

基于文本挖掘的在线医疗平台综合评价研究

李小红, 刘梦丹, 蒋俊

华北电力大学, 数理学院, 北京

收稿日期: 2023年2月27日; 录用日期: 2023年3月23日; 发布日期: 2023年3月30日

摘要

随着防疫政策的放开, 传统医疗资源有限, 在线医疗一定程度上成为人们“病有所医”的保障。然而, 我国目前主流的在线医疗平台存在产品创新不足、服务质量不高、各类产品同质化严重等问题, 导致可能无法给用户提有建设性有效的医疗建议。随着互联网技术的广泛应用以及人们对便利的医疗服务的需要, 完善在线医疗平台的服务是十分有必要的。影响用户对在线医疗平台的满意度的因素是什么? 在线医疗平台如何有针对性地完善平台功能? 用户如何选择合适的在线医疗平台? 这些都是当下需要解决的问题。本文基于从网络上爬取的主流在线医疗平台的评论文本, 运用文本分析技术、聚类算法、层次分析法、情感分析等技术搭建满意度模型, 对影响用户对平台满意度的因素进行研究, 为平台提高自身功能和服务, 以及为用户选择在线医疗平台提供建议和参考。

关键词

在线医疗平台, 文本分析, 聚类算法, 层次分析法, 情感分析, 满意度模型

Research on Comprehensive Evaluation of Online Medical Platform Based on Text Mining

Xiaohong Li, Mengdan Liu, Jun Jiang

School of Mathematics and Physics, North China Electric Power University, Beijing

Received: Feb. 27th, 2023; accepted: Mar. 23rd, 2023; published: Mar. 30th, 2023

Abstract

With the release of epidemic prevention policies, traditional medical resources are limited, and online medical treatment has become a guarantee for people to have access to medical services to

some extent. However, China's mainstream online medical platforms have problems such as insufficient product innovation, poor service quality and serious homogenization of various products, which may lead to the inability to provide users with constructive and effective medical advice. With the wide application of Internet technology and people's need for convenient medical services, it is very necessary to improve the service of online medical platform. What factors affect users' satisfaction with online medical platforms? How to improve the function of online medical platform? How do users choose the right online medical platform? These are the issues that need to be addressed right now. Based on the comments of mainstream online medical platforms crawled from the Internet, this paper uses text analysis technology, clustering algorithm, analytic hierarchy process, emotion analysis and other technologies to build a satisfaction model, and studies the factors that affect users' satisfaction with the platform, so as to provide suggestions and references for the platform to improve its own functions and services and for users to choose online medical platforms.

Keywords

Online Medical Platform, Text Analysis, Clustering Algorithm, Analytic Hierarchy Process, Sentiment Analysis, Satisfaction Model

Copyright © 2023 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 引言

后疫情时代，疫情时有反复，加之防疫政策放开，传统医疗资源紧缺的情况或时有发生。在传统医疗不能满足人们看病就医的需求时，在线医疗成为大众接受医疗服务的理想之选。在线医疗可以合理利用医疗资源为人们提供优质的医疗服务，且能够有效预防医患因近距离接触产生交叉感染的问题，一定程度上成为大众“病有所医”的保障。但目前，我国在线医疗平台产品创新不足、服务质量不高、各类产品同质化严重。因此，如何有针对性地提高用户满意度，是在线医疗平台各大管理者需解决的重点问题。本文运用数学方法深入探究影响用户对在线医疗平台满意度评价的因素，有利于为用户选择在线医疗平台提供参考，有助于管理者有针对性地提高在线医疗平台的服务质量，从而进一步为人们的看病就医提供便利和保障。

所谓文本挖掘，是抽取有效的、新颖的、可理解的、散布在文本文件中的有价值的知识，并且利用这些知识更好地组织信息的过程。通过阅读相关文献我们了解到，一部分学者挖掘海量的用户在线评论文本数据后，对数据进行处理，并将其应用于信息提取、情感分析和文本分类三个方面。

洪玲[1]以“平安好医生”APP为例，对爬取的该APP的评论文本数据进行文本分析，筛选出评论中的高频词，得到用户对该APP满意度评价的影响因素。陈婷[2]以移动医疗APP的在线评论文本作为切入点，采用文本挖掘技术和内容分析法探究影响用户满意度评价的因素，构建移动医疗APP用户服务满意度模型。

学者们对满意度的研究都较为深入，丁甜甜[3]以覆盖“互联网+医疗健康”全产业链的数字健康平台微医为例，通过用户网络评论分析在线医疗现状，研究在线医疗用户满意度影响因素，以期为了解用户需求、提升医疗服务质量与用户满意度、提高在线医疗效果提供参考。周露莎[4]对“好大夫在线”的数据进行研究，取出患者满意度的主题维度，计算每个医生各主题维度的满意度得分。

4.2. TF-IDF 算法

TF-IDF 是一种评估一个字词在一个文件集, 或者在一个语料库中的其中一份文件的重要程度的统计方法。一般情况下, 我们认为, 词语的词频越大, 则词语在文本中的重要程度越高[6]。

首先, 我们对分词后的高频词的词频进行归一化的处理:

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (1)$$

其中, $n_{i,j}$ 是该词语在文本 d_j 中出现的次数, $\sum_k n_{k,j}$ 则是文本 d_j 中所有词语出现的次数总和。

$$\text{词频: } TF_w = \frac{\text{某词语 } w \text{ 在文本中出现的次数}}{\text{文本的总次数}} \quad (2)$$

继而, 衡量一个词语是否为常见词时, 若该词语在日常使用中相对少见, 但在文本中多次出现, 则它很可能反映了文本的特性, 可视为文本的关键词。基于该思路, 我们采用逆文档频率分配高频词权重, 即对于较常见的词语分配较小的权重; 对于较少见的词语分配较大的权重。

为避免词语在文本中不存在而导致公式出错的情况, 计算时一般使用 $|\{j:t_i \in d_j\}|+1$ 计算, 即:

$$idf_i = \log \frac{|D|}{|\{j:t_i \in d_j\}|+1} \quad (3)$$

其中, $|D|$ 为所爬取评论的文件总数, $|\{j:t_i \in d_j\}|$ 为包含词语 t_i 的文件数目。

$$\text{逆文档频率 (IDF)} = \log \left(\frac{\text{总文档数}}{\text{包含该词语的文档数}+1} \right) \quad (4)$$

由此, 可计算出高频词的 TF-IDF 值:

$$\text{TF-IDF 值} = \text{TF (词频)} * \text{IDF (逆文档频率)} \quad (5)$$

词语的 TF-IDF 值与其在所有爬取的评论文本文档中出现的次数呈正相关, 与其在语言使用中的次数呈负相关。

将文本中的词语转换为词频矩阵, 其中, 矩阵元素 $a[i][j]$ 表示 j 词在 i 类文本下的词频。对于每一个训练文本, 只考虑每个词汇在该训练文本中出现的频率。将文本中的词语转换为词频矩阵后, 计算各个词语出现的次数。继而, 统计每个词语的 TF-IDF 权值。将文本转为词频矩阵并计算 TF-IDF 后, 再将 TF-IDF 矩阵抽取出来, 元素 $a[i][j]$ 表示 j 词在 i 类文本中的 TF-IDF 权重。然后, 输出词典以及文本向量。

由所得结果可知, 每一行的数值的平方相加等于 1, 由该结果我们初步得到应给各个高频词语分配的权重数值。

4.3. 语义网络知识图谱

利用 jieba 库中的 posseg 对高频词进行词性标注, 筛选出其中的名词并对其进行初步分类, 将该分类作为后续高频词划分的参考。

语义网络知识图谱中, 每一个数据模型可以看成是由节点和边构成的有向图。通常情况下, 用实体表达图中的节点, 用关系表达图中不同实体间的关系, 即属性。

基于清晰文本后筛选出的高频词, 绘制相应的语义网络图(见图 2)。

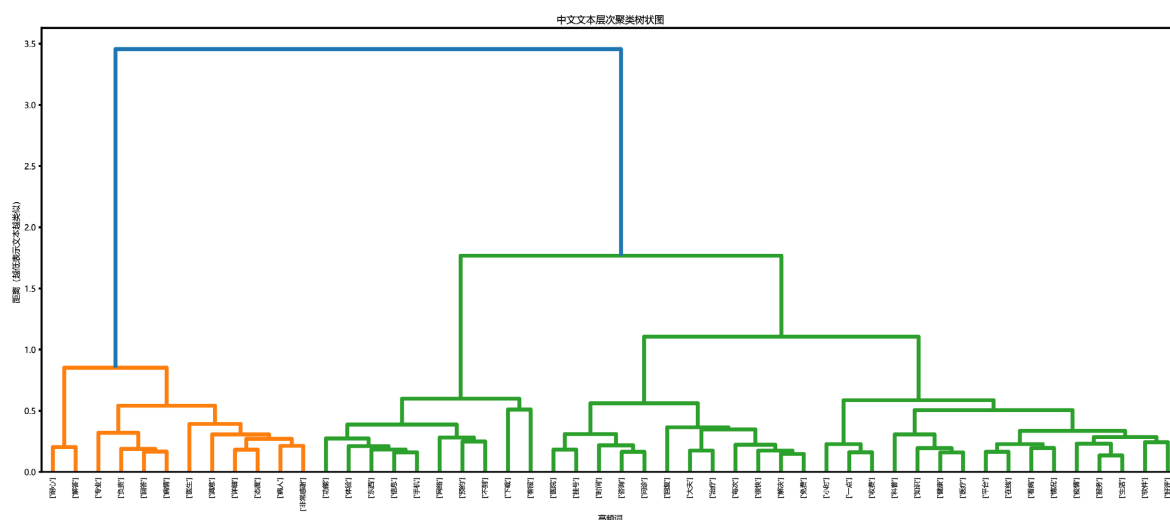


Figure 3. Visualization of clustering results

图 3. 聚类结果可视化

Table 1. Preliminary clustering results

表 1. 初步聚类结果

类别	高频词
1	耐心、解答、专业、负责、回答、病情、医生、满意、 详细、态度、病人、非常感谢
①	功能、体验、东西、信息、手机、网络、预约、不到、 下载、客服
2	② 医院、挂号、时间、咨询、问诊、回复、大夫、治疗、 每次、很快、解决、免费
③	小时、一点、收费、科普、知识、健康、医疗、平台、 在线、看病、情况、疫情、服务、生活、软件、好评

4.5. LDA 模型

LDA 是一种包含词、主题和文档三层结构的文档主题生成模型,该模型是一种无监督的贝叶斯模型,经过推导可得一个链式关系:

$$P(\text{词}|\text{文档}) = P(\text{词}|\text{主题})P(\text{主题}|\text{文档}) \quad (6)$$

某个词在同一主题中出现的概率,以及某个主题在同一文档中出现的概率,两者相乘,即可得到某篇文档出现某个词的概率。

我们通过多次的调试,根据困惑度变化情况,并观察抽取的各个主题的主题词的混淆情况来确定最优主题数目 K 值[2]。通过主题数目 K 依次取值 3、4、5、8 来训练模型时,发现当 K 值取 4 时,模型效果最为理想。

利用 LDA 主题模型对高频词进行聚类,得到的可视化结果(见图 4):

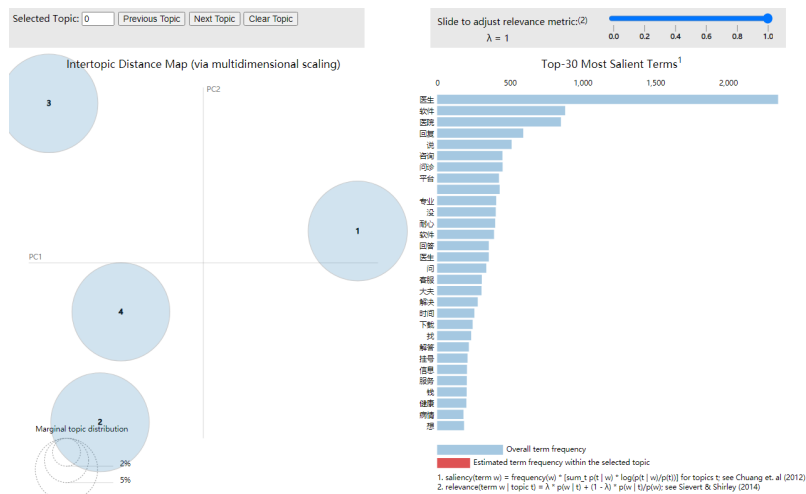


Figure 4. Visualization of LDA model clustering results
图 4. LDA 模型聚类结果可视化

通过分析 LDA 模型的结果可知，主题 1 的高频词，即“问诊”“客服”“挂号”“医疗”等，主要体现出用户对平台方便性及其功能的关心；主题 2 的高频词，即“医生”“平台”“专业”“态度”等，主要体现用户对医生专业能力的重视；主题 3 的高频词，即“医生”“服务”“解决”“病情”等，主要体现用户对医生态度即能力的重视；主题 4 的高频词，即“软件”“医院”“资讯”“耐心”等，主要体现用户对软件平台咨询有效性和客服关怀的关注。

4.6. K-means 算法聚类

自然语言理解转化为机器学习的第一步通常是将文本信息数字化。本文通过 word2vec 获得词语的低维实数向量，向量的每个维度都代表了该词的一个潜在特征，而该特征捕获了有用的句法及语义信息。通过 word2vec 模型得到筛选出的 50 个高频词的词向量。

文本特征聚类是数据挖掘领域的一个重要应用，它根据文本的相似性，将相似度高的文本划分为一个簇，各个簇之间关联程度较小，各个簇内部关联程度较大，因此可实现文本的快速检索，从而帮助检索者快速有效地检索出所需的文本信息。本文采用 K-means 算法对文本进行聚类。

K-means 算法将数据点到原型的某种距离作为优化的目标函数，利用函数求极值的方法得到迭代运算的调整规则。我们采用欧式距离作为相似度测度。

欧氏距离： n 维欧式空间中的点 X 可表示为 (x_1, x_2, \dots, x_n) ，空间中的点 A, B 可分别表示 (a_1, a_2, \dots, a_n) ， (b_1, b_2, \dots, b_n) ，则 A, B 间的欧氏距离为

$$d(AB) = \sqrt{\sum (a_i - b_i)^2} \tag{7}$$

算法采用误差平方和准则函数作为聚类准则函数。

误差平方和：

$$SSE = \sum_{i=1}^k \sum_{x \in S_i} (d(c_i - x))^2 \tag{8}$$

其中， d 是欧氏空间中两个点的欧氏距离， k 是簇的个数， x 是数据点， c_i 是第 i 个簇的中心， S_i 表示第 i 个簇中数据点的集合。

算法的基本步骤[8]为：① 从 N 个文档中随机选取 K 个文档作为质心；② 对其余的每个文档，测量其到每个质心的距离，将其归并到距离最小的质心的类；③ 重新计算已得到的类的质心；④ 迭代②~③步直至新的质心与原质心相等或小于指定阈值，算法结束。

基于上述 K-means 算法步骤，结合词向量模型，我们利用之前选出的 50 个高频词，通过计算轮廓系数，确定最佳聚类个数。轮廓系数用于评价聚类效果的好坏，越接近 1，则聚类效果越好[9] (见图 5)。

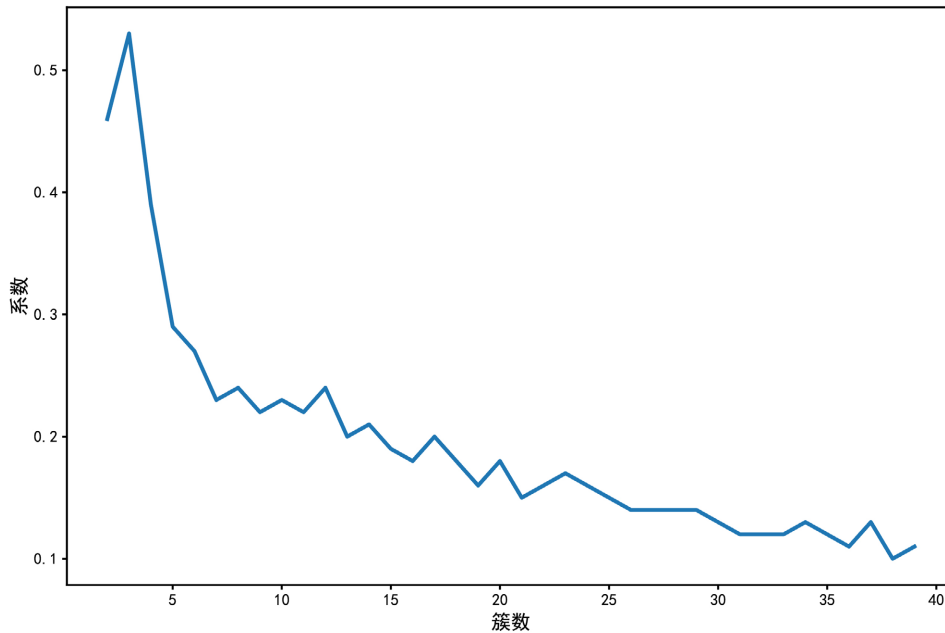


Figure 5. K-means algorithm contour coefficient

图 5. K-means 算法轮廓系数

观察上图可知，轮廓系数在聚类簇数为 3 或 4 时达到最大值，故分别选择聚类簇数为 3 和 4 对高频词进行聚类。通过对比发现，聚类簇数为 4 是聚类效果较理想。相应的可视化结果如下(见图 6)：

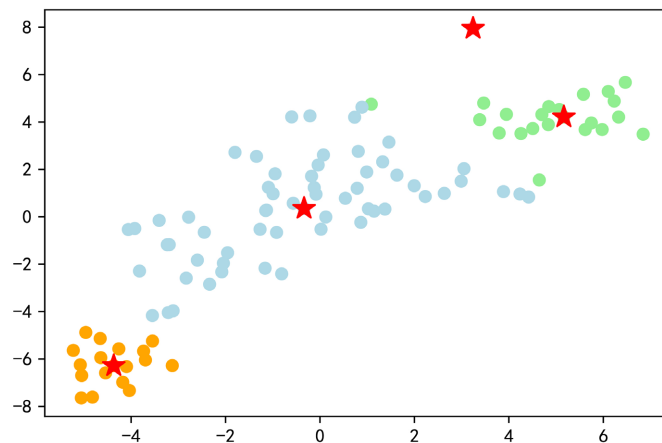


Figure 6. Visualization results of clustering cluster number = 4 by K-means algorithm

图 6. K-means 算法聚类簇数 = 4 的可视化结果

分析上图可明显看出,图中左下角存在一离群点,故可考虑将 50 个高频词划分为三个类别。该结果可作为高频词分组的参考。

4.7. 层次分析法

层次分析法[6]是将与决策有关的元素分解成目标、准则、方案等层次,在此基础上进行定性和定量分析的决策方法。该方法将决策问题按总目标、各层子目标、评价准则、具体方案的顺序分解为不同的层次结构,继而求解判断矩阵特征向量,求出每一层次的各个元素对上一层次的某元素的优先权重,最后再运用加权和方法,递阶归并各方案对总目标的最终权重,最终权重最大者的方案即为最优方案。

具体建模步骤如下:

① 建立层次结构模型。将决策的目标、决策准则、决策对象按照其相互关系分为最高层、中间层和最低层,绘出层次结构图。其中最高层为决策的目的,即要解决的问题;中间层为决策时的备选方案;最低层为决策的准则。

② 构建判断矩阵。层次分析法中构造判断矩阵的方法是一致矩阵法,即将所有因素两两相互比较。此时采用相对尺度(见表 2),以求尽可能减少性质不同因素相互比较的困难,提高准确度。

Table 2. Determine the scale of matrix elements

表 2. 判断矩阵元素标度

判断矩阵元素标度	
标度	含义
1	两因素相比,两者同样重要
3	两因素相比,其中一个比另一个稍微重要
5	两因素相比,其中一个比另一个明显重要
7	两因素相比,其中一个比另一个强烈重要
9	两因素相比,其中一个比另一个极端重要
2、4、6、8	上述两相邻判断的中值
倒数	A 和 B 相比,若标度为 3,则 B 和 A 相比标度为 1/3

③ 层次单排序及其一致性检验。对于判断矩阵最大特征根 λ_{\max} 的特征向量,将向量中各元素进行归一化处理,将特征向量记为 W 。定义一致性指标 $CI = \frac{\lambda - n}{n - 1}$, $CI = 0$, 有完全的一致性; CI 接近于 0, 有满意的一致性; CI 越大,不一致越严重。为了衡量 CI 的大小,引入随机一致性指标 RI (见表 3)。

Table 3. Average random consistency index RI

表 3. 平均随机一致性指标 RI

平均随机一致性指标 RI															
n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RI	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49	1.52	1.54	1.56	1.58	1.59

定义一致性比率 $CR = \frac{CI}{RI}$ ，一般认为一致性比率 $CR < 0.1$ 时，认为判断矩阵的不一致程度在容许范围之内，通过一致性检验，可用其归一化特征向量 W 作为权向量。

④ 层次总排序及其一致性检验。计算某一层次所有因素对于最高层(总目标)相对重要性的权值，称为层次总排序。

层次总排序的一致性比率为 $CR = \frac{a_1CI_1 + x_2CI_2 + \dots + a_mCI_m}{a_1RI_1 + a_2RI_2 + \dots + a_mRI_m}$ ，当 $CR < 0.1$ 时，认为层次总排序通过一致性检验。

基于以上层次结构模型构造成对比较矩阵。对于各级指标，对该层下面的各方案进行成对比较，并按照相对的重要性量化评级。通过对同一准则下的方案层指标出现的频数大小差异进行对比，来衡量每个方案的相对重要性。每一个具有向下隶属关系的元素作为判断矩阵的第一个元素，位于左上角，隶属于它的各个元素依次排列在其后的第一行和第一列。

判断矩阵如下：

两个一级指标的判断矩阵：

一级指标	医生	软件
医生	1	5
软件	1/5	1

“医生”维度判断矩阵：

医生	专业性	医德医风	患者信任感
专业性	1	5	3
医德医风	1/5	1	1
用户信任感	1/3	1	1

“软件”维度判断矩阵：

软件	有用性	可靠性	软件性能	费用	客服关怀
有用性	1	5	7	3	5
可靠性	1/5	1	5	3	3
软件性能	1/7	1/5	1	1/5	1/3
费用	1/3	1/5	5	1	3
客服关怀	1/5	1/3	3	1/3	1

层次单排序是指每一个判断矩阵各因素针对其准则的相对权重，所以本质上是计算权向量。在实际中要求判断矩阵满足大体上的一致性，需通过计算一致性比例 CR ，并根据平均随机一致性指标 RI 进行判断，进行一致性检验。

最终, 分别利用算术平均法、几何平均法和特征值法计算权重, 可以得到二级指标相对一级指标、一级指标相对于总的准则的权重(此处取算术平均法得到的权重, 保留三位小数)。

其中, 软件部分的各个二级指标具体含义为:

- 有用性: 软件使用是否便捷, 软件是否能提供附加服务如预约疫苗等, 以及软件是否能提供有用的科普知识和文章等;
- 可靠性: 软件是否能有效保护用户隐私, 软件提供的服务是否值得信赖, 软件是否值得用户推荐给他人使用等;
- 软件性能: 软件下载是否流畅, 软件画面是否美观, 软件使用时网络是否流畅等;
- 费用: 在软件上进行挂号、买药、资讯所需的费用是否合理;
- 客服关怀: 软件客服回复态度是否良好, 回复是否及时。

各指标权重如下图所示(见表 4):

Table 4. Weights of primary and secondary indexes

表 4. 一级指标、二级指标权重

要素层	权重	指标层	权重	综合权重
医生	0.833	医生专业性	0.659	0.549
		医德医风	0.156	0.13
		患者满意度	0.185	0.154
软件	0.167	有用性	0.461	0.077
		可靠性	0.256	0.043
		性能	0.042	0.007
		费用	0.157	0.026
		客服关怀	0.084	0.014

由上表可知, 影响用户满意度的最主要因素为医生的专业性及软件的有用性, 相比较而言, 软件的客服关怀及软件性能对用户满意度影响较低。因此, 在线医疗平台在进一步升级 APP 时, 可着重提高医生的专业性及软件的有用性。

4.8. 情感分析

由于用户评价中所包含的态度通常不是完全积极或完全消极, 因此, 为得到在线医疗平台的用户满意度评分, 我们不能简单计算评分结果的均值。机器计算的结果是基于词语字面表达的态度(积极或消极)得来的, 仅依据少量的文字表达难以得出理想的具体评判分值。为此, 我们建立评价模型, 该模型可以包容情感分析带来的文字到数值转化的误差。

4.8.1. 情感分组

对主流在线医疗平台的评论进行情感分析, 首先需要对评论进行情感分组。人工评定情感分值后, 利用机器学习对余下评论进行打分。通过构建情感词典的方式对在线评论文本数据进行情感分析, 并将情感倾向性词语分为积极、中肯和消极倾向三类[10]。继而, 我们计算三组评价数在中评价数中的占比, 计算结果见表 5。

Table 5. The proportion of positive, pertinent and negative comments on mainstream online medical platforms
表 5. 主流在线医疗平台积极、中肯、消极评论数占比

在线医疗平台	积极评论数	比例	中肯评论数	比例	消极评论数	比例
京东健康	58	55.77%	0	0	46	44.23%
平安健康	115	46.56%	3	1.21%	129	52.23%
春雨医生	442	80.34%	2	0.45%	106	19.27%
医鹿	287	62.26%	4	0.87%	170	36.88%
好大夫在线	259	59.40%	6	1.38%	171	39.22%
快速问医生	629	91.42%	5	0.73%	54	7.85%
丁香医生	552	70.86%	19	2.44%	208	26.70%

相比于其他在线医疗平台，“快速问医生”的积极评论占比最大，其次是“春雨医生”；而差评比例最大的医疗平台是“平安健康”。该结果表明，“快速问医生”和“春雨医生”两个医疗平台用户满意度较高，而“平安健康”的用户满意度最低，需要继续完善平台服务。

4.8.2. 各在线医疗平台各二级指标情感分组

利用人工标记类别的评论训练模型，并用模型余下评论数进行分类。经整理后，得到每个在线医疗平台二级指标的评论数量及比例。下表为“丁香医生”各类的评论数量及比例(见表 6):

Table 6. The proportion of positive, pertinent and negative comments on Lilac Doctor
表 6. 丁香医生积极、中肯、消极评论数占比

丁香医生(779 条评论)						
二级指标	积极评论	比例	中肯评论	比例	消极评论	比例
专业性	73	94.81%	1	1.30%	3	3.89%
医德医风	42	95.45%			2	4.55%
患者满意度	57	78.08%			16	21.92%
有用性	237	92.22%	1	0.39%	19	7.39%
可靠性	86	77.48%	1	0.90%	24	21.62%
性能	27	30.00%	11	12.22%	52	57.78%
费用	24	26.67%	5	5.55%	61	67.78%
客服关怀	5	13.51%			32	86.49%

下图 7 为本文所研究的 7 个主流在线医疗平台各类评论数目的条形图:

从条形图可知，“快速问医生”和“丁香医生”的评论数目最多且积极评论数目占比较大，“春雨医生”的积极评论数目占比也明显高于其中肯评论数和消极评论数；相比之下，“医鹿”和“好大夫在线”的积极评论数占比较低。

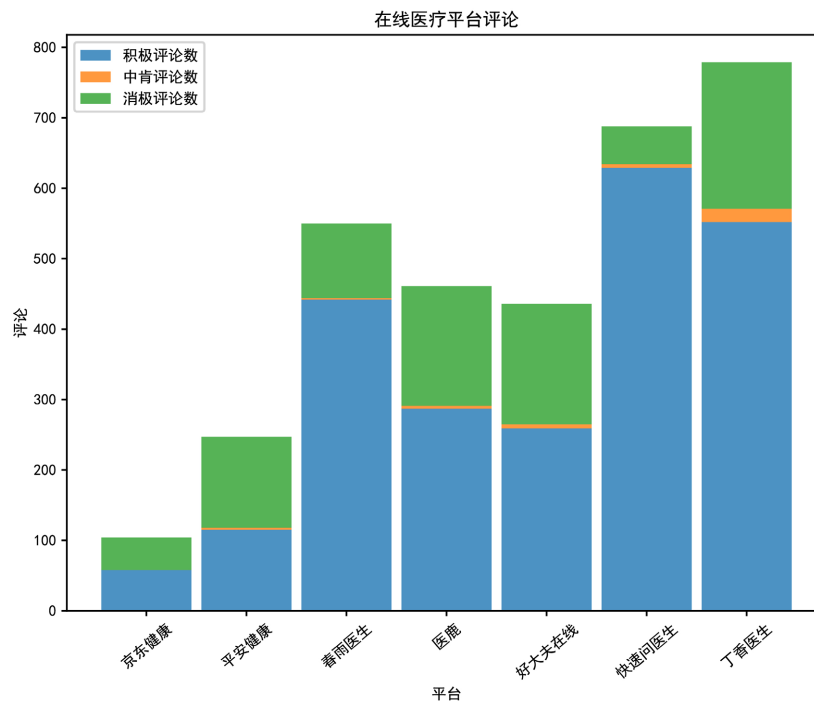


Figure 7. Bar chart of reviews of each type for each medical platform
图 7. 各个医疗平台各类型评论数条形图

下图 8 为本文所研究的 7 个主流在线医疗平台各属性相关的评论数条形图:

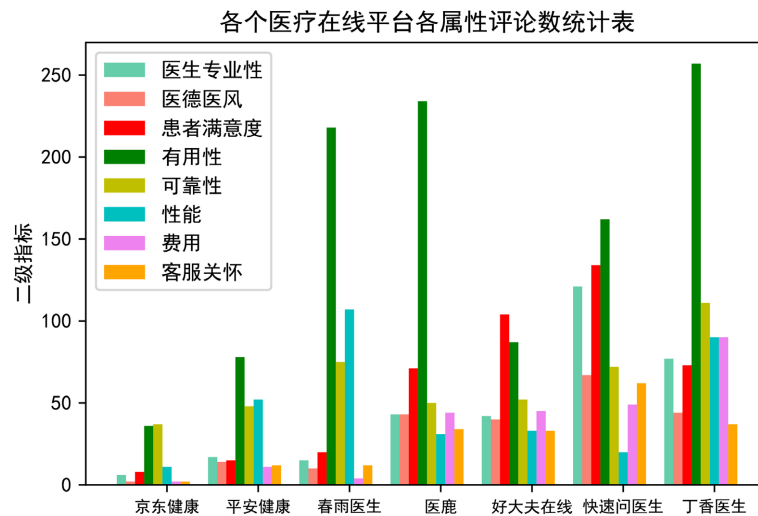


Figure 8. Bar chart of the number of comments for each attribute of each medical platform
图 8. 各个医疗平台各属性评论数条形图

从条形图可明显看出, 关于软件有用性的评论数目在“平安医生”、“春雨医生”、“医鹿”、“丁香医生”中明显多于关于其他属性的评论数目; 关于医生专业性的评论在各个在线平台的数目都比较多; 关于患者满意度的评论在“医鹿”、“好大夫在线”、“快速问医生”和“丁香医生”中的数目较多;

关于软件可靠性的评论是各个平台评论的重要组成；相比之下，各平台中关于其他属性的评论数较少。

给各组评论赋分，积极评价为 5 分，中肯评价为 3 分，消极评价为 1 分，计算各个属性的情感分值，得出结果(见表 7)：

Table 7. Emotional scores of each attribute of online medical platform

表 7. 在线医疗平台各属性情感分值

平台	专业性	医德医风	满意度	有用性	可靠性	性能	费用	客服关怀
京东健康	5	3	4.5	3.89	2.73	1.73	3	3
平安健康	4.76	4.71	4.67	3.77	1.96	1.62	1	1.33
春雨医生	4.07	4.02	4.66	4.88	3.6	2.68	2.64	3.59
医鹿	4.47	4.2	5	3.95	3.4	4.68	3	1.67
好大夫在线	3.57	3.9	4.85	4.03	2.69	1.73	1.44	1.85
快速问医生	4.74	4.76	5	3.82	4.83	3.6	3.94	4.63
丁香医生	4.82	4.82	4.12	4.7	4.12	2.44	2.18	1.54

4.8.3. 满意度的计算

依据已经得出的权值和特征词情感值，用户满意度计算公式为：

$$C = \sum W_i \sum W_j X_j \quad (9)$$

式中 C 代表用户满意度； W_i 代表要素层的权重； W_j 代表指标层的权重； X_j 代表各个特征词的综合情感值。

根据上述公式，计算得到各个指标下的满意度，最终得到各个在线医疗平台的用户综合满意度评分(见表 8)。

Table 8. Online medical platform scores and comprehensive scores of each index

表 8. 在线医疗平台各评价指标的得分及综合得分

平台	专业性	医德医风	满意度	有用性	可靠性	性能	费用	客服关怀	综合
京东健康	2.745	0.39	0.693	0.2995	0.1174	0.0121	0.078	0.042	4.37703
平安健康	2.6132	0.612	0.7192	0.2903	0.08428	0.01134	0.026	0.01862	4.37525
春雨医生	2.23443	0.5226	0.71764	0.37576	0.1548	0.01876	0.06864	0.05026	4.14289
医鹿	2.45403	0.546	0.77	0.30415	0.1462	0.03276	0.078	0.02338	4.35452
好大夫在线	1.95993	0.507	0.7469	0.31031	0.11567	0.01211	0.03744	0.0259	3.71526
快速问医生	2.60226	0.6188	0.77	0.29414	0.20769	0.0252	0.10244	0.06482	4.68535
丁香医生	2.64618	0.6266	0.63448	0.3619	0.17716	0.01708	0.05668	0.02156	4.54164

由表 8 的综合评分可看出，本文选取的在线医疗平台中，用户综合满意度评分，“快速问医生” > “丁香医生” > “京东健康” > “平安健康” > “医鹿” > “春雨医生” > “好大夫在线”，“快速问医生”的综合评分较高，主要是因为其医生专业性、软件有用性、软件可靠性的评分较高；“丁香医生”的综合评分同样较为理想，由表格看出，其医生专业性、医德医风、患者满意度、软件有用性的评分都比较

高；而对于“好大夫在线”而言，其综合评分明显低于其他在线医疗平台的主要原因是它的医生专业性评分较低，同时其他方面的评分不占优势，从而导致其综合评分不理想。由此我们可知，“快速问医生”和“丁香医生”app的综合功能更全面，提供的服务更为周到，而“好大夫在线”则仍需进一步升级平台服务，尤其需要着重提高平台中提供医疗服务的医生的专业性。

5. 结论及建议

5.1. 结论

本文以后疫情时代为背景，利用文本分析处理技术从应用商店(如华为应用市场、360手机助手等)中获取主流在线医疗服务平台的用户评论文本信息，探究影响用户满意度的主要因素，并对各个平台进行满意度评分，为用户选择合适的在线医疗平台提供帮助。

本文首先对爬取的评论文本数据进行分词、去停用词、词性标注等预处理；然后利用TF-IDF算法提取特征词，并用特征词来构建评论文本语义网络；运用聚类算法、LDA算法、K-means算法聚类对高频词进行聚类分组，将影响用户服务满意度的因素分为两大类：“医生”、“软件”；其中影响对医生的满意度的因素有：医生专业性、医德医风、患者满意度；影响对软件的满意度的因素有：有用性、可靠性、性能、费用、客户关怀。运用层次分析法搭建在线医疗平台的用户满意度模型，结合情感分析得到用户对主流在线医疗平台的满意度评分。

研究表明，在各个满意度影响因素中，医生专业性和软件有用性所占比例较高，而软件的性能和客服关怀所占比例较低；在所研究的七个主流在线医疗平台中，“快速问医生”和“丁香医生”的综合评分较高，而“春雨医生”和“好大夫在线”的综合评分较低。

5.2. 建议

从研究所得的结论可看出，用户在评价在线医疗平台时，比较看重医生的专业性(即医生能否提供有建设性的专业建议)和软件的有用性(即软件使用是否便捷或是否能在软件中获得有用的科普知识等)；相比较而言，软件的性能(即软件下载是否流畅、软件页面是否美观)和软件的客服关怀对用户评价在线医疗平台时的影响较小。因此，各在线医疗平台在进一步提升软件时，我们建议各平台可着重提升医生的专业性和软件的有用性等，如引入更有经验专业水平更高的医生为用户解答疑惑和提供就诊建议，以及在软件中提供更多便民服务。

鉴于各个医疗平台的综合得分情况，我们建议“春雨医生”和“好大夫在线”可进一步有针对性地提高平台的各项性能，为用户提供更优质的服务。同时，用户在选择医疗平台时，建议优先选择“快速问医生”和“丁香医生”。

基金项目

华北电力大学大学生创新创业训练项目(202209001)。

参考文献

- [1] 洪玲. 基于文本挖掘的“平安好医生”APP用户满意度分析[D]: [硕士学位论文]. 武汉: 中南财经政法大学, 2019.
- [2] 陈婷. 基于文本分析的移动医疗APP用户服务满意度研究[D]: [硕士学位论文]. 济南: 山东财经大学, 2019.
- [3] 丁甜甜, 钱爱兵. 在线医疗用户满意度影响因素研究[J]. 医学信息学杂志, 2021, 42(10): 22-25.
<https://doi.org/10.3969/j.issn.1673-6036.2021.10.005>
- [4] 周露莎. 在线医疗社区患者满意度挖掘及其对患者择医行为的影响研究[D]: [硕士学位论文]. 武汉: 武汉大学,

-
- 2018.
- [5] 旷开金, 廖海琳, 裴文庆. 基于文本挖掘的影评数据情感分析——以《我和我的祖国》为例[J]. 武夷学院学报, 2022, 41(5): 68-74.
- [6] 刘梦月. 基于文本挖掘的外卖类 APP 用户满意度研究[D]: [硕士学位论文]. 武汉: 中南财经政法大学, 2020.
- [7] 韩培文. 基于文本挖掘技术的顾客体验满意度研究——以京东书城为例[D]: [硕士学位论文]. 哈尔滨: 黑龙江大学, 2016.
- [8] 韩晓红, 胡彧. K-means 聚类算法的研究[J]. 太原理工大学学报, 2009, 40(3): 236-239.
- [9] 赵谦益. K-means 算法中文文献聚类的 Python 实现[J]. 软件, 2019, 40(8): 89-94.
<https://doi.org/10.3969/j.issn.1003-6970.2019.08.021>
- [10] 邢丹, 屈仁均. 基于主题模型和情感分析的进口跨境电商顾客满意度评价研究[J]. 科技和产业, 2023, 23(1): 58-65. <https://doi.org/10.3969/j.issn.1671-1807.2023.01.010>