

基于文本挖掘的生鲜电商顾客满意度研究

肖慧莲, 徐 锐

(湖北大学 商学院, 武汉 430000)

摘要:运用 Python 爬虫获取在线评论,通过词频统计和 K-means 方法得到顾客满意度评价体系的指标,通过 TF-IDF 方法计算出各指标的权重,从而得到生鲜产品的总体满意度。结果表明:消费者对生鲜产品的服务和价格的满意度相对较高,对包装的满意度相对较低;消费者对海鲜水产的满意度最高,对新鲜水果的满意度最低。最后为提高生鲜平台顾客满意度提出针对性建议。

关键词:生鲜电商;顾客满意度;在线评论;情感分析;文本挖掘

中图分类号:F274.6 **文献标志码:**A **文章编号:**1671-1807(2022)01-0288-07

网购已经成为一种重要的购物形式。不仅越来越多的消费者选择在网上购物,而且消费者网购的品类也越来越丰富。经过近几年的发展,各个电子商务平台都在如火如荼地进行中,但被誉为电商行业“最后一块蓝海”的生鲜电商还在孕育中。生鲜电商与其他电商最大的区别在于,产品更注重新鲜,不易保存。艾瑞咨询数据显示,中国生鲜电商市场发展迅速,平均每年保持 50% 以上的增长率。并且,2020 年的疫情给生鲜电商带来了极大的机遇。因此,如何把握这一机遇,需要深入分析影响生鲜电商平台上顾客满意度的因素。从这些因素入手,改善生鲜电商的产品和服务,提高顾客满意度,增强用户黏性,从而推动生鲜电商的发展。

1 文献综述

1.1 顾客满意度研究

Cardozo 在 1964 年首次提出了顾客满意度的概念^[1],然后世界各地开展了顾客满意度的研究和讨论。瑞典首先设计出瑞典顾客满意指数(SCSB)模型,模型中有预期质量、感知价值、用户满意度、用户抱怨和用户忠诚 5 个变量^[2]。随后,美国顾客满意指数(ACSI)和欧洲顾客满意指数(ECSI)模型也产生了。在上述模型的基础上,中国创建了中国顾客满意指数(CCSI)模型,如图 1 所示。

目前国内外学者通过多种方法对顾客满意度的评价开展了研究。陈范娇^[3]依据 Oliver 期望-实绩理论模型和全面顾客满意理论,建立了生鲜电商

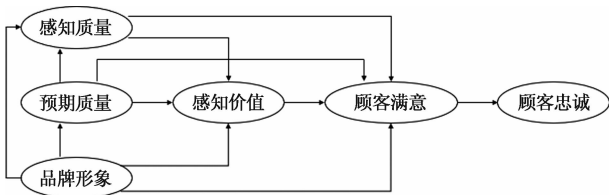


图 1 中国顾客满意指数模型

消费者满意度模型。通过问卷调查的方法获取数据,利用结构方程模型进行实证分析,结果表明,影响生鲜电商消费者满意度的主要因素为商品质量、网站功能和售后服务。程焕焕^[4]依据美国顾客满意指标(ASCI)模型,构建了消费者满意度模型,根据问卷调查数据,通过层次分析法分解出影响消费者满意度的主要因素,通过模糊综合评价法计算满意度分值,最后给出了针对天猫商城的营销建议。沈长霞等^[5]以京东运动鞋为例,运用贝叶斯网络对京东运动鞋的购买评价进行分析,得出电商平台顾客满意度的影响因素。毛志勇^[6]基于 BP 神经网络算法,建立了 B2C 电子商务企业顾客满意度评价模型,研究结果证明这个模型的误差在预定范围内,具有可行性。邵为爽等^[7]通过构建 RS-BP 神经网络模型,约简影响顾客满意度的因素,从而实施了顾客满意度评价。

1.2 在线评论情感分析研究

文本情感分析是通过分析文本,提取文本背后的态度和情绪,判断出该文本的情感倾向,一般包

收稿日期:2021-09-18

作者简介:肖慧莲(1997—),女,湖北荆门人,湖北大学商学院,硕士研究生,研究方向为电子商务与物流;徐锐(1964—),男,湖北孝感,湖北大学商学院,教授,研究方向为电子商务与物流。

括积极、中立和消极 3 种情感倾向状态。近年来,国内外学者对情感分析的研究发展迅速。常用的情感分析方法有基于情感词典的方法和基于机器学习的方法。

1.2.1 基于情感词典的情感分析

基于情感词典的情感分析方法是通过对情感词典对文本中的词语进行分析和匹配,根据一定的规则计算出情感值,从而判断出文本的情感倾向。构建情感词典的方法有很多种。其中,目前已有的知名情感词典有知网(HowNet)情感词典、台湾大学(NTSUSD)简体中文情感极性词典、清华大学李军中文褒贬义词典和大连理工大学中文情感词汇本体等。

李钝等^[8]运用语言学的知识,采用“情感倾向定义”权重优先的计算方法,从而获取词语的语义倾向程度。王振宇等^[9]将 HowNet 和 PMI 相融合计算词语情感极性。王洪伟等^[10]基于 HowNet 的语义相似度极性计算方法和副词量级划分方法,对在线评论进行了情感极性和强度分析,通过计算对快递企业服务质量进行了评价。王志涛等^[11]利用统计信息和点互信息(PMI)识别新词汇,将其添加到已有的情感词典中。根据不同的语言方式给予不同的规则,在情感计算时还考虑表情符号。张克亮等^[12]利用情感词典和概念层次网络(HNC)语境框架分两个阶段来分析文本的情感倾向性,第一个阶段是特征词判定阶段,第二个阶段是语句和句群判定阶段。

1.2.2 基于机器学习的情感分析

基于机器学习的情感分析方法是抽取评论情感特征后,用常用分类算法对文本进行情感分类,从而得到文本的情感倾向。常用分类算法有 K 最邻近算法(KNN)、支持向量机(SVM)、朴素贝叶斯(NB)、最大熵(ME)算法、决策树(DT)等。

Pang 等^[13]通过对比多种方法对电影评论进行情感分类,研究结果表明支持向量机(SVM)算法效果最好。Baltas 等^[14]分别运用朴素贝叶斯、决策树和逻辑回归 3 种方法对 Twitter 文本进行情感分析,结果显示,朴素贝叶斯的效果最好。李明等^[15]通过比较朴素贝叶斯、决策树、支持向量机(SVM)、K 最邻近算法(KNN)4 种常见算法,发现支持向量机(SVM)的召回率和精确率最高,高达到 90%以上,然后结合支持向量机(SVM)和点互信息(PMI)对商品进行细粒度情感分析。赵杨等^[16]采用卷积神经网络(CNN)和支持向量机(SVM)相结合的方法

进行情感分类,对海淘 App 各个维度的用户满意度进行了评分。鲍继彬^[17]通过比较不同方法进行分类,得出长短期记忆网络(LSTM)的准确率最高,其次是支持向量机(SVM)和 K 最邻近算法(KNN),最差的是朴素贝叶斯(NB)的分类效果。

2 文本特征分析

2.1 数据采集与预处理

京东生鲜作为综合性生鲜电商平台,生鲜品类非常全面,并且拥有大量的忠实用户。该平台商品的文本评论数量比较多,并且质量较好。在一定程度上,从京东生鲜产品评论中能够获取较为全面的评价信息。因此,选择从京东生鲜上采取数据,分别从新鲜水果、海鲜水产、精选肉类、冷饮冻食和蔬菜蛋品这 5 大商品分类中各选取一个产品,具体爬取产品及评论来源见表 1。

表 1 采集的数据来源

商品分类	商品名称	商品网址
新鲜水果	甄选应季鲜橙	https://item.jd.com/4176410.html
海鲜水产	国联翡翠生虾仁	https://item.jd.com/100001243976.html
精选肉类	双汇猪五花肉块	https://item.jd.com/3411176.html
冷饮冻食	伊利大白瓶新鲜纯牛奶	https://item.jd.com/100015295314.html
蔬菜蛋品	家美舒达滕州土豆	https://item.jd.com/4019068.html

编写 Python 程序进行爬虫,爬取各个商品从 2020 年 3 月 1 日到 2021 年 6 月 30 日的评论,总共获得 8 580 条评论数据,其中新鲜水果 2 990 条、海鲜水产 1 400 条、精选肉类 1 400 条、冷饮冻食 1 390 条、蔬菜蛋品 1 400 条。包括 4 990 条好评,1 790 条中评,1 800 条差评。

一部分用户在发表评论时,遣词造句比较随意。一部分用户虽然进行了评论,但其态度极其不认真,敷衍了事,甚至复制粘贴他人评论,因此爬取的数据中有很多无意义的评论。为了防止这类无意义数据对研究结果产生影响,必须首先对获得的评论数据进行数据清洗。主要包括以下 6 类评论数据:①文本中含有特殊字符和无意义的英文字符的评论,清除掉无意义字符;②由于用户未发表评论系统自动默认的“此用户未填写评价内容”此类评论,直接进行删除;③用户直接复制粘贴的重复评论,保留第一条评论;④文本中存在大量重复用词,用户存在凑字数的嫌疑,如“我认为这个橘子非常

新鲜非常新鲜非常新鲜非常新鲜非常新鲜”，通过机械压缩为“我认为这个橘子非常新鲜”；⑤评论字数非常少的评论，如“很好”“还行”“一般”，这类评论表述模糊不清，不能确定是对商品某方面的评价，用于产品分析时没有什么意义，在进行短句删除处理时，删除掉原评论字符数或者机械压缩后的评论字符数小于等于 3 个字符的评论；⑥与商品评价完全无关的评论，这类评论需要手动剔除。

通过数据清洗后爬取的评论数见表 2。

表 2 处理后的评论数

商品分类	评论数
新鲜水果	2 932
海鲜水产	1 323
精选肉类	1 350
冷饮冻食	1 295
蔬菜蛋品	1 319
总计	8 219

2.2 文本特征分析

2.2.1 基于词云图的特征分析

词云图将文本数据中的高频词汇清晰明了地展示出来，直观表达文本数据的意思，对大量文本数据进行可视化分析。通过对词云图进行分析，可以清楚得到在线评论的主题，可以了解消费者比较关注的因素。通过利用 Python 分别对新鲜水果、海鲜水产、精选肉类、冷饮冻食和蔬菜蛋品这 5 类商品制作词云图，如图 2~图 6 所示，同时综合所有商品对京东生鲜制作词云图，如图 7 所示。



图 2 新鲜水果的词云图

从图 7 可知，京东生鲜的消费者对商品的关注点的关键词为新鲜、好吃、不错、满意、味道、质量、包装、口感、价格、物流、客服、服务、性格比等，说明消费者对产品全方面关注。



图 3 海鲜水产的词云图



图 4 精选肉类的词云图

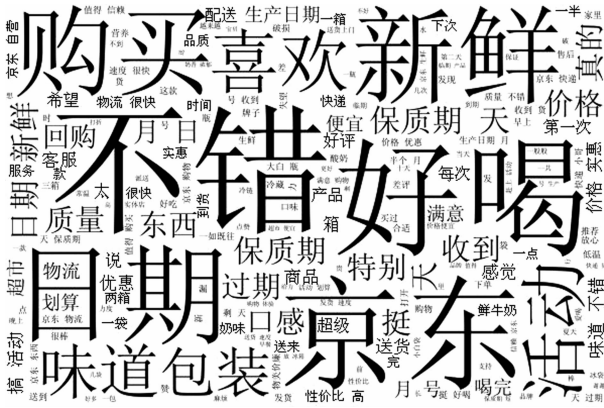


图 5 冷饮冻食的词云图

2.2.2 基于网络语义图的特征分析

语义网络由节点和有向线段组成，节点代表常识概念，有向线段代表这些概念之间的关系。通过对网络语义图的分析，可以突出中心词，以及展示与中心词紧密联系的词语。采用 ROST CM 6 文本挖掘软件对在线评论文本进行社会网络和语义网络分析。

从图 8 可以看出，整体上京东生鲜的语义网络

联系紧密,没有孤立点,而且节点相对突出。整个图都围绕“京东”这个节点。与“京东”紧密联系的关键词是“信赖”“方便”“放心”“值得”,说明总体上消费者在京东上购买生鲜产品是满意的,认为在京东生鲜上购物值得信赖,认为在京东生鲜上购物是美好的。

此外,“质量”“物流”“包装”“服务”等是网络语



图 6 蔬菜蛋品的词云图

义图的核心节点,说明在线评论文本大量提及质量、物流、包装和服务。与“质量”一词紧密联系的情感词是“满意”和“新鲜”,与“物流”一词紧密联系的情感词是“很快”和“满意”,与“包装”一词紧密联系的情感词是“满意”,与“服务”一词紧密联系的词是“态度”。总的来说,消费者对京东生鲜各方面都较为满意。持正面评价。

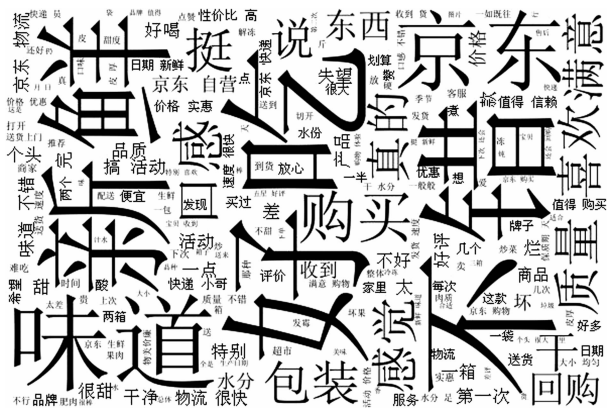


图 7 京东生鲜产品的词云图

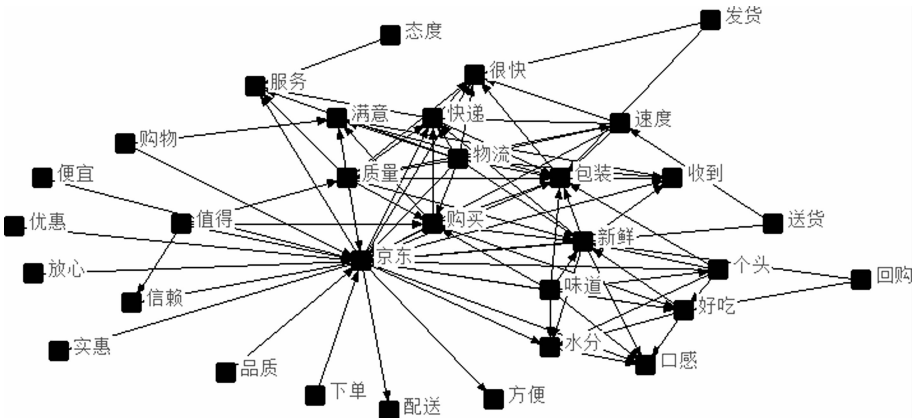


图 8 京东生鲜评论语义关系网络

3 顾客满意度测算

3.1 特征词选取

通过结巴分词对在线评论文本进行分词,然后进行词频统计。将高词频的词汇中明显与主题不相关的词语进行筛选,提取出在线评论文本中的高频特征词。通过 K -means 聚类算法对在线评论进行聚类,最终确定 $k=5$,则将从物流、包装、服务、质量和价格这 5 个维度来衡量生鲜电商的顾客满意度。

由于中文博大精深,一个词语拥有众多表达方式,拥有大量近义词。如果对此类近义词忽略不计的话,将对研究结果产生一定的影响,因此,运用词向量法(Word2vec)对特征属性进行相似词语的统

计,见表 3。

表 3 特征词词典

特征词	相似词
物流	快递、送货、冷链、发货、速度
包装	破损、严实、完好、精美、箱内
服务	态度、热情、一如既往、客服、周到
质量	品质、品牌、不错、值得、信赖
价格	实惠、活动、经济、划算、便宜

3.2 情感词典构建

情感词典主要包括情感词词表、副词词表和否定词词表。本文的词库的主要来源是知网中文词库(HowNet)。

1)情感词词表。将知网中文词库(Hownet)中的正面情感词语和正面评价词语合并构成正面情感词语,将知网中文词库(Hownet)中的负面情感词语和负面评价词语合并构成负面情感词语。

2)副词词表。在文本中,程度副词表达了情感程度,比如“速度慢”“速度好慢”和“速度超级慢”表达了3种不同的情感,对速度的不满程度逐渐加深。为了精准识别文本的情感程度,需要对程度副词赋予相应的权重。依照知网中文词库(Hownet)将情感副词分为5个等级,并且赋予5个不同的权重,见表4。

表 4 情感副词词典

程度级别	举例	权值大小
超	百分之百、绝对、完全、万分、至极	2
非常	格外、颇为、实在、特别、尤其	1.5
较	更加、还要、较为、愈加、越发	1.25
稍	略、略微、稍微、相当、有点	0.5
不足	不怎么、轻度、丝毫、微、相对	0.25

3)否定词表。否定词将句子的情感倾向完全改变,比如“牛奶不新鲜”是“牛奶新鲜”完全相反的意思,将对牛奶的正面评价变成了负面评价。本文根据之前分词的结果,选出一些具有否定含义的词,构建了否定词表,见表5。

表 5 部分否定词

否定词	权值大小
不、勿、否、无、非、没、不能、不是、从不、不要、禁止、没有、禁止、绝不、从未、几乎不	-1

3.3 产品特征情感分析

依据提取出来的5类特征词,对在线评论文本进行筛选,将各类产品中含有该特征词的在线评论文本进行分类,从而计算各类产品各特征属性的顾客满意度。情感值的具体计算规则如下:假设存在特征词 W_i ,对某一含有 W_i 特征词的在线评论文本中,如果修饰该特征词的情感词是正面情感词语,则分值为2;如果修饰该特征词的情感词是负面情感词语,则分值为-2。如果该文本中存在程度副词,则按照副词词表进行相应赋值。如果该文本中存在否定词,统计否定词的个数,分值为 $(-1)^n$ 。因此,含有特征词 W_i 的该文本情感值的计算公式为

$$S(i)=E(i)D(i)I(i) \tag{1}$$

式中: $E(i)$ 为该评论中情感词的分值(2分或者-2分); $D(i)$ 为该评论中程度副词的分值; $I(i)$ 为该评论中否定词的分值。

依次计算出各类生鲜产品的每条评论的情感值后,取含有 W_i 特征词的该类产品满意度的平均值即为该产品在 W_i 特征上的满意度。具体的计算结果见表6。

表 6 各类产品各特征的满意度

满意度	物流	包装	服务	质量	价格
新鲜水果	2.56	2.12	2.98	2.34	2.96
海鲜水产	2.67	2.34	2.89	2.88	2.89
精选肉类	2.56	2.23	2.67	2.87	2.87
冷饮冻品	2.67	2.34	2.89	2.88	2.89
蔬菜蛋品	2.65	1.98	3.23	2.51	2.81

表6显示,消费者对新鲜水果的服务和价格相对比较满意,对海鲜水产和冷冻饮品的服务、价格和质量相对比较满意,对精选肉类的质量和价格相对比较满意,对蔬菜蛋品的包装和价格相对比较满意。整体上,消费者对各类产品的价格和服务的满意度比较高,对包装满意度最低。

3.4 产品总体满意度计算

3.4.1 指标权重的确定

权重反映了消费者在生鲜电商平台上对各项指标的重要程度。常用的权重计算方法有TF-IDF权重、布尔权重和频度权重。

TF-IDF是由TF和IDF两部分构成的,在文本挖掘领域被广泛使用。TF即词频,表示关键词在文本中出现的频率。IDF即逆向文件频率,度量一个词的普遍重要性。IDF越大,说明关键词的区分度越高。

TF-IDF的计算公式为

$$TF-IDF=TF\times IDF \tag{2}$$

TF-IDF的值越大,则关键词的权重越高。

因此,运用TF-IDF算法计算物流、包装、服务、质量、价格这5个影响因素在新鲜水果、海鲜水产、精选肉类、冷饮冻食和蔬菜蛋品这5类产品中所占的权重,结果见表7。根据表7绘制雷达图,如图9所示。

表 7 各类产品各特征的权重

产品类别	物流	包装	服务	质量	价格
新鲜水果	0.29	0.17	0.12	0.37	0.25
海鲜水产	0.37	0.13	0.13	0.56	0.41
精选肉类	0.30	0.10	0.07	0.64	0.44
冷饮冻品	0.32	0.14	0.11	0.43	0.51
蔬菜蛋品	0.34	0.11	0.11	0.60	0.22

通过图9可以看出,京东生鲜平台上,在质量、价格和物流这3个维度的权重相对更多。

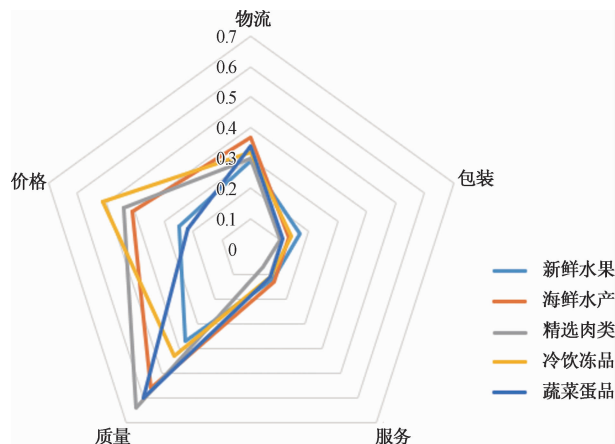


图 9 各类产品各特征的雷达图

3.4.2 产品满意度的计算

为了比较消费者对京东生鲜上各类产品的满意程度,分别计算出新鲜水果、海鲜水产、精选肉类、冷饮冻食和蔬菜蛋品这 5 类产品的顾客满意度。总体满意度的计算公式为

$$S_i = \sum W_{i,j} F_{i,j} \tag{3}$$

式中: S_i 代表第 i 类产品消费者的满意度; $W_{i,j}$ 表示第 i 类产品第 j 个特征的权重; $F_{i,j}$ 表示第 i 类产品第 j 个特征的顾客满意度。消费者对各产品满意度情况计算结果见表 8。

表 8 各产品的满意度

产品类别	新鲜水果	海鲜水产	精选肉类	冷饮冻食	蔬菜蛋品
满意度	3.066 2	4.465 5	4.277 5	4.212 2	3.598 3

通过表 8 可以看出,消费者对京东生鲜上这 5 类产品的总体满意度排名最高为海鲜水产,排名最低为新鲜水果。

4 结论与建议

4.1 结论

通过 Python 爬取京东生鲜产品的在线评论,通过对在线评论进行情感分析,研究得到影响消费者对生鲜电商平台上产品的满意度因素依次为物流、包装、服务、质量和价格。研究结果表明,消费者在电商平台购买生鲜产品时最满意的是海鲜水产,其次为精选肉类,最不满足的是新鲜水果。整体上,消费者对各类产品的价格和服务的满意度比较高,对包装满意度最低。

消费者对京东生鲜上不同种类的产品看重的特征不同。比如,对冷冻产品,消费者最看重价格,对其他种类产品,消费者最看重质量。整体上,消费者对京东生鲜的产品更看重质量、价格和物流。

4.2 建议

根据研究结果,针对京东生鲜平台,提出以下几点提高顾客满意度的建议。

4.2.1 进一步扩大优势

消费者对京东生鲜平台的质量、价格和物流相对满意,可以进一步扩大这几个方面的优势。

1)质量方面。京东生鲜拥有稳定的品牌供应商,产品来源可控。京东生鲜减少了从产地到顾客手中过程中的损耗,使得农产品比较新鲜。因此在质量方面有一定的优势。

生鲜电商平台还可以通过区块链、物联网等技术,实现质量溯源。消费者可以随时观察产品的生产、加工、包装等全过程。信息可视化后,消费者对产品的质量更加信任,平台也会更注重生产过程中的质量问题。

2)价格方面。京东生鲜的有些产品原产地直采,有些产品建立了自己的生鲜农产品基地,因此能提供相对优惠的价格。

生鲜电商平台还可以在选择供应商时,进行多方对比,在保证质量的前提下,和供应商进行长期合作,同时优化物流、仓储等方案,减少成本,从而给顾客最优惠的价格。

3)物流方面。京东生鲜自建物流设施,打造冷链体系,采取社会化冷链协同网络、标准化产品矩阵和智能化供应链管理来解决生鲜产品需要保鲜的问题。目前有“标准达”和“京准达”两种配送方式。因此,消费者对京东生鲜平台的物流满意度相对较高。

4.2.2 加强商品包装

京东生鲜平台应当重视生鲜产品的包装。从功能和环境保护角度出发,针对不同种类的生鲜产品,结合产品的特点给予合理的包装。另外,企业还可以加大研发投入,研制新的包装产品,不仅使得包装结实完好,而且保证产品的新鲜。

4.2.3 提供人性化服务

1)提高客服服务水平。平台可以对客服工作人员进行统一专业化的培训,提高工作人员的专业知识。目标是当每位顾客在任意时间段需要客服时,客服都能及时有效地提供咨询和解决疑问,同时争取做到从咨询到售后由同一个客服团队负责同一位顾客。针对消费者可能出现的退换货问题,平台应做好预见性的处理方案,及时解决,避免消费者产生不好的购买体验。同时,平台应该进行跟踪和分析,避免类似问题再次发生。

2)可以实施线上和线下相结合的全渠道零售模式。生鲜电商平台建立优质的线下实体店,提供优质的商品和服务,提高消费者的满意度。通过采取线上线下同价等措施,打通线上和线下,对线上和线下资源进行充分整合。

参考文献

[1] CARDOZO R. Consumer satisfaction; Laboratory studyand marketing action[J]. Journal of Marketing Research,1964(2):244-249.

[2] 裴飞,汤万金,咸奎桐. 顾客满意度研究与应用综述[J]. 世界标准化与质量管理,2006(10):4-7.

[3] 陈范娇. 生鲜电商消费者满意度影响因素实证研究[J]. 商业经济研究,2016(8):49-51.

[4] 程焕焕. B2C 模式下网购顾客满意度测评研究[D]. 济南: 山东财经大学,2014.

[5] 沈长霞,车万留,桂海霞. 基于贝叶斯网络的电商客户满意度影响因素分析;以京东运动鞋为例[J]. 数学的实践与认识,2020,50(23):285-294.

[6] 毛志勇. 一种基于 BP 神经网络的 B2C 电子商务顾客满意度评价模型[J]. 科技和产业,2008(5):49-52.

[7] 邵为爽,李晓红. RS-BP 神经网络在 C2C 电子商务顾客满意度评价中的应用 [J]. 科技通报,2013,29 (5): 72-75,172.

[8] 李钝,曹付元,曹元大,等. 基于短语模式的文本情感分类研究[J]. 计算机科学,2008(4):132-134.

[9] 王振宇,吴泽衡,胡方涛. 基于 HowNet 和 PMI 的词语情感极性计算[J]. 计算机工程,2012,38(15):187-189,193.

[10] 王洪伟,宋媛,杜战其,等. 基于在线评论情感分析的快递服务质量评价[J]. 北京工业大学学报,2017,43(3):402-412.

[11] 王志涛,於志文,郭斌,等. 基于词典和规则集的中文微博情感分析[J]. 计算机工程与应用,2015,51(8):218-225.

[12] 张克亮,黄金柱,曹蓉,等. 基于 HNC 语境框架和情感词典的文本情感倾向分析[J]. 山东大学学报(理学版),2016,51(7):51-58,73.

[13] PANG B,LEE L,VAITHVANATHAN S. Thumbs Up? Sentiment classification using machine learning techniques[C]//Proceedings of the 2002 ACL Conference on Empirical Methods in Natural Language Processing,2002:79-86.

[14] BALTAS A,KANAVOS A,TSAKALIDIS A K. An Apache Spark implementation for sentiment analysis on twitter data [M]. Springer International Publishing,2017:15-25.

[15] 李明,胡吉霞,侯琳娜,等. 商品评论情感倾向性分析[J]. 计算机应用,2019,39(S2):15-19.

[16] 赵杨,李齐齐,陈雨涵,等. 基于在线评论情感分析的海淘 APP 用户满意度研究[J]. 数据分析与知识发现,2018,2(11):19-27.

[17] 鲍继彬. 基于机器学习的网购平台产品评论情感分析 [D]. 哈尔滨:哈尔滨工业大学,2019.

Research on Customer Satisfaction of Fresh Food E-commerce Based on Text Mining

XIAO Huilian, XU Rui

(School of Business,Hubei University,Wuhan 430000,China)

Abstract: Python crawler was used to obtain online reviews, the index of customer satisfaction evaluation system was obtained through word frequency statistics and K-means method, calculate the weight of each index through TF-IDF method, and the overall satisfaction status of fresh products was get. The results show that consumers have relatively high satisfaction with the service and price of fresh products, and relatively low satisfaction with packaging. Consumers have the highest satisfaction with seafood and aquatic products, and the lowest satisfaction with fresh fruits. Finally, targeted suggestions were put forward for improving customer satisfaction of fresh food e-commerce.

Keywords: fresh food e-commerce;customer satisfaction;online reviews;sentiment analysis;text mining