
- [72] 李少波, 全华凤, 胡建军, 等. 基于在线评论数据驱动的产品感性评价方法[J]. 计算机集成制造系统, 2018, 24(第 3): 752-762.
- [73] 冯坤, 杨强, 常馨怡, 等. 基于在线评论和随机占优准则的生鲜电商顾客满意度测评[J]. 中国管理科学, 2021, 29(2): 205-216.
- [74] 史丽丽, 林军, 朱桂阳. 基于混合神经网络的中文在线评论产品特征提取及消费者需求分析[J]. 数据分析与知识发现, 2023, 7(10): 63-73.
- [75] 王安宁, 张强, 彭张林, 等. 在线评论的行为影响与价值应用研究综述[J]. 中国管理科学, 2021, 29(12): 191-202.
- [76] Dhenakaran S S, Sambanthan K T. Web crawler-an overview[J]. International Journal of Computer Science and Communication, 2011, 2(1): 265-267.
- [77] hu X, Ilyas I F, Krishnan S, et al. Data cleaning: Overview and emerging challenges[C]//Proceedings of the 2016 International Conference on Management of Data. 2016: 2201-2206.
- [78] Thakur R. Customer engagement and online reviews[J]. Journal of Retailing and Consumer Services, 2018, 41: 48-59.
- [79] Wu Y, Ngai E W T, Wu P, et al. Fake online reviews: Literature review, synthesis, and directions for future research[J]. Decision Support Systems, 2020, 132: 113280.
- [80] 奉国和, 郑伟. 国内中文自动分词技术研究综述[J]. 图书情报工作, 2011, 55(002): 41-45.
- [81] Chu X, Ilyas I F, Krishnan S, et al. Data cleaning: Overview and emerging challenges[C]//Proceedings of the 2016 International Conference on Management of Data. 2016: 2201-2206.
- [82] Hernández M A, Stolfo S J. Real-world data is dirty: Data cleansing and the merge/purge problem[J]. Data Mining and Knowledge Discovery, 1998, 2: 9-37.

- [83] Wu Z, He Q, Li J, et al. Public attitudes and sentiments towards new energy vehicles in China: A text mining approach[J]. *Renewable and Sustainable Energy Reviews*, 2023, 178: 113242.
- [84] Liu Z, Hao H, Cheng X, et al. Critical issues of energy efficient and new energy vehicles development in China[J]. *Energy Policy*, 2018, 115: 92-97.
- [85] Huang L, Ma J, Chen C. Topic detection from microblogs using T-LDA and perplexity. In 2017 24th Asia-Pacific software engineering conference workshops (APSECW)(pp. 71-77)[J]. 2017: 71-77.
- [86] Jo T. Text mining[J]. *Studies in Big Data*, 2019, 45: 1-373.
- [87] Yuan X, Liu X, Zuo J. The development of new energy vehicles for a sustainable future: A review[J]. *Renewable and Sustainable Energy Reviews*, 2015, 42: 298-305.
- [88] Ghassabeh Y A, Rudzicz F, Moghaddam H A. Fast incremental LDA feature extraction[J]. *Pattern Recognition*, 2015, 48(6): 1999-2012.
- [89] Wankhade M, Rao A C S, Kulkarni C. A survey on sentiment analysis methods, applications, and challenges[J]. *Artificial Intelligence Review*, 2022, 55(7): 5731-5780.
- [90] Yin J, Cao X J, Huang X, et al. Applying the IPA–Kano model to examine environmental correlates of residential satisfaction: A case study of Xi'an[J]. *Habitat International*, 2016, 53: 461-472.
- [91] Mao Y, Liu Q, Zhang Y. Sentiment analysis methods, applications, and challenges: A systematic literature review[J]. *Journal of King Saud University-Computer and Information Sciences*, 2024: 102048.
- [92] Patle A, Chouhan D S. SVM kernel functions for classification[C]//2013 International Conference on Advances in Technology and Engineering (ICATE). IEEE, 2013: 1-9.
- [93] Lai Z, Liang G, Zhou J, et al. A joint learning framework for optimal feature extraction and multi-class SVM[J]. *Information Sciences*, 2024, 671: 120656.
- [94] Ke T, Ge X, Yin F, et al. A general maximal margin hyper-sphere SVM for multi-class classification[J]. *Expert Systems with Applications*, 2024, 237: 121647.

- [95] Yang Y, Li Q, Li C, et al. User requirements analysis of new energy vehicles based on improved Kano model[J]. *Energy*, 2024, 309: 133134.
- [96] Li S, Xiao Q. Classification and Improvement Strategy for Design Features of Mobile Tourist Guide Application: A Kano-IPA Approach[J]. *Mobile Information Systems*, 2020, 2020(1): 8816130.
- [97] Joung J, Kim H M. Approach for importance–performance analysis of product attributes from online reviews[J]. *Journal of Mechanical Design*, 2021, 143(8): 081705.
- [98] Jiang X, Zhang J, Yang C, et al. Evaluating the service quality of insular and coastal recreational fisheries by integration of the SERVQUAL-fuzzy Kano model and importance-performance analysis[J]. *Ocean & Coastal Management*, 2023, 243: 106753.
- [99] Cai R, Qin B, Chen Y, et al. Sentiment analysis about investors and consumers in energy market based on BERT-BiLSTM[J]. *IEEE Access*, 2020, 8: 171408-171415.
- [100] Wankhade M, Rao A C S. Opinion analysis and aspect understanding during COVID-19 pandemic using BERT-Bi-LSTM ensemble method[J]. *Scientific Reports*, 2022, 12(1): 17095.
- [101] Gou Z, Li Y. Integrating BERT embeddings and BiLSTM for emotion analysis of dialogue[J]. *Computational Intelligence and Neuroscience*, 2023, 2023(1): 6618452.
- [102] Li W, Du Y, Li X, et al. UD_BBC: Named entity recognition in social network combined BERT-BiLSTM-CRF with active learning[J]. *Engineering Applications of Artificial Intelligence*, 2022, 116: 105460.
- [103] Kousta S T, Vinson D P, Vigliocco G. Emotion words, regardless of polarity, have a processing advantage over neutral words[J]. *Cognition*, 2009, 112(3): 473-481.
- [104] Cao H, Kang J. Study on improvement of recommendation algorithm based on emotional polarity classification[C]//2020 5th International Conference on Computer and Communication Systems (ICCCS). IEEE, 2020: 182-186.

- [105] 陈铁华, 于飞, 李红霞. 基于 KANO-IPA 的矿工安全沟通满意度需求层次研究[J]. 煤炭工程, 2022: 181-187.
- [106] 毕建武, 刘洋, 樊治平. 依据在线评论的商品排序方法 [J]. 系统工程学报, 2018, 33 (03): 422-432.
- [107] Opricovic S, Tzeng G H. Extended VIKOR method in comparison with outranking methods[J]. European Journal of Operational Research, 2007, 178(2): 514-529.
- [108] Fung R Y K, Ren S, Xie J. The prioritisation of attributes in customer requirement management[C]//1996 IEEE International Conference on Systems, Man and Cybernetics. Information Intelligence and Systems (Cat. No. 96CH35929). IEEE, 1996, 2: 953-958.
- [109] Hancock J, Khoshgoftaar T M. Leveraging lightgbm for categorical big data[C]//2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (Big Data Service). IEEE, 2021: 149-154.
- [110] Zhang W, Wu C, Tang L, et al. Efficient time-variant reliability analysis of Bazimen landslide in the Three Gorges Reservoir Area using XGBoost and LightGBM algorithms[J]. Gondwana Research, 2023, 123: 41-53.

附录

附录 1 新能源汽车用户在线评论示例

会员ID	会员名称	品牌	车系	购买车型	裸车购买价	购车地点	动力类型	油耗	电耗	评论时间
278371657	定西车友6340695	比亚迪	秦PLUS	2024款 荣耀版 DM-i 55KM领先型	7.98万	定西	插电式混动	3.4	13.2	2024/4/24
234973549	潍坊车友0293593	比亚迪	秦PLUS	2024款 荣耀版 DM-i 120KM超越型	11.58万	日照	插电式混动	0	16.2	2024/4/23
277333623	孝感车友7548786	比亚迪	秦PLUS	2024款 荣耀版 DM-i 55KM领先型	7.98万	重庆	插电式混动	8	15	2024/4/23
234972481	东营车友6521148	比亚迪	秦PLUS	2024款 荣耀版 DM-i 55KM领先型	7.98万	德州	插电式混动	0	16	2024/4/23
278362734	小飞机	比亚迪	秦PLUS	2024款 荣耀版 DM-i 120KM卓越型	12.58万	东莞	插电式混动	5	14.8	2024/4/23

会员ID	会员名称	评论内容
278371657	定西车友6340695	首先我需要说明一点，我呢不是很懂车，买这台车呢主要是因为我父亲比较偏向于新能源，我哥呢也比较关注比亚迪。
234973549	潍坊车友0293593	我个人觉得混动车挺实用的，又能省钱，跑长途也没问题。
277333623	孝感车友7548786	车漆有点薄，遇到一点不好的路况，小石子溅起来很容易刮花。
234972481	东营车友6521148	车里的隔音效果有点差，在高速开车的情况下有明显的风噪、胎噪音。
278362734	小飞机	秦plus的内饰让我很满意，因为我很喜欢这种设计感。而且车内的颜色搭配也很时尚大气，用料上也比较高级，整体看起来档次感十足。

附录 2 在线评论数据预处理示例

评论内容
首先我需要说明一点，我呢不是很懂车，买这台车呢主要是因为我父亲比较偏向于新能源，我哥呢也比较关注比亚迪。
我个人觉得混动车挺实用的，又能省钱，跑长途也没问题。
车漆有点薄，遇到一点不好的路况，小石子溅起来很容易刮花
车里的隔音效果有点差，在高速开车的情况下有明显的风噪、胎噪音
秦plus的内饰让我很满意，因为我很喜欢这种设计感。而且车内的颜色搭配也很时尚大气，用料上也比较高级，整体看起来档次感十足。

评论文本预处理
说明、懂车、买车、父亲、新能源、关注、比亚迪
混动车、实用、省钱、长途、无问题。
车漆、薄、路况、小石子、刮花
隔音、效果差、高速、风噪、胎噪
秦plus、内饰、满意、设计感、颜色搭配、时尚、大气、用料、高级、档次感

附录 3 在线评论属性划分方法示例

属性	在线评论
舒适性	真皮座椅舒适性还是可以的！
外观	我喜欢它那犀利的前大灯，像是一双炯炯有神的眼睛。车身线条流畅，一看就是那种不好惹的角色。
空间	对后排空间我并不是很满意，并排坐的人多了就稍显紧张，三个大人长途挤下来会很不舒服。
驾驶感受	驾驶感受的话，我觉得这个车的方向盘指向性还是比较好的。
性价比	性价比基本认可，虽然买车没什么优惠，但是很省油。
续航	不得不说，车子的纯电续航还是有所欠缺，也就是跑80公里。
内饰	内饰设计的还是挺经典的，搭配了悬浮式的中控，这种设计会显得更科技一些。
智能化	我感觉智能化吧，没什么挑毛病的，毕竟是一台车，你不可能像手机一样全部功能都有。
操纵	八十迈方向盘就有点飘了。

附录 4 情感分析结果部分示例

属性	在线评论	情感极性
舒适性	后排硬的要死！尤其中间的位置，无处安放的手脚啊！	异常消极
外观	符合我的审美观。	积极
空间	第二排空间中间位置不适合做成年人，尤其是远行的时候。	消极
功能	整车宽敞，该有的功能都有了，一些很细节的功能也考虑到了。	中性
性价比	性价比很高，车子价位不高，关键是电动汽车不用承担高额的汽油费。	异常积极
动力	车子动力果然名不虚传，很强劲，加速也很快。	积极
售后政策	不到十万的入门价格，还有售后政策，有什么理由不去选择呢？	积极
能耗	一般3天充一次电吧 虽然充电桩离家800米 还是懒。	中性

属性	在线评论	情感极性
舒适性	真皮座椅舒适性还是可以的！	正面
外观	我喜欢它那犀利的前大灯，像是一双炯炯有神的眼睛。车身线条流畅，一看就是那种不好惹的角色。	正面
空间	对后排空间我并不是很满意，并排坐的人多了就稍显紧张，三个大人长途挤下来会很不舒服。	负面
驾驶感受	驾驶感受的话，我觉得这个车的方向盘指向性还是比较好的。	正面
性价比	性价比基本认可，虽然买车没什么优惠，但是很省油。	中性
续航	不得不说，车子的纯电续航还是有所欠缺，也就是跑80公里。	负面
内饰	内饰设计的还是挺经典的，搭配了悬浮式的中控，这种设计会显得更科技一些。	中性
智能化	我感觉智能化吧，没什么挑毛病的，毕竟是一台车，你不可能像手机一样全部功能都有。	中性
操纵	八十迈方向盘就有点飘了。	负面

攻读硕士学位期间发表或录用的学术论文

发表论文

[1]盛志远,沈超,董丰,等.基于 Kano-IPA 模型的新能源汽车产品需求优先级排序方法[J].佳木斯大学学报(自然科学版),2024,42(09):61-65.

申请发明专利

[1]基于 Kano-IPA 模型的新能源汽车产品需求优先级排序方法和系统（已实审，申请号：202420837358.3）

参与科研项目

安徽省自然科学基金青年项目《新兴信息技术环境下基于产品大数据的客户需求获取与融合方法研究》（2008085QG336）参与人。

致谢

在完成本篇硕士论文之际，回望这段求学之路，心中百感交集。三年的时光转瞬即逝，留下的是知识的积淀与成长脚步，更是无数关怀与帮助的温暖。在此，我谨向所有在我学习与生活中给予我指导、支持与鼓励的师长、亲人和朋友们致以最诚挚的谢意。

首先，我要衷心感谢我的导师沈超老师。在整个研究生阶段，沈老师不仅在学术上给予我深入而细致的指导，更在生活中给予我真挚的关心与鼓励。沈老师严谨治学的态度、渊博的学识和宽厚仁爱的品格深深影响了我。论文从选题到完成，沈老师始终耐心指导、精益求精，每一次讨论都让我受益匪浅。是沈老师的悉心教诲让我学会了科研的思维方式，也让我更加坚定了求学与探索的信心。在此，谨向沈老师致以我最深的敬意和衷心的感谢！

同时，我也要特别感谢课题组的张文刚老师、丁一老师和吕伟老师。在课程学习、论文写作及科研过程中，三位老师都给予了我极大的帮助与启发。他们不仅在学术方面为我提供了宝贵的建议，更在生活上给予我亲切关怀。三位老师严谨认真的治学精神和耐心细致的指导让我深感敬佩，令我在学术之路上不断砥砺前行。感谢你们在我成长路上的默默付出与悉心引导。

此外，我要特别感谢我的父母。你们一直是我坚实的后盾，是我不断追梦的力量源泉。正是你们无条件的支持和理解，才让我得以心无旁骛地投入到学习与科研中。你们的爱与鼓励，是我前行道路上最温暖的灯塔。

我要感谢我亲爱的师兄姐妹们，正是你们的陪伴、鼓励与帮助，使我在研究生生活中倍感温暖与充实。

最后，感谢学校和学院提供的良好学术氛围与学习环境，感谢 901 实验室，使我顺利完成各项研究任务。

这段求学旅程将成为我人生中最宝贵的回忆。我将带着这份感恩之心，继续在未来的人生道路上努力前行，不负众望，勇攀高峰。