

现代情报
Journal of Modern Information
ISSN 1008-0821, CN 22-1182/G3

《现代情报》网络首发论文

题目：基于主题挖掘和情感分析的在线健康社区用户评论研究
作者：郭羽婷，姚宣合
网络首发日期：2024-12-11
引用格式：郭羽婷，姚宣合. 基于主题挖掘和情感分析的在线健康社区用户评论研究[J/OL]. 现代情报. <https://link.cnki.net/urlid/22.1182.G3.20241210.1728.002>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于主题挖掘和情感分析的在线健康社区用户评论研究

Research on User Reviews of Online Health Community Based on Topic Mining and Sentiment Analysis

郭羽婷* 姚宣合

Guo Yuting* Yao Xuanhe

(东北师范大学信息科学与技术学院, 吉林 长春, 130117)

(Department of Information Science and Technology, Northeast Normal University, Changchun, 130117, China)

摘要: [目的/意义] 在线健康社区为用户提供线上健康服务, 分析其用户评论的潜在信息, 对于医疗服务质量的提高和健康社区信息建设的优化具有重要意义。[方法/过程] 本文提出了一个在线健康社区用户评论分析模型。首先通过 LDA 主题模型挖掘患者评论的主题; 然后使用分类模型对患者评论进行主题分类; 最后通过词频筛选、TF-IDF 关键词提取以及 SO-PMI 方法构建领域情感词典, 计算各个主题的患者评论文本的情感得分, 分析不同情感倾向的评论信息。[结果/结论] 通过分析“好大夫在线”综合性三甲医院的用户评论数据进行实证研究, 根据实验结果的信息内容和规律, 提出了改进医疗服务和信息建设的相关参考建议。

关键词: 健康社区; 用户评论; 主题挖掘; 文本分类; 情感分析

分类号: G353

Abstract: [Objective/Significance] Online health communities provide online health services for users, and analyzing the potential information of user comments is of great significance for improving the quality of medical services and optimizing the information construction of healthy communities. [Method/Process] An online health community user comment analysis model was proposed. Firstly, the LDA topic model was used to dig the topic of patient comments. Then the classification models were used to classify the patient comments into different topics. Finally, word frequency screening, TF-IDF and SO-PMI method were used to construct a domain sentiment dictionary which was employed to calculate the sentiment score of patient comments in each topics. And comment information of different sentiments were further analyzed. [Result/Conclusion] Empirical research is conducted by analyzing the user comment data of General Grade Three hospital from "Good Doctor Online". According to the information content and rules of the experimental results, relevant reference suggestions for improving medical service and information construction are put forward.

Key words: healthy communities; user reviews; topic mining; text classification; sentiment analysis

在科技发展、政策支持以及民众就医习惯改变的背景下, 中国医疗健康互联网化趋势明显, 以移动健康管理以及在线健康社区为代表的互联网健康实践发展迅猛^[1]。据第 54 次《中国互联网络发展状况统计报告》显示, 截至 2024 年 6 月, 我国互联网医疗用户规模达 3.65 亿人, 占网民整体的 33.2%^[2]。在线健康社区是用户进行健康信息交流、经验分享、问答咨

基金项目: 国家自然科学基金项目“基于多目标优化的 MOOCs 课程视频摘要生成框架研究”(项目编号: 62107009); 东北师范大学校内青年基金项目项目“基于深度特征表示和注意力机制的视频摘要框架研究”(项目编号: 2109213)。

作者简介: 姚宣合 (1996 年-), 女, 硕士, 研究方向: 文本挖掘。

通信作者: 郭羽婷 (1985 年-), 女, 博士, 讲师, 研究方向: 文本挖掘、智慧教育。

询的开放式网络平台^[3]，为公众快捷地获取和利用健康信息提供了极大的便利。患者评论是在线健康社区信息的重要内容之一，能够充分体现患者在诊疗过程中的真实感受；且患者是在线健康社区使用的直接用户，因此，挖掘当前在线健康社区用户发表评论的规律，分析患者评论信息的内涵和结构，对于提升医疗服务质量、优化健康社区信息组织建设具有重要价值。

本文提出了一个在线健康社区用户评论分析模型，对患者评论文本进行主题挖掘和情感分析。以“好大夫在线”综合性三甲医院患者评论数据为例，基于 LDA 方法挖掘患者关注的主题，采用机器学习方法，较为精确地对用户评论进行主题分类，并构建健康社区领域情感词典，分析不同主题下患者评论的情感倾向。根据负面评论主题内容回溯、评论信息的潜在规律，从患者、健康社区用户角度出发，提出改善医疗服务、优化在线健康社区信息组织和建设的参考建议。

1 相关研究

随着我国互联网医疗的快速发展，许多健康社区平台都推出了患者评论功能，如“好大夫在线”“39 健康网”“微医网”等^[4]，患者在医院接受了医疗服务后，可以通过在线健康社区平台分享就医感受。患者评论是由已就诊的患者或患者家属撰写，通过相关手机应用或者网页发表于在线健康社区平台上的评论，是在线评论的一个分支。患者评论的形式包括文字、图片、语音等，内容包括就诊体验、感受、实况等。患者评论作为一种医疗评论信息，对解决传统医疗服务环境中的信息不完全和信息不对称问题起着至关重要的作用^[5]，为其他用户提供了择医、就医、愈后相关的参考信息。患者评论具有实时性强、易于获取、成本低等特点，蕴含巨大的研究潜力。然而健康社区并未对海量患者评论内容进行挖掘和特征标识，导致用户若想参考此类评论，需要耗费大量的时间和精力阅读评论内容。为了挖掘在线健康社区用户评论蕴含的医疗服务和健康信息，相关研究者分别从情感分析、主题挖掘等方面展开研究。

情感分析是自然语言处理研究领域中的一个重要的研究方向，主要研究如何从文本中发现或挖掘人们对于某种事物、产品或服务所表达出的情感或意见^[6]。高慧颖等^[7]根据在线患者评论文本具有专业性强、差异性大、非规范性的特点，提出了基于特征加权词向量的在线患者评论情感分析方法。该方法利用 Word2vec 构建词向量模型，抽取情感词集合，完善医疗服务领域情感词典，有效地获知在线医疗评论的情感倾向。Rahim A I 等^[8]对影响马来西亚医院 Facebook 评论中表达积极情感的决定因素，以及医院认证与 Facebook 评论中表达的

情感之间的关系进行实证研究。首先构建情感分类器和服务质量分类器,自动对评论的情感和 SERVQUAL 维度进行分类,利用逻辑回归分析来探究影响表达积极情感的决定因素。王辉等^[9]通过 SOPMI 算法扩充基础情感词典构建领域情感词典来分析“好大夫在线”健康社区中重大慢病患者评论的情感倾向,分析重大慢病患者评论负面倾向的重要影响因素。熊回香等^[10]就如何从海量针对医生服务的用户评论中挖掘出医生服务的相关特征,并使用 BiLSTM 模型获取用户对目标医生每一服务特征类别的情感分值,结合权重计算得出医生服务的评价。

主题挖掘是指对文本内容进行整理、分析与逻辑推断,从而分析出文本主题内涵的过程^[11]。Shah A M 等^[12]基于双因素理论分析在线医疗评论,探讨了患者对英国医疗服务的满意度和不满意度。该研究基于 SentiNet 和 LDA 相结合的文本挖掘方法来揭示患者评论的语义,发现了高风险和低风险疾病类别中的潜在主题。周欢等^[4]基于社会网络视角,分析在线健康社区评论的有用性。首先,使用 TF-IDF、TextRank 以及 LDA 方法分析患者评论的主题分布;然后,将各类患者评论文本分别转换为文本关联网,使用社会网络分析方法进一步分析患者评论的有用性特征。文献[13]和文献[14]采用情感分析和主题挖掘相结合的方式对健康社区用户评论进行分析。叶艳等^[13]以“好大夫在线”平台上的高血压患者评论作为研究数据,采用 LDA 主题模型得到 13 个患者评论主题,再结合相关文献将得到的患者评论主题合并为 6 个医疗服务质量主题,然后根据 BiLSTM 模型得到各服务质量主题的情感倾向分布,并分析产生负面情感的原因。余佳琪等^[14]在 LDA 主题模型的基础上融入情感信息与时间信息,对甜蜜家园糖尿病患者的评论数据进行挖掘。考虑了评论数据的时间属性对主题与情感可能产生影响,识别疾病确诊阶段、并发症阶段等共 7 个时间片内的评论热点主题与伴生的情感强度及类型,揭示了主题、情感随时间演化的特征。

综上所述,已有健康社区用户评论相关的研究分别基于情感分析和主题挖掘对医疗评价、情感倾向影响因素、特殊疾病患者关注主题、信息有用性分析、主题和情感演化特征等方面展开工作。在相关研究中,将主题挖掘和情感分析相结合来研究患者评论的工作较少。文献[13]分析了不同服务质量主题的情感倾向分布,但文献[13]未考虑同一评论可能涉及不同的主题,而将 LDA 主题模型中概率最大的主题作为该评论的主题,文档所属主题类别归类方式较为粗略。文献[14]在 LDA 主题模型的基础上融入情感信息,但文献[14]侧重于分析评论数据的时间属性对主题与情感可能产生的影响、主题情感随不同时间段演化的特征情况。且文献[13]和[14]针对特定群体(高血压患者和糖尿病患者)的在线评论进行研究,仅仅分

析了某种特殊疾病群体的情感和关注主题。为了突破现有工作的局限性，本文提出一种针对患者评论分析的主题挖掘与情感分析框架。首先，采用 LDA 主题模型对患者评论进行主题挖掘；其次，将患者评论切分为子句序列，通过人工进行主题类别标注；第三，使用人工标注的患者评论数据训练机器学习分类模型，再使用分类模型对未标注的评论子句进行较高精度的主题分类；最后，构建领域情感词典，使用领域情感词典对各个主题下的患者评论文本进行情感分类。在实证研究中，本文对“好大夫在线”平台上北京、上海的 5 家综合性三甲医院的患者评论进行分析，探讨在不同主题下患者的情感倾向，挖掘当前在线健康社区用户发布评论的信息和规律。

2 研究设计

本文提出了一种基于 LDA 主题挖掘和情感分析的在线健康社区患者评论分析框架，以中文在线健康社区的患者评论为数据源，分析患者关注的主题内容和情感倾向。本文的研究框架如图 1 所示。

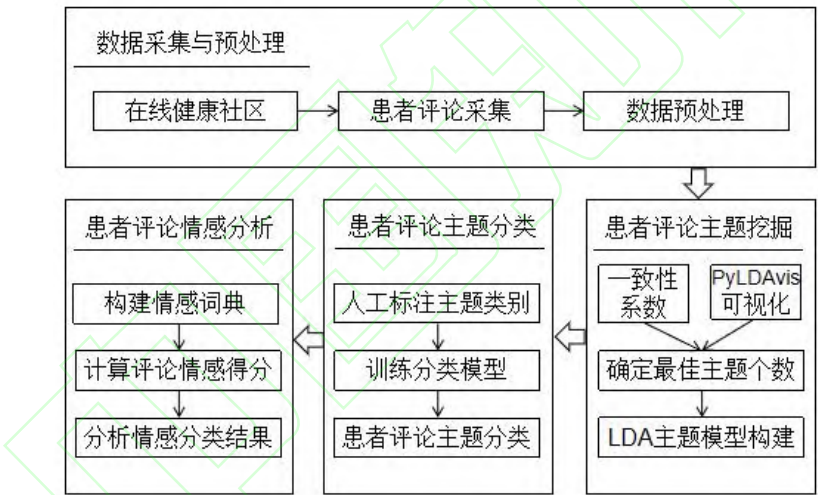


图 1 研究框架

Fig.1 Research Framework

2.1 评论数据的采集与预处理

2.1.1 患者评论采集

本文使用后羿采集器软件，采集在线健康社区平台上的患者评论数据。首先，先浏览多个在线健康社区上患者发表的相关评论内容，主要考虑患者评论的数量；其次，确定目标在线健康社区平台，获取该平台上患者发表评论的页面地址；最后，设计后羿采集器的爬虫模式，在爬虫的目标网页中点击选中所需的字段内容，并制定后羿采集器采集网页信息的页面加载规则，采集患者评论数据。

2.1.2 数据的预处理

在线评论文本包含 URL、表情等噪声数据，且为无结构文本形式。如不加以处理，可能会影响后续主题挖掘的结果，以及文本分类和情感分类的准确性。因此，需要对在线评论进行数据预处理^[15]。数据预处理是指对采集到的原始数据进行清理、加工，去除具有干扰性的无效信息，将原始数据转化为完整、一致的形式。本文的数据预处理方法包括数据剔除、分词、去除停用词 3 个步骤。使用 Jieba 分词工具下的精确模式对患者评论文本进行分词。

2.2 患者评论主题挖掘

LDA (Latent Dirichlet Allocation) 模型是由 Blei D M 等于 2003 年提出的一种文档集（或语料库）的生成式概率模型^[16]，其基本思想是将文档表示为潜在主题的随机混合，其中每个主题都以单词分布为特征。LDA 模型采用词袋模型降低文本维度，通过无监督方式挖掘“文档—主题”分布和“主题—词”分布，在文本挖掘中常用于识别大规模、非结构化文档集（或语料库）中的隐藏主题信息。LDA 主题模型原理如图 2^[16]所示。

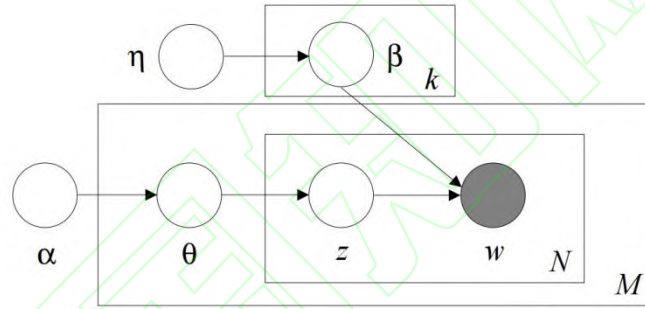


图 2 LDA 模型

Fig.2 LDA Model

M 表示文档的总数、 N 表示文档包含词总数， K 是主题数。 α 和 η 是两个参数。 θ 表示文档 m 的主题分布， $m = 1, 2, \dots, M$ ； β 表示第 k 个主题的词分布， $k = 1, 2, \dots, K$ 。由参数为 α 的狄利克雷分布采样生成 θ ， $\theta \sim \text{Dir}(\alpha)$ ；由参数为 η 的狄利克雷分布采样生成 β ， $\beta \sim \text{Dir}(\eta)$ 。文档和主题之间服从多项式分布，主题和词之间也服从多项式分布。由参数为 θ 的多项式分布采样生成主题 z ， $z \sim \text{Multinomial}(\theta)$ ；由以 z 、 β 为条件的多项式分布采样生成词 w 。然后如此循环直到生成一个文档，并最终生成 M 个文档。最终生成文档与原始文档一致性越高，模型的性能越好。LDA 模型的联合分布如公式 (1)^[16] 所示：

$$p(\theta, z, w | \alpha, \beta) = p(\theta | \alpha) \prod_{n=1}^N p(z_n | \theta) p(w_n | z_n, \beta) \quad (1)$$

最佳主题数 K 可根据计算一致性来确定，一致性最大时，所对应的主题数为最佳主题，一般情况下，当一致性不再增加或处于拐点时的主题数为最优。

2.3 患者评论主题分类

2.3.1 文本分类的数据处理

经过对患者评论文本进行主题挖掘，得到主题挖掘的结果。然而，在患者的一条评论中，往往会包含多个主题信息。这种现象在较长的评论中尤为常见。而在较短的评论子句中，表达的主题通常会比较单一。因此，在本文的主题分类中，为了提高分类的准确性，需要把每条评论分割为多个子句。在汉语的表达中，除了顿号之外，其他的标点符号通常用于将前后表达的不同内容分开。因此，本篇论文将逗号、分号、句号、感叹号、问号作为每条患者评论的切割点^[20]，按照切割点对患者评论进行切割，把每条患者评论分割成为多条子句，如表 1 所示。

表 1 切割子句示例

Tab.1 Cutting Sentence Samples

评论文本	评论文本 ID 值	切割后的子句	子句 ID 值
对病人很耐心，很关照病人的情绪，手术做得非常好，恢复快，术后平整漂亮。	1	对病人很耐心	1
		很关照病人的情绪	1
		手术做得非常好	1
		恢复快	1
		术后平整漂亮	1

2.3.2 人工标注

随机抽取 10% 的评论子句，按照主题挖掘的结果对子句进行主题类别的人工标注。为了保证人工标注的客观性，由两位研究人员在互不干扰的情况下，根据挖掘出的主题特征，对样本子句所属的主题类别进行标注。人工标注完成后，使用 SPSS 软件计算标注结果的 Kappa 系数，从而判断人工标注的可靠性。计算得到的 Kappa 系数值若是在可靠的范围内，则用其进行下一步研究；计算得到的 Kappa 系数值未在可靠的范围之内，则重新归纳总结挖掘出的主题特征，再次进行人工标注和计算人工标注结果的 Kappa 系数值，直至其数值在可靠的范围内。此外，使用 Python 数据分析工具对两位标注者的标注结果做差异对比，找出标注结果不同的评论子句。两位标注者针对标注结果不一致的评论子句进行协商讨论，最终使用意见一致的标注结果作为本篇论文实验的数据。

2.3.3 训练分类模型

本篇论文使用 Python Sklearn 库中的 CountVectorizer 包对人工标注的子句评论文本进行特征提取，分类器采用 Sklearn 库中的决策树模型、支持向量机模型、K 近邻模型以及朴素贝叶斯模型。本文用训练集训练 4 种分类模型，用测试集测试 4 种模型的分类效果，以准确

率、精准率、F1-Score 这 3 个指标作为评估分类模型的评价指标，采用 5 折交叉验证策略调参。

选择最优分类模型对其余未人工标注的子句进行主题分类后，将归属于同一个主题下的子句，按照相同的子句 ID 值重新将其拼接，如图 3 所示，再对各个主题的评论进行归类汇总，能更客观地、可靠地分析患者对于不同主题评论的信息。

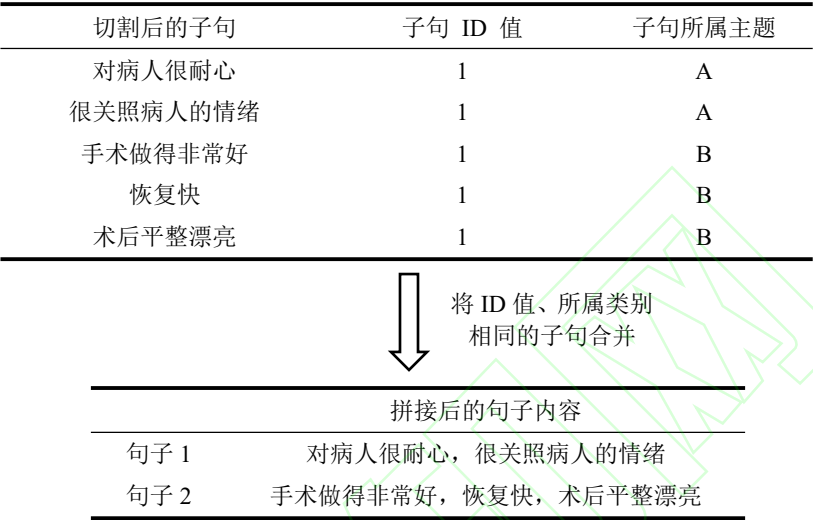


图 3 合并子句示例图

Fig.3 Merging Short Sentences Sample

2.4 患者评论文本情感分析

2.4.1 构建领域情感词典

首先，将知网的 HowNet 情感词典和 NTUSD 情感词典合并作为基础情感词典^[17]；其次，对患者评论语料文本按照前文介绍的数据预处理方法进行分词和过滤停用词处理，观察后总结出在该领域中，情感词出现频率较高的词性；然后，对评论文本进行 TF-IDF 关键词提取，并同时添加词性筛选条件，获得词性为形容词、名词、动词、副词的关键词^[21]，并对关键词进行人工筛选，确定种子情感词；最后，使用 SO-PMI 方法确定其他词语与种子情感词的关联程度，设定关联阈值，从而筛选出候选情感词，将种子情感词与候选情感词合并加入基础情感词典中，构建领域情感词典。

2.4.2 计算患者评论情感得分

将正面情感词权重设定为 1，负面情感词权重设定为-1。程度词划分为 6 个级别，从小到大依次赋予的分值为：0.5、0.8、1.2、1.5、1.7、2。当评论中出现“没有”“不”等否定

词时，否定词会颠倒整条语句的情感倾向，并且每出现一次否定词，情感倾向就会发生一次反转。因此，将否定词权重设定为-1，情感得分的计算公式如公式（2）、公式（3）所示：

$$score_{\text{分句}} = \sum (-1)^n \bullet w_a \bullet w_s \quad (2)$$

$$score_{\text{总句}} = \sum score_{\text{分句}} \quad (3)$$

其中， w_a 表示程度词权重， w_s 表示情感词权重， n 为否定词出现的次数，总句指被计算情感得分的评论语句，分句指根据分割符号将总句分割后得到的子句。

3 实证研究

3.1 数据准备

我国比较知名的在线健康社区有“好大夫在线”“春雨医生”“39健康网”“丁香园”等。其中，“好大夫在线”平台于2006年创建，目前已有18年的历史，具备丰富的健康信息数据和较为成熟的运行模式。该平台收录了全国10 083家医院905 175位医生数据，且医生在该平台上的活跃度较高^[18]。文本挖掘研究需要大量的数据支持，而综合型三甲医院的病患相对较多，其患者评论数据也较为丰富。因此，本文选取了北京和上海的5家三甲综合性医院（包括：北京朝阳医院、北京同仁医院、上海第六人民医院、上海瑞金医院、上海新华医院）在“好大夫在线”平台上的患者评论作为研究的数据来源。

使用后裔采集器对这5所医院在2020年10月1日—2022年9月22日的患者评论数据进行采集，共获得40 293条评论文本。评论前面存在“看病描述：”“看病过程：”“康复状况：”3个提示词，这是“好大夫在线”平台标准的患者评论内容描述，因此将这些词语删除。进行分词和去停用词处理后，用词频统计方法观察文本预处理的效果，发现其中“医生”“医院”等词出现的频数排在前列。“医生”“医院”等词在患者评论中的出现频率虽然很高，但是不具备实际意义，对信息挖掘帮助不大，因而将词频高且无特殊性的词添加到停用词表中，再次使用过滤停用词的方法删除评论文本中的这类词语。

3.2 患者评论的主题挖掘

本文通过计算LDA模型在不同主题数量下的一致性分值来确定最佳的主题个数。当LDA主题模型的主题个数为4时，如图4所示，一致性分数达到最高，为0.6576。同时，运用pyLDAvis将主题个数为4的LDA主题模型结果可视化，主题的4个气泡分布均匀，并且无重叠现象，表明将LDA模型主题个数设置为4较为合理。

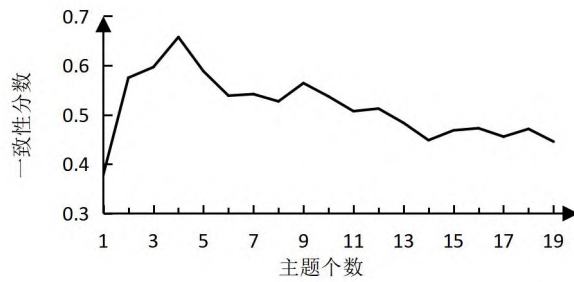


图 4 一致性系数

Fig.4 Consistency Coefficient

LDA 模型提取到各个主题下的主题词属于某个主题的概率值越大，表明该主题词属于这一主题的可能性越大。同一主题下主题词之间的排序也是由概率值而定，概率越大，主题词排序越靠前。表 2 展示了 4 个主题的前 15 个高概率主题词。其中，在 T1 主题下的 15 个高频词主要是描绘医者的医术技能水平、态度和品德，因此将 T1 主题命名为“医术医德”；T2 主题下的 15 个高频词中“治疗”“诊断”“方案”等词与看病的过程有关，“用药”“效果”“好转”“痊愈”等词与用药效果有关，因此将 T2 主题命名为“用药诊断”；T3 主题下的 15 个高频词主要是描述手术及术后情况，因此将 T3 主题命名为“手术”；T4 主题下的 15 个高频词中出现“挂号”“预约”“检查”等词，说明 T4 主题与流程有关，“手术室”“地方”等词表示 T4 主题与环境有关，因此将 T4 主题命名为“环境流程”。

表 2 患者评论 LDA 主题分布

Tab. 2 The LDA Topic Distribution of Patient Reviews

主题	前 15 个高概率词汇
医术医德 (T1)	耐心、医术、态度、详细、医德、负责、精湛、解答、高尚、专业、热情、着想、理解、关心、技术
用药诊断 (T2)	治疗、效果、方案、吃药、病因、好转、诊断、医嘱、用药、症状、痊愈、药物、保守、控制、判断
手术 (T3)	手术、恢复、顺利、康复、成功、出院、很快、疼痛、良好、伤口、切除、活动、愈合、一流、锻炼
环境流程 (T4)	时间、地方、挂号、预约、方便、检查、等待、平台、手术室、环境、排队、地方、报告、加号、检查

3.3 患者评论的文本分类

从 40 293 条患者评论文本中随机抽取 20 200 条评论进行下一步研究。按照 2.3.1 介绍的文本分类数据处理方法对 20 200 条患者评论进行处理后，得到 51 002 条评论子句。

将 51 002 条评论子句随机打乱顺序后，抽取其中 10%的数据进行人工标注。人工标注

完成后，使用 Kappa 系数进行人工标注数据的一致性校验，SPSS 软件计算 Kappa 系数的结果为 0.796，根据一致性参考指标（Kappa \geq 0.8 为十分可靠，Kappa \geq 0.69 较为可靠）^[19]，说明本文人工标注的数据结果较为可靠。

分类模型采用决策树、朴素贝叶斯、支持向量机和 K 近邻，使用随机搜索和网格搜索相结合的策略寻找 4 个分类模型的最优参数并训练分类模型。为了减少评估结果的随机性，使用 5 折交叉验证的方式，最终结果取准确率、精准率、F1-Score 的平均值，如表 3 所示。通过观察 4 个分类模型 3 个指标的平均得分，发现支持向量机的 3 个评价指标的平均得分都是最高。因此，选择支持向量机作为患者评论文本的分类模型。

表 3 分类器分类结果

Tab. 3 The Classifier Classification Results

分类模型名称	准确率	精准率	F1-Score
决策树	89.37%	89.09%	0.890 9
朴素贝叶斯	89.44%	89.34%	0.891 1
支持向量机	91.42%	91.31%	0.912 3
K 近邻	85.48%	85.69%	0.843 3

使用训练好的支持向量机分类模型对未人工标注的评论进行分类。然后将文本 ID 值和所属主题类别相同的子句进行合并。合并后，得到 4 个主题的患者评论共计 29 767 条。其中，提及“医术医德”主题的评论文本有 16 001 条；提及“用药诊断”主题的评论文本有 5 584 条；提及“手术”主题的评论文本有 5 297 条；提及“环境流程”主题的评论文本有 2 885 条。4 个主题的患者评论比例如图 5 所示。

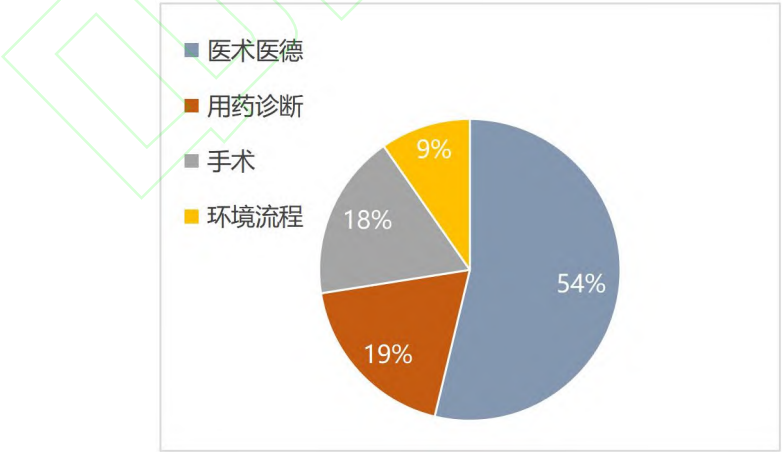


图 5 各个主题的患者评论占比

Fig.5 The Proportion of Patient Reviews for Each Subject Category

由图 5 可知，患者发表的评论中，“医术医德”主题评论的占比最高，“用药诊断”和“手术”主题评论的占比紧随其后，“环境流程”主题的评论占比是最低的。因此，“医术

医德”是患者真实关注度最高的主题。该主题的评论大多数是描绘医方态度的，态度方面容易给患者留下较为深刻的印象，尤其是有不少患者在评论中表示，医生作为专家，却没有架子，使患者感到亲近感和尊重。患者在就医之后，治疗效果可能需要一段时间才能呈现。大多数用户习惯在就诊后不久发表评论，这个时间段发表的评论内容可能只有对医生态度的感受。“用药诊断”和“手术”两个主题下的评论占比相差不大，原因之一可能是在实际的治疗过程中，用药和诊断与手术是紧密联系的，因此提及率类似。此外，“好大夫在线”平台的患者评论发表的页面中，该平台设置了“看病过程：”“康复情况：”等提示，因此，患者发表关于“环境流程”类的评论较少，从而导致该主题的评论占比较低。

3.4 患者评论的情感分类

3.4.1 领域情感词典的构建结果

HowNet 情感词典包含 4 566 个正面情感词，4 370 个负面情感词。NTUSD 情感词典包含 2 810 个正面情感词，8 276 个负面情感词。首先将 HowNet 和 NTUSD 情感词典合并、去重得到基础情感词典，然后对基础情感词典进行观察，发现在基础负面情感词表中，存在“疾病”“病弱”等词汇，这些词汇在医疗服务领域中不具备负面情感。据统计，其中共有 71 个词汇在医疗服务领域不具备负面情感，因此将这些词从基础负面情感词表中删除。最终得到的基础情感词典中共包含 7 176 个正面情感词，11 991 个负面情感词。使用 TF-IDF 在患者评论语料库中提取正面种子情感词 389 个，负面种子情感词 68 个，然后采用 SO-PMI 方法确定语料库中其他词语与种子情感词的关联程度，获得 129 个正面候选情感词，111 个负面候选情感词。最终经过去重后，形成医疗服务领域情感词典，如表 4 所示。

表 4 医疗服务领域情感词典构建情况

Tab. 4 The Construction of Emotional Dictionary in Medical Service Fields

情感词典	正面情感词	负面情感词
基础情感词典	7 176	11 991
种子情感词	389	68
SO-PMI 获取词汇	129	111
去重形成领域情感词典	7 694	12 170

3.4.2 情感分类结果

用上述构建好的医疗服务领域情感词典，计算 4 个主题下患者评论的情感得分。将情感得分大于 0、小于 0、等于 0 的患者评论分别划分为正面评论、负面评论和中性评论。各个主题下的患者评论情感分类占比结果如图 6 所示。结果发现，4 个主题下的患者评论均为正

面评论居多，中性评论和负面评论占比较低，并且中性评论的占比大于负面评论占比。其中，“用药诊断”和“手术”主题类的正面评论、中性评论、负面评论占比类似。“医术医德”主题类的正面评论占比位居 4 个主题的首位，高达 98.37%，“环境流程”主题类的正面评论占比最低，为 77.44%，且该主题的负面评论占比最高，为 10.12%。

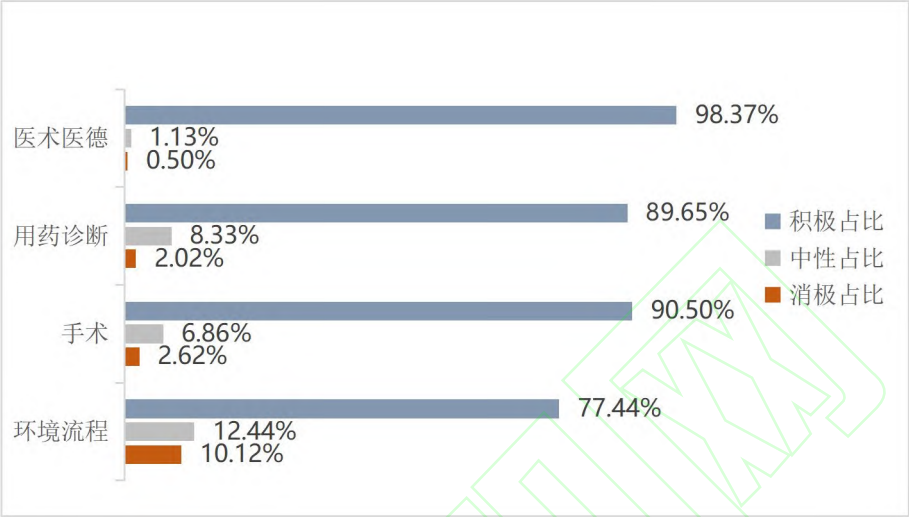


图 6 各个主题下的患者评论情感占比

Fig.6 The Proportion of Emotional Patient Comments Under Each Topic

3.4.3 情感分类结果分析

对上述 4 个主题的正面评论、负面评论文本进行词云图展示，展示结果如图 7~图 10 所示。评论的情感极性是对该评论包含情感词综合衡量后的最终判定，但在评论文本中可能同时包含了负面情感词和正面情感词，因此，词云图中的单独关键词本身的极性可能并不能代表其所在评论的情感极性。根据下文词云图中关键词，回溯到相应的评论原文中，归纳出患者发表关于各类主题的正面评论、负面评论的原因。



图 7 “医术医德”评论词云

Fig.7 The “Medical Skill and Medical Ethics” Comment Word Cloud

观察“医术医德”的正面评论词云图可知，“医术”“医德”“态度”等词占据词云图中较大的面积。根据这些词定位回评论原文中，发现“医术”主要是与“精湛”“高超”等词相连，称赞医生医术好。“医德”主要是与“高尚”等词相连，赞美医护人员的品德好。在“医术医德”的负面评论词云图中，“态度”一词仍然占据词云图中较大的面积，意味着“态度”在该类主题的评论文本中是比较重要的，并且也分为待人态度和工作态度。在“医术医德”类的负面评论文本中，形容医护人员的待人态度的词主要有“不耐烦”“恶劣”“冷漠”等词。形容医护人员的工作态度的词主要有“不负责任”“打发”“草草了事”等。



图8 “用药诊断”评论词云

Fig.8 The “Medication and Diagnosis”Comment Word Cloud

在“用药诊断”主题下，观察正面评论词云图，得到“问诊”“详细”“耐心”“效果”等关键词，用上述关键词定位回“用药诊断”类的正面评论文本，发现在用药方面，让患者满意的主要是医生能精准用药，且用药效果良好，患者病情得到好转。在诊断方面，让患者满意的原因有：医生详细地询问患者情况；医生仔细检查和看片子；能快速且准确找到病因；耐心地给患者分析讲解病情；制定适合的治疗方案；给予患者合理的建议和指导。观察“用药诊断”主题下的负面评论的词云图，基于图中“检查”“看病”“治疗”“不适”等词，定位到该主题的负面评论文本中，归纳总结出患者发表关于“用药诊断”主题不满意的缘由。对于用药方面，患者不满意的原因有：医生开的药物过多，患者认为已经属于过度医疗；药物治疗效果差，患者病情无好转迹象，甚至出现严重的不良反应；医生开的药物贵，且不在可报销药物之列。对于诊断方面，患者不满意的原因有：医生看病敷衍，问诊不详细，甚至没有问诊；医生不给患者解释病情；医生给出的治疗方案不合格；就诊的时间太急促，患者没有充足的时间与医生交流；医生无法确诊病因，甚至给出错误诊断；患者挂了号，医生不接诊，没有向患者提供其所需的治疗。

观察“环境流程”主题的正面评论词云图发现：“安排”“回复”“住院”“手术”“网上”“顺利”等词占据正面评论词云图的大部分面积。回溯定位到“环境流程”主题的正面评论文本中，总结对于环境方面的发表正面评论原因主要有：医院干净卫生；环境安静，适合养病；医院设备先进齐全；饭菜清淡可口有营养，适合患者恢复健康。总结对于流程方面的正面评论原因主要包括：网上挂号快捷，预约方便；医生帮忙加号；检查过程高效；网上诊室回复问题及时，减少病人往返跑的次数；安排手术、检查、住院等事宜迅速；流程简单、有序、紧凑、速度快；设置医导人员指导患者就诊；医院整体秩序管理到位；主动为经济困难的患者申请慈善基金。患者发表负面评论的原因包括：排队时间太长，等待时间久；患者人太多；患者挂不上号；患者就诊需要多次往返跑，很麻烦；医疗机器设备有故障；病床不足；患者不知道如何咨询相关事项，就诊指南不明确；医院环境差，卫生条件堪忧；就诊程序复杂，手续烦琐；秩序混乱，管理不到位；节假日医院人手不足；加号费用太贵；只能线上挂号，不能线下挂号，导致老年人无法挂号。

4 建议策略

根据实验结果的信息内容和规律，分别从医院医疗服务视角、健康社区信息组织视角提出针对三级甲等医院和在线健康社区提高服务质量的相关建议。

医院医生方面：①关注患者情绪变化，加强人文关怀。医方应关注患者情绪变化，及时给予患者必要的包容和关爱，加强对患者门诊治疗的心理关注与疏导。加大对疾病、药物、就诊过程、医生常规操作的宣传，尽量减少由于信息不对等患者产生的消极情绪。②增设业务办理场所，分散业务办理，加大资金投入，确保设备供应，设置温馨导航，院内进一步设置清晰的路标，明示就医流程，增加导诊服务人员，缩短患者等待时间。③鼓励开展线上诊疗和线下相结合。利用线上诊疗提供有效的患者咨询反馈渠道，及时帮助患者解决问题和困扰，为患者带来极多便利。

在线健康社区方面：①判断患者发布评论信息的情感倾向，为在线健康社区用户构建用户画像，并对情绪特征鲜明的患者进行有针对性的情感支持。②将患者评论按照主题进行分类，利于患者和在线健康社区更好地利用评论信息。③为患者提供个性化的检索服务，推出基于主题的智能查询功能，从而辅助患者快速找到自身所需的主题健康信息，提高健康信息检索的效率与查询的质量。

5 总结

本文首先基于 LDA 主题模型对患者评论进行主题挖掘，得到“医术医德”“手术”“环

境流程”“用药诊断”4个主题；然后，采用支持向量机对患者评论文本进行分类，将评论文本划分到4个主题下；最后，使用词性筛选、TF-IDF和SO-PMI的组合方法构建了医疗服务领域情感词典，用该词典对患者评论进行情感分析，得到各个主题下不同情感倾向的患者评论文本。对4个主题的正面评论、负面评论文本进行词云图展示，根据词云图中关键词，回溯到相应的评论原文中，归纳出患者发表各类主题正面评论、负面评论的原因。利用“好大夫在线”平台上三甲医院的患者评论数据进行实证研究，根据对负面评论及相关结论的分析，提出了帮助在线健康社区改进医疗服务和帮助在线社区改进信息建设的相关参考建议。

参 考 文 献

- [1] 孟猛, 尤剑, 刘晨晖, 等. 健康信息采纳行为研究——概念界定、理论模型与未来展望[J]. 现代情报, 2024, 44(6): 157-167.
- [2] 中国互联网络信息中心. 第54次《中国互联网络发展状况统计报告》[EB/OL]. <https://www.cnnic.net.cn/NMediaFile/2024/0911/MAIN1726017626560DHICKVFMS6.pdf>. 2024-10-25.
- [3] 翟冉冉, 韩正彪. 在线健康社区用户信息交互的驱动因素研究[J]. 图书情报工作, 2024, 68(3): 15-26.
- [4] 周欢, 刘嘉, 张培颖, 等. 复杂网络视角下在线健康社区评论有用性研究[J]. 情报科学, 2022, 40(9): 88-97.
- [5] 李昂, 赵志杰. 基于信号传递理论的在线评论有用性影响因素研究[J]. 现代情报, 2019, 39(10): 38-45.
- [6] Messaoudi C, Guessoum Z, Ben Romdhane L. Opinion Mining in Online Social Media: A Survey[J]. Social Network Analysis and Mining, 2022, 12(1): 25. [\[LinkOut\]](#)
- [7] 高慧颖, 公孟秋, 刘嘉唯. 基于特征加权词向量的在线医疗评论情感分析[J]. 北京理工大学学报, 2021, 41(9): 999-1005.
- [8] Rahim A I, Ibrahim M I, Musa K I, et al. Assessing Patient-Perceived Hospital Service Quality and Sentiment in Malaysian Public Hospitals Using Machine Learning and Facebook Reviews[J]. International Journal of Environmental Research and Public Health, 2021, 18(18):9912.
- [9] 王辉, 王晓玉, 顾东晓, 等. 在线健康社区重大慢病患者负面评论倾向的关键影响因素分析[J/OL]. 情报科学, 1-19[2024-12-05]. <http://kns.cnki.net/kcms/detail/22.1264.g2.20240129.0942.010.html>.
- [10] 熊回香, 向瀛泓, 孟璇, 等. 在线健康社区中基于用户评论挖掘的医生服务评价研究[J/OL]. 情报科学, 1-19[2024-12-05]. <http://kns.cnki.net/kcms/detail/22.1264.G2.20240126.1758.004.html>.
- [11] 祁瑞华, 付豪. 新冠肺炎疫情相关政策扩散特征与主题挖掘研究[J]. 智库理论与实践, 2021, 6(6): 55-64.
- [12] Shah A M, Yan X B, Tariq S, et al. What Patients Like or Dislike in Physicians: Analyzing Drivers of Patient Satisfaction and Dissatisfaction Using a Digital Topic Modeling Approach[J]. Information Processing & Management, 2021, 58(3): 102516.
- [13] 叶艳, 吴鹏, 周知, 等. 基于LDA-BiLSTM模型的在线医疗服务质量识别研究[J]. 情报理论与实践, 2022, 45(8): 178-183, 168.
- [14] 余佳琪, 赵豆豆, 刘蕤. 在线健康社区慢性病患者评论主题情感协同挖掘研究——以甜蜜家园为例[J]. 数据分析与知识发现, 2023, 7(10): 95-108.

- [15] 庞庆华, 董显蔚, 周斌, 等. 基于情感分析与 TextRank 的负面在线评论关键词抽取[J]. 情报科学, 2022, 40(5): 111-117.
- [16] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet Allocation[J]. Journal of Machine Learning Research, 2003, 3(4/5): 993-1022.
- [17] 吴琦, 李阳. 融入领域风险词典的社会安全事件网络舆情风险评估研究[J]. 情报理论与实践, 2024, 47(6): 175-183.
- [18] 马骋宇. 在线医疗社区医患互动行为的实证研究——以好大夫在线为例[J]. 中国卫生政策研究, 2016, 9(11): 65-69.
- [19] Carletta J. Assessing Agreement on Classification Tasks: The Kappa Statistic[J]. Computational Linguistics, 1996, 22(2): 249-254.
- [20] 徐海燕, 姜瑛. 针对复杂用户评论的代码质量属性判断[J]. 软件学报, 2021, 32(7): 2183-2203.
- [21] 邓淑卿, 李玩伟, 徐健. 基于句法依赖规则和词性特征的情感词识别研究[J]. 情报理论与实践, 2018, 41(5): 137-142.