

基于情绪分析模型的市场数据跟踪测量

和 ARIMA 模型

总结

亚马逊为客户提供评分和评分的机会。其他客户可以利用这些评论来帮助他们做出购买决定。基于客户评论等信息，公司可以做出增加产品吸引力的决策。如果公司掌握了产品未来的发展趋势，就可以为之提供依据

业务决策。

至于任务一，为了准确捕捉数据集中的有效信息，对原始数据集中的缺失值、异常值和重复观测进行了一系列的数据处理方法处理，并采用定性和定量相结合的方法筛选不相关的变量。此外，为了对产品市场满意度做出初步的了解，采用了描述性统计。然后采用基于词典的方法对文本数据进行分析，建立情感评分评价体系，能够从每条评论中快速捕捉顾客满意度。

任务 II 根据问题的相关性分为三个小版块。在第 1 小节中，建立了一个星级评级的三级评估体系。在使用相关性检验判断评论和评分对明星的影响后，通过主成分分析建立了包含投票信息、评论内容和客户身份的综合评价体系。在此基础上，借鉴亚马逊平台原油之星算法，引入时间跨度指标，对原有三级体系进行更新，使体系更加丰富

全面。

对于第 2 小节，引入时间序列模型，研究时间对评论倾向性的影响。首先，我们将时间序列可视化，得到产品的情感趋势图和其他变量的时间趋势图。通过图像分析，掌握不同商品的情感得分趋势，得到三种商品的星级评分、累计评论数、累计点赞数的时间趋势。接下来，在时间序列模型的基础上，我们引入了差分自相关移动平均模型(difference autocorrelation moving average model, ARIMA)，该模型可以准确预测 9 月和 10 月的星级评分情况和评论情感趋势

2015.

最后讨论了极端评论对其他变量的影响以及极端评论与极端明星的互动关系。在研究后续评论对极端明星数量和情感方向的影响时，由于数据不符合正态分布的特征，非参数检验成为本文的主要方法。从两个方向分析评论中极端情绪词汇对后续明星的影响：

正面和负面。

综上所述，本文所开发的三个产品的实证分析和模型，对于其他电子商务平台也是一个有力的工具，显示了一个

优秀的性能。

关键词:基于词典的方法情感评分评估系统

主要

成分分析

时间序列模型

华宇电
脑

非参数检验



关注数学模型
获取更多资讯

有道文档翻译
pdf.youdao.com

致阳光公司营销总监的信亲爱的先生或女士:

根据您的网上销售需求, 我们团队为您提供一些解决相关问题的建议。我们提出了数学模型来分析情感极性, 并找出三种客户反馈之间的关系。我们还考虑了时间因素如何影响评分变化, 并提供了一个简单的模型, 可以用来预测短期内的评分和情感趋势。最后我们形成一个 COM_S 参数来定义推荐的产品。

具体来说, 基于词典的方法是在情感词典的基础上, 对客户提供的评论文本进行文本情感分析。并通过该方法得到一个重要的情感得分, 以发现评论和评分变量之间的内在联系。然后, 我们建立了两个参数客户评论(CR)和客户满意度(CS), 以帮助呈现三个变量之间的具体关系。这两个参数之一, 客户评论, 通过主成分分析方法进行分析。三个主成分分别是投票信息、评论内容和客户。另外两个参数中的一个, 客户满意度(CS)也可以通过详细的计算公式来确定。最后, 综合满意度(COM_S)可以从同一产品的 CS 汇总中推导出来。

而当我们考虑时间因素时, 可以利用时间序列模型来刻画不同变量的趋势。此外, 我们还关注了表现为一星或五星的极端情绪如何改变自身或影响其他评分变量。结果显示, 极端的态度会影响其他客户, 尤其是在很长一段时间内。

然后, 我们对产品目前存在的问题提出了一些建议, 并确定为今后需要修改的重点。在对每个产品的评论进行分析的过程中, 一些常见的产品问题需要多加注意。对于微波炉, 要注意产品本身的性能和价格, 并及时对产品进行改进。对于电吹风来说, 更要关注产品的性能和外观。对于安抚奶嘴, 评价普遍较高, 公司要注意产品使用寿命的问题。我们相信我们的方法可以帮助贵公司选择最有可能成功的产品, 避免风险的产品。以下是推荐列表:

| | 成功的 | 不成功的 |
|-------|--|---|
| 奶嘴 | cosco alpha、Fisher、ameda 和 Sealy | Chicco keyfit、宜家 和蝙蝠侠 |
| 微波吹风机 | sharp、lg、magic chef、 whirlpool、profile、Panda Small 沙龙、Waterpik、微镊子、 DuWop、夏威夷岛、Oral-B | 三星、丹比、 jx7227sfss Andis, conair, xtava , 磁流体动力,雷 |

你真诚的,
#2009116 团队



关注数学模型
获取更多资讯

内 容

1 介绍5

1.1 问题的声明5

1.2 文献综述5

1.3 概述我们的工作5

1.4 符号6

1.5 假设6

2.任务 16

2.1 问题分析6

2.2 方法7

2.3 解决问题7

2.3.1 细节的数据预处理8

2.3.2 描述性统计的结果9

2.3.3 评论的情感分析11

3.任务 211

3.1 问题分析11

3.2 方法11

3.3 解决问题12

3.3.1 确定星级指数 CR12

3.3.2 综合满意度的分析14

4.任务 316

4.1 问题分析16

4.2 解决问题16

4.2.1 时间序列可视化和准备三个变量的预测16

4.2.2 预测星评级18

4.2.3 预测评论情绪18

4.2.4 时间序列分析18

4.2.5 star_rating 和 审查之间的相互作用19

5.模型评价22

5.1 模型的局限性22

时间序列分析:22



关注数学模型
获取更多资讯

Lexicon-based 方法 : 22

主成分分析的缺点 : 22

5.2 模型的优点 22

时间序列分析: 22

Lexicon-based 方法 : 22

主成分分析的优点 : 23

6.结论 23

7.参考文献 24



关注数学模型
获取更多资讯

1 介绍

1.1 问题陈述

以亚马逊为首的在线市场已经成为大多数公司最重要的交易平台。作为消费者相互交流购物体验的机制，各种评论和评分为潜在消费者提供了主观证据。它还可以为新进入者提供选择目标市场和目标产品的指导。此外，在社交媒体时代，消费者的声音比以往任何时候都更大、更强，从数据可以看出，超过 50% 的消费者在购买产品前经常会考虑在线评论。[1]因为这个过程可以看作是与前客户的直接链接，获取产品的第一手信息，以避免卖家的夸大或炒作。

但是，在这个阶段，还不清楚他们的互动，以及他们如何影响企业战略。因此，我们的目标是探索三个变量的内在关系，并设计一个复合的数据度量，以协助企业选择潜在的成功产品。同时，讨论了客户群体评论对个体评论态度取向的影响。最后，根据我们的方法对备选产品进行排序，并向阳光公司提供一个推荐列表。

1.2 文献综述

目前，情感分析主要有两种方法，一种是基于机器学习的方法，另一种是基于词典的方法。近年来，基于机器学习的几种方法逐渐成为人们关注的焦点，其中大部分都是基于社交媒体上大量的个性化评论。由于社交媒体上的评论不仅是英语，还有阿拉伯语等其他语言。[2]很多来自非英语地区的研究者也运用这种方法来解决问题，可见这种方法是相当成熟的。

另一种方法是基于词典的方法。这种方法也利用了人们在社交媒体上表达情感或情感的文本，如 Twitter 帖子和微博文本。[3,4]对情感极性词典的最新研究也表明，该词典的资源被扩展到特定领域，即股票市场，无需人工干预，并解决了缩放和阈值化问题。[5]因此，我们认为这种方法具有广泛应用的潜力。

在我们比较了这两种方法之后，我们决定选择基于词典的方法。我们认为，这种方法在时间和资源有限的情况下最适用，即与乌尔都语情感分析研究中得到的结论相似。[6]

在引入时间因素之后，我们提出了一个时间序列模型来讨论三个变量在时间维度上的变化。在比较了几种适用的模型后，我们选择使用传统的时间序列模型和自回归综合移动平均 (ARIMA) 模型来观察时间因素对不同评论的影响，并预测未来的指标趋势。我们认为在预测阶段，ARIMA 模型可以像 Matyjaszek, M. et al.[7]的理论一样实现相当准确的预测同时，我们还对不同产品进行了逐产品的产品预测。这与 Nguyen, H.等人的观点一致。该模型对不同产品价格的预测与[8]相似。

1.3 我们的工作概述

我们的主要目标是解决三个需要探索的问题:(1)文本评论的情感分析;(2)评论和评分指标之间的关系以及(3)时间的影响。然后提出了这三个内容相结合的产品排序方法。

为了解决这些问题，我们首先分析了每个数字参数的数据特征，然后应用基于情感词典的方法来分析文本评论的文本情感态度，提取其中的情感词



关注数学模型
获取更多资讯

句子通过内连接，再计算句子情感得分。这个情感评分的过程，就是用来区分句子的积极态度和消极态度的。然后通过比较文本评论和两个评分变量，得到客户评论(Customer Review, CR)来讨论星级评分如何受到评论因素的影响，然后讨论客户满意度(Customer Satisfaction, CS)的影响因素，并得出定量关系。之后，我们根据每个客户的满意度，根据产品满意度形成综合满意度(COM_S)。并且我们用这个数据衡量来筛选我们的候选产品，并为阳光公司提供一个适合在线销售的替代产品列表。

1.4 符号

| 象征 | 定义 |
|-------|-------------------|
| P | 安抚奶嘴电吹风微波帮助投票 |
| H | |
| 米 | total_votes |
| 高压 | SentimentQDAP |
| 电视 | star_rating |
| 平方 | number |
| 老 | Vine |
| n | verified_purchase |
| v | customer_id |
| 副总裁 | product_parent |
| ci | 日第一主成分第 |
| 页 | 二主成分第三主成分 |
| D | |
| 1 | 客户审查 |
| 2 | 客户满意度综合满意度 |
| 3 | |
| CR | |
| CS | |
| COM_S | |

1.5 假设

- 我们假设事物过去的趋势将延续到未来。事物的现实是历史发展的结果，事物的未来是现实的延伸。事物的过去和未来是相通的。
- 我们假设所预测的数据是不规则的。
- 我们假设市场发展之间没有因果关系。
- 我们假设不存在其他影响时间序列分析的相关随机变量变化。
- 我们假设给定数据集中所包含的变量能够充分反映产品的综合得分。
- 我们假设基于词典的情感分析方法选择的字典包含所有所需的单词。
- 我们假设当客户评论中有多个句子时，所有句子的情感得分的平均值可以代表客户的情感得分。● 我们假设客户评论是简单的逻辑关系，没有使用贬义词来表示积极的含义，也没有否定词、程度副词和情感词的复杂组合。
- 我们假设情感词典中情感词的定义是没有争议的。

2.任务 1

2.1 问题分析

由于本课题提供的数据量大，且多以文本的形式呈现，因此本课题的重点是对原始数据的预处理和提取



关注数学模型
获取更多资讯

有道文档翻译
pdf.youdao.com

文本信息。本文在解决任务 1 的处理中，首先使用一系列数据处理方法对原始数据集中的缺失值、异常值和重复观测进行处理，并采用定性和定量相结合的方法对不相关变量进行筛选。这一步骤既可以去除杂项信息的干扰，又可以准确捕捉数据集中的有效信息。接下来，本文对三种商品的不同变量进行描述性统计，对产品市场满意度做出初步了解。最后，本文采用基于词典的方法对文本数据进行分析，并建立情感评分评价体系，能够从每条评论中快速捕捉顾客满意度。

2.2 方法

本节分为三个部分，数据预处理，描述性统计和基于词典的方法。在这些小节中，介绍了任务 1 中建模的思想。

2.2.1 数据预处理

由于在设计这一问题时数据量巨大，为了去除干扰信息的影响，提取数据中的有效信息，对原始数据的数据预处理是解决问题的关键一步。本文的数据预处理可以分为以下几个步骤：

步骤 1. 检查原始数据集中的重复观测值

步骤 2. 缺失值和异常值的处理

步骤 3. 观测数据的标准化

步骤 4. 删除无用的变量

2.2.2 基于词典的方法

目前解决情绪分析方法有两种：一种是基于机器学习的方法，一种是基于情绪词典的方法。

基于机器学习的方法是输入大量人工标注的语料库作为训练集，通过提取文本特征和构造分类器来实现情感分类。由于个人电脑的限制和比赛的长度，我们很难做到这一点。同时，由于我们不清楚已知分类和评论文本之间的内在联系，这种方法并不适合解决这个问题。

因此，我们选择了基于情感词典的方法，先只使用已知文本和情感词典，不涉及其他评分指标。本文引用的三种词典分别为：

(1)通用问询软件中使用的来自 Harvard-IV 词典的观点词典

根据通用问询软件中使用的心理学哈佛-iv 字典，带有积极和消极词汇列表的字典。这是哈佛大学开发的通用词典。

(2)来自《亨利金融词典》的观点词典

根据《亨利金融词典》，词典中包含了褒贬义词。这个词典最初是由金融学科中最早采用文本分析的人之一在《商业交流杂志》上提出的。

(3)来自 Loughran-McDonald 金融词典的观点词典

根据 Loughran-McDonald 金融特定词典，字典包含积极、消极和不确定性词汇列表。该词典首次发表在《金融杂志》上，此后在金融领域被广泛使用。

一般步骤如下。首先，通过收集相关的文本信息，他们创建了包含文本、用户、情感、情感信息的特定数据集。然后，通过提取和构建一系列相关词典，可以对情感词典进行扩展。接下来，通过权重的计算就可以得到一篇文本的情感值。最终，研究人员得到了影响力分数，从而实现了他们的最终目标。

2.3 解决问题



关注数学模型
获取更多资讯

为方便后续工作，复制原电吹风内容。tsv,微波炉。TSV 和安抚奶嘴。tsv 文件到 excel，保存为 CSV 文件。数据集 pacifier.csv 包含 18939 个观测值和 16 个变量;Microwave.csv 包含 1615 个观测值和 16 个变量;Hair_dryer.csv 包含 11470 个观测值和 16 个变量。三个文件中的 16 个变量相同，变量分为字符型变量和数值型变量。在变量中，星评、助人为乐的声音和评论主体是最重要的研究对象;评论 ID 是评论号，是识别变量，唯一。变量还包括产品信息、市场信息、客户信息和评论信息。

2.3.1 数据预处理细节

步骤 1. 检查原始数据集中的重复观测值

将文件导入 SAS 系统，使用 proc 排序过程中的 nodupkey 选项检查数据集中是否存在重复观测，并使用 review [ID]作为识别变量。如果有重复观测，则将其存储在工作中。Dup 数据集。三个文件的检查结果显示，在数据集工作中没有观察到。Dup，所以说明在原始数据集中没有重复观察。

步骤 2. 缺失值和异常值的处理

(1)缺失值

在 Excel 中，你可以简单地检查缺失值并填写缺失值。pacifier.csv 中的产品“category”是缺失值。Pacifier 是婴儿产品，可以直接填充。此外，还发现了 8 个遗漏重要信息的观察结果。这 8 个观测被删除，剩下 18931 个观测。电吹风的评审日期变量中有 7 个缺失值。Csv，可以手动补充。微波.csv 没有缺失值。

(2)异常值

首先，我们使用 proc 内容过程来生成变量描述表工作。var，然后我们使用 proc SQL 过程将字符类型变量定义为宏变量 var_char，将数值类型变量定义为宏变量 var_num，以便在后期过程中更高效地调用字符类型变量和数值类型变量;然后使用 proc means 进程的 nmiss 选项来检查数值类型变量中的缺失值，使用 proc freq 进程中的 tables 语句的缺失选项来检查字符类型变量中的缺失值，并检查字符类型变量的频率分布。结果显示，字符类型变量中没有缺失值，packer.csv 文件中的 review [data]中有 14 个缺失值，其他变量中也没有缺失值。没有明显的错误值。在 pacifier 文件中删除 14 个有缺失值的观测值，在微波文件中删除 18917 个观测值和 1615 个观测值，其中 11470 个观测值。

步骤 3. 观测数据的标准化

亚马逊是一个评论自由度很强的平台，所以有很多垃圾评论需要删除。根据观察，评论中有很多垃圾评论或者无用的评论，我们需要过滤掉。另外，我们需要通过评论内容来分析情感和态度，所以长度过短的无用评论也需要被删除。为了方便后续工作，有必要对评论文本进行规范。

(1)删除过短的评论

使用 length 语句检查三个文件中观察到的审查主体的字符串长度。如果字符串长度小于 20，则删除注释。奶嘴文件中串长小于 20 的观察评论正文有 1674 篇，电吹风文件中串长小于 20 的观察评论正文有 738 篇，微波文件中串长小于 20 的观察评论正文有 79 篇。

(2)数据标准化。

为了便于后续提取评论体中的重要信息，需要删除评论中的标点符号，将全部大写改为小写，使用 tranwrd 函数删除标点符号，并使用 translate 函数



关注数学模型
获取更多资讯

将评论中的大写改为小写。此外，使用 `translate` 函数修改变量 `vine` 和变量 `veried_` 的大写观察值。**步骤 4。** 删除无用的变量

文件中有一些无用的变量可以删除。三个文件市场相同，可以删除。产品 ID、产品标题、产品父级可以唯一标识产品。只保留三者中的一个，并保留 `product parent`。产品类别和评论标题都是无用的变量，可以删除。使用 `drop` 语句删除无用变量。

最后，从 `pacifier` 文件中删除 1674 个观测值，保留 17243 个观测值和 10 个变量。毛发家伙文件中删除了 738 个观测，保留了 10732 个观测和 10 个变量。微波文件中删除了 79 个观测，保留了 1536 个观测和 10 个变量。**2.3.2 描述性统计结果**

(1) Star_ratings

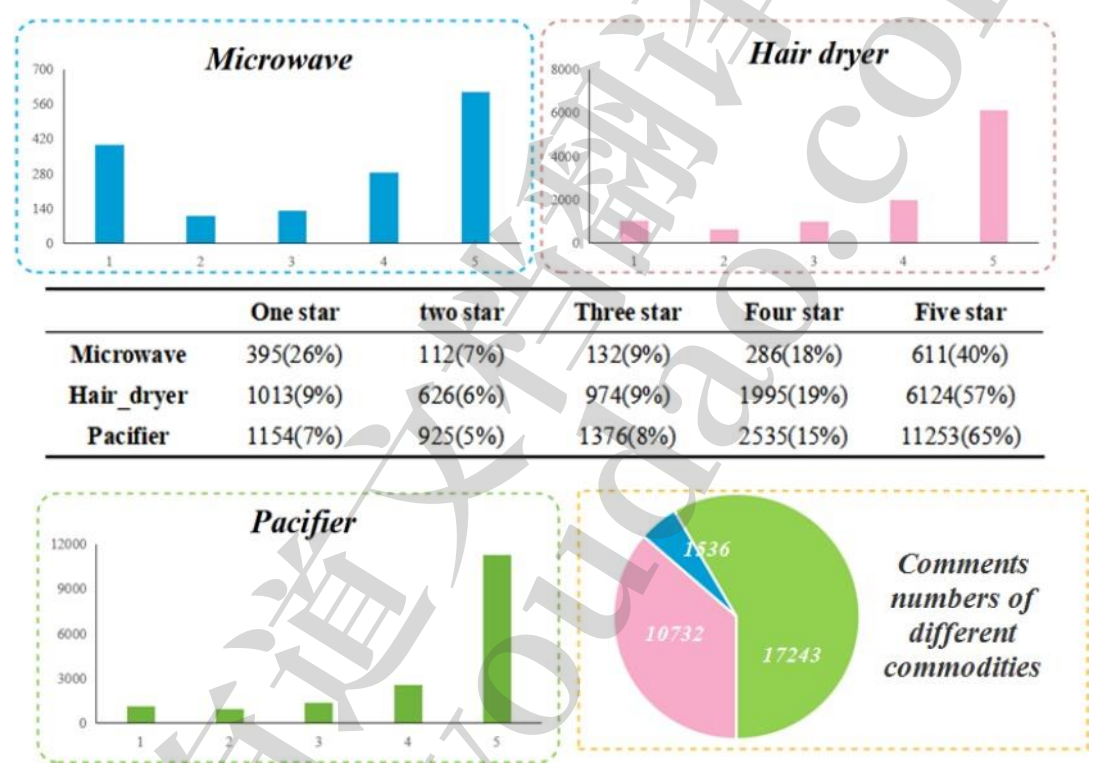


图 1 产品分析图的 satr_rating

如上表所示，是三个产品的星级评级频次表和占比。根据表中的数据，我们可以初步知道，五星评分在三款产品中占比最大，二星评分占比最小。对于微波而言，星级评分集中在 1 星和 5 星，而发乐者和安抚者的星级评分则集中在 5 星，其他星级的占比很低。

上图为三款产品的星级评分频率分布图。根据分布图可以得出一个结论，安抚奶嘴和头发播放器的人群购买者更容易给出好评。`Pacifier` 和 `hair player` 是亚马逊生产的比较成功的产品。而购买微波炉的买家则更倾向于给出较低的评分，说明亚马逊的微波炉并不是很成功。

(2)评论

顾客对产品的评价直接反映了产品的成功。根据产品评论，我们可以探究客户对产品的使用满意度，从而发现产品的受欢迎程度以及存在的问题。根据评论中排名前 50 的单词，我们绘制了一幅单词云图。如图所示



关注数学模型
获取更多资讯

下图是根据三款产品的评论绘制的词云图。



图 2 根据 3 条商品评论绘制的词云图

对于 microave 来说, 客户讨论的产品有按钮、空间、月份、价格、年份、尺寸、作品、柜台、小。反映客户情感态度的词有 great、well、problem、good、small、sampling、o old 等。它表示顾客对微波炉的评价是好是坏。阳光公司要关注产品本身的性能和价格, 对产品及时进行改进。

客户谈论的更多的是电吹风产品本身，比如线绳、长、定、热、空气、功率、价格、好看等。反映客户情感态度的词语包括 great、love、right、like、dry 等。说明客户对电吹风的评价普遍较好。阳光公司要注重产品的性能和外观，找出产品成功的因素。

对于安抚奶嘴来说，产品本身经常被客户从月、时间、萌度等方面讨论。反映顾客情感态度的词语包括 great、love 等。说明客户普遍对安抚奶嘴评价较高，阳光公司在婴儿市场做得不错。公司要注意产品的使用寿命。

(3)有用性评分

表 1 有用性评分一般分析表

产品 helpful_votes 总票数比例

微波 8996 10635 0.846 电吹风 24938 29202 0.854 奶嘴 15244 20708 0.736

表 2 star_ratings 和有用性评分的关系表

| | Ratings | helpful_votes | total_votes | helpful/total |
|---|---------|---------------|-------------|---------------|
| 1 | 9418 | 12798 | 0.735896234 | |
| 2 | 3052 | 4284 | 0.712418301 | |
| 3 | 4362 | 5778 | 0.754932503 | |
| 4 | 7948 | 9539 | 0.833211028 | |
| 5 | 24398 | 28146 | 0.866837206 | |

如上表所示，总有用票数，总票数

描述了投票和有用投票占子产品总投票的比例。从统计结果来看，电吹风客户的有效评论占比最大，安抚奶嘴的有效评论占比最小。

此外，我们思考星级评分和有用性评分之间是否存在关系。因此，我们通过星级来计算有用评论的百分比。



关注数学模型
获取更多资讯

有道文档翻译
pdf.youdao.com

表 3 星级有用评论占比分析表

| 比例 1 2 3 4 5 | | | | | | | | | | | | | |
|---|--|--|--|--|--|--|--|--|--|--|--|--|--|
| 安抚奶嘴 0.61 0.66 0.65 0.78 0.805 微波 0.806 0.70 0.80 0.82 0.91 电吹风 0.78 0.78 0.801 0.88 0.90 | | | | | | | | | | | | | |

根据明星的有效投票百分比分析，for
安抚奶嘴和吹风机，可以提供更有效的信息，给予高分的客户评价。而对于
microave 来说，无论是高分还是低分的客户评论，它都可以提供更有效的信息。

2.3.3 评论的情感分析

步骤 1. 把句子

因为这个话题设计中的评论等变量都存在于文本情境中，所以本体首先根据语法对要分析的文本进行划分。在本文中，使用内连接来提取文本中的情感词步骤 2。情感词的分类

根据情感词典，将上一步提取的词分为褒义词表、贬义词表、中性词表、否定词和程度副词，然后得到情感分类器表及其情感强度，

步骤 3. 删除功能词

虚词是指一类没有实际意义的词。与其他词相比，功能词使用频率较高，如“the”、“is”、“at”、“which”、“on”等。除此之外，还有一类词，比如“want”这个词。在搜索引擎中，这样的词对搜索结果帮助不大。它们的出现不仅难以帮助用户缩小搜索范围，还会降低搜索效率。因此，本文将这些词语从问题中剔除，从而提高文本利用效率。

步骤 4. 计算每条评论的情感得分

比如 hair_dryer 中 review_id 的评论标题是 R100TKSFK51G4K，标题是：“了不起的吹风机了不起的价格和很轻的重量”。分词去除功能词后的活跃词数为 7，负性情绪得分为 0，正性情绪得分为 0.2857，最终情绪得分为 0.2857。

3.任务 2

3.1 问题分析

为了探究评论和评级对客户评论(Customer reviews, CR)的影响，我们整理了七个变量并测试了它们的相关性。由于变量之间存在相关性，我们使用主成分法将周期变量划分为三个主成分。基于此，我们得到了客户评论(CR)的计算公式。结合 CR 和评论时间，我们计算出每条评论的客户满意度(CS)，并对 CS 进行标准化。然后，根据产品分类，可以得到购买者对特定产品的综合满意度(COM_S)。最后，我们根据综合满意度对安抚奶嘴、电吹风、微波炉等每个产品进行了排序，并分析了成功产品信息和不成功产品信息。

3.2 方法

在上述模型研究的基础上，我们引入了一个时间序列模型来研究时间因素对评论倾向性的影响。

(1)自回归模型(Autoregressive model):自回归模型是描述当前值和历史值之间关系的模型。它是一种利用变量本身的历史事件数据来预测自身的方法。



关注数学模型
获取更多资讯

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \epsilon_t$$

为当前值; μ 为常数项; p 是阶数; γ_i 是自相关系数, 和 ϵ_t 为误差值

(2)综合:ARIMA 模型最重要的部分是时间序列数据的平滑性。平稳性是要求通过样本时间序列得到的拟合曲线在未来短时间内能够沿着现有形态继续惯性, 即数据的均值和方差理论上不应该发生太大的变化。

(3)移动平均:移动平均模型关注自回归模型中误差项的累积。它可以有效地消除预测中的随机波动。

$$y_t = \mu + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i}$$
$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i}$$

3.3 解决问题

3.2.1 星级评定指标 CR 的确定

review_body 中包含情感信息和评论的字符数。我们选择基于情感词典的方法, 利用内连接提取情感词, 并根据情感词的得分计算每篇文本的情感得分。使用 SentimentQDAP 来表达情感得分。此外, 评论中的字符数也可以反映客户的星级评分结果。评论的字数为 number。

步骤 1. 影响星级的指标

在讨论星级评分的最终结果时, 如果不首先考虑时间因素, 只讨论评论和评分对星级评分的影响。我们称这个星级评分指标为客户评论。影响客户评论的主要因素是 helpful_votes (hv)、total_votes (tv)、sentimentqdap (sq)、star_rating (sr)、number (n)、vine (v)、verified_purchase (vp)、customer_id (ci)和 product_parent (pp)。变量之间可能存在相关性, 因此我们首先对变量的相关性进行检验。如果相关系数大于等于 0.3, 我们认为变量之间有很好的线性相关性。通过相关性检验, 我们发现部分变量之间存在相关性。

表 1 相关检验

| | star_rating | helpful_votes | total_votes | vine | verified_purchase | SentimentQDAP | number |
|-------------------|-------------|---------------|-------------|-------|-------------------|---------------|--------|
| star_rating | 1.000 | -.048 | -.069 | .020 | .142 | .335 | -.132 |
| helpful_votes | -.048 | 1.000 | .994 | .040 | -.093 | -.062 | .153 |
| total_votes | -.069 | .994 | 1.000 | .042 | -.100 | -.070 | .163 |
| vine | .020 | .040 | .042 | 1.000 | -.244 | -.028 | .117 |
| verified_purchase | .142 | -.093 | -.100 | -.244 | 1.000 | .117 | -.273 |
| SentimentQDAP | .335 | -.062 | -.070 | -.028 | .117 | 1.000 | -.322 |
| number | -.132 | .153 | .163 | .117 | -.273 | -.322 | 1.000 |

步骤 2. 主成分分析

主成分分析是利用降维的思想, 用较少的变量来解释原始数据中的大部分变量, 并将我们拥有的许多高相关性的变量转化为独立或不相关的变量。它通常会选择几个可以解释大部分的新变量



关注数学模型
获取更多资讯

变量，即所谓的主成分，用它们来解释数据的综合指标。

对我们方法的论证

根据 Bartlett 检验的结果，Bartlett 检验的 p 值小于 0.001，拒绝了零的假设，也就是说，研究数据可以通过主成分提取出来。

| 表 2 KMO 和 Bartlett 检验 Kaiser-Meyer-Olkin 检验抽样充分性。524 Bartlett 球形检验卡方值 | | |
|---|----|----|
| 145929.849 | | |
| | df | 21 |
| | 意义 | 组 |

主

占数据变异的 60-70%。主成分提取结果显示，前 3 个主成分的特征值大于 1，分别占总数据变异量的 31.025%、22.117% 和 16.597%，占总数据的 69.738%。提取的主成分个数为 3 个。

表 3 展示了变异统计量

| 初始特征值 截距和载荷方回路和载荷方 | | | | | | | | | |
|--|-------|--------|---------|-------|--------|--------|-------|--------|--------|
| 总变差累积-总变差累积-总变差累积% lative% tion% lative% tion% lative% | | | | | | | | | |
| 1 | 2.172 | 31.025 | 31.025 | 2.172 | 31.025 | 31.025 | 2.006 | 28.663 | 28.663 |
| 2 | 1.548 | 22.117 | 53.141 | 1.548 | 22.117 | 53.141 | 1.518 | 21.687 | 50.350 |
| 3 | 1.162 | 16.597 | 69.738 | 1.162 | 16.597 | 69.738 | 1.357 | 19.389 | 69.738 |
| 4 | 0.813 | 11.620 | 81.359 | | | | | | |
| 5 | 0.742 | 10.605 | 91.964 | | | | | | |
| 6 | 0.557 | 7.959 | 99.923 | | | | | | |
| 7 | 0.005 | 0.077 | 100.000 | | | | | | |

判读能力

提取的主成分应该具有一定的意义，即对研究内容的解读能力。解读能力说明前三个主成分符合研究的实际需求。各主成分对对应变量(已剔除相关系数小于 0.3 的数据)的解释能力，如下表所示：

主成分对应变量的解释能力
表 4 主成分对应变量的解释能力

| | 1 | 2 | 3 |
|--------------------|------|------|------|
| helpful_votes | .996 | | |
| total_votes | .995 | | |
| SentimentQDAP | | .801 | |
| star_rating_number | | .743 | |
| vine | | .503 | |
| verified_purchase | | | .786 |
| | | | .726 |

由上表可以看出，前三个主成分解释的变量信息与分类的基本相同。第一主成分(F1)是 helpful_votes 和 total_votes。第二主成分(F2)是 SentimentQDAP、star_rating 和 number。第三主成分(F3)是 vine、经过验证的购买和数量。相应的，第一主成分主要反映评论的投票信息，第二主成分主要反映评论的内容，第三主成分



关注数学模型
获取更多资讯

有道文档翻译
pdf.youdao.com

组件主要体现客户身份信息。前三个主成分的提取具有很好的结果解读能力。

基于此，我们可以推导出三个主成分的基本表达式：

$$\begin{aligned} F_1 &= 0.996 * hv + 0.995 * tv \\ F_2 &= 0.801 * sq + 0.743 * sr + 0.503 * n \\ F_3 &= 0.786 * v + 0.726 * vp \end{aligned}$$

步骤 3。客户评论的数据元素

根据主成分分析的结果，第一主成分反映评论的投票信息，第二主成分反映评论的内容，第三主成分反映客户身份信息。由此可见，投票信息是决定客户评论最终得分的主要因素，其次是评论内容，最后是客户身份。客户评论的高低直接反映了客户对产品的满意程度。通过给不同的变量赋予不同的权重，可以得到客户评论(customer reviews, CR)分数。CR 的计算公式如下：

$$CR = \alpha_1 * F_1 + \alpha_2 * F_2 + \alpha_3 * F3$$

3.2.2 综合满意度分析

步骤 1。客户满意度影响因素

作为经验最丰富的电商，亚马逊肯定会充分了解顾客的心理，尤其是会提取出被顾客认可最多的评论。据了解，亚马逊有一种专门的智能机器(机器学习模型)，通过复杂的 star 算法给产品打分。星型算法中权重最高的参数为：

- (1)评审日期。日期越长，越有价值；
- (2)有帮助的点击。当顾客点击“有用”，就意味着这篇评论是有用的。
- (3)该评论是否来自已验证购买(VP)。

除了上述三个 Amazon rolls 权重较大的参数外，star 算法还参考了其他权重较小的指标：

- (1)评论中的字数，即句子的长度；
- (2)评论原分，即各买家给出的评论评分；
- (3)评论的内容。在构建模型时，我们选择了一种基于情感词典的方法，利用内连接提取情感词，并根据情感词的得分计算每条评论的情感得分。

在上文中，我们使用主成分分析方法将决定客户评论(CR)得分的因素分为三类：评论投票信息、评论内容和客户身份信息。

这三个数据集的评审时间为 2002 年 3 月 02 日至 2015 年 8 月 31 日。我们选取 01/01/2016 作为标准时间，利用 SAS 软件的 intck 函数计算评论时间到 01/01/2016 之间的天数。这个区间内的天数就是评论的“年龄”。“年龄”越长，评论就越有价值。天与天之间的间隔用天(D)表示。

步骤 2。计算客户满意度

根据分析，我们可以引入影响顾客满意度的因素(CS)：helpful_votes (hv)，total_votes (tv)，SentimentQDAP (sq)，star_rating (sr)，number (n)，vine (v)，verified_purchase (Vp)和 day (D)，因此顾客满意度(CS)的基本表达是：

$$CS = \alpha * D + \alpha_1 * F_1 + \alpha_2 * F_2 + \alpha_3 * F3$$

将上述公式(1)、(2)、(3)代入

$$\begin{aligned} CS &= 0.3 * D + 0.3 * (0.996 * hv + 0.995 * tv) + 0.1 * (0.801 * sq + 0.743 * sr + 0.503 * n) \\ &\quad + 0.3 * (0.786 * v + 0.726 * vp) \\ CS &= 0.3 * D + 0.2988 * hv + 0.2985 * tv + 0.0801 * sq + 0.0743 * sr + 0.0503 * n \\ &\quad + 0.2358 * v + 0.2178 * vp \end{aligned}$$



关注数学模型
获取更多资讯

我们使用 0- 1minmax 归一化，将客户满意度得分缩放到[0,1]区间。归一化后的客户满意度用 CS *表示。

步骤 3。综合满意度分析

在获得每个客户的满意度后，可以根据产品分类得到产品购买者的整体满意度。综合满意度(COM_S)等于购买者对同一产品满意度的平均值，

$$COM_S = \sum_{i=1}^n CS *_i/n$$

根据产品综合得分分析，Pacifier 产品得分最高为 0.713，最低为 0。微波炉的产品综合得分最高为 0.533，最低为 0.01。电吹风的整体得分为 0.557，最低为 0.013。文字显示，对于安抚奶嘴、微波炉、电吹风这三个产品，公司最好的产品是安抚奶嘴。

步骤 4。分析成功的产品

分别分析三款产品综合满意度排名前十的产品。我们发现，parafierde 有很长的成功产品评论历史。评论字数在 150 字左右。正面评论占有效投票的比重很大。星评集中在 3-5。顾客几乎不是 verified_purchase 或 Amazon Vine 的会员。Microwave 在成功的产品评论方面有着悠久的历史。评论字数在 100 字左右。差评占比增加，有效票数减少，星评无集中趋势。大部分买家都是 verified_purchase，几乎没有亚马逊 Vine 会员。电吹风历史悠久，评论字数 50 字左右，正面评论占比大，有效投票数很少，星评集中在 3-5，半数买家是 verified_purchase，几乎没有。

这说明《奶嘴》和《电吹风》中成功产品的正面情感评价较多，顾客给出的星级评分也较高。微波既有正面情绪评论，也有负面情绪评论，星级评分也比较分散。这与我们在分析星级和有效票数时初步得出的结论是一致的。

对于微波的成功产品来说，拥有 verified_purchase 身份的买家更多，其次是电吹风，最后是安抚奶嘴。微波成功产品的评价最靠谱，其次是电吹风，Pacifier 成功产品的评价不太靠谱。而且评论人是不是 Amazon Vine 的会员，和产品的成功没有任何关系。

也可以看到，verified_purchase 或 Amazon Vine 会员的评论与有效投票数没有显著关系。说明客户判断评论是否有用，不会受到评论人身份的影响。

此外，所有评分较高的产品在正面情感评论中所占比例较大。而所有产品中评分较低的产品，正面情感评论占比比之前更大。我们无法直接判断评论情感与综合评分的关系。

表 5 安抚奶嘴 COM_S 最高的 10 款产品

| COM_S | PRODUCT_TITLE | PRODUCT_PARENT |
|-------|-----------------------------------|----------------|
| 0.713 | 中远 alpha omega 精英敞篷汽车座椅 | 485572042 |
| 0.703 | Sealy Baby Soft 超床垫 | 409107492 |
| | Fisher45;价格水族馆监视器 | 61825188 |
| 0.696 | Velour Boppy 护理枕-绿色 | 108369347 |
| 0.693 | Fisher-price 婴幼儿摇滚乐手蓝色/红色 | 32051425 |
| 0.691 | 海军蓝的终极婴儿包裹 | 234474039 |
| 0.691 | Ameda 纯你的吸奶器，随身携带 | 005510005 |
| | Baby jogger performance 系列双 20/海军 | |



关注数学模型
获取更多资讯

安抚奶嘴产品中，中远阿尔法、费雪、阿米达、西利等品牌表现较好。同理，我们可以得出以下结论。在微波炉产品中，夏普、lg、魔厨、惠而浦、profile、熊猫小等品牌表现不俗。电吹风产品中，沙龙、Waterpik、微镊子、DuWop、ISLAND HAWAII、Oral-B 等品牌做得比较好。安抚奶嘴产品中，chicco keyfit、宜家 kladd、batman 等品牌表现不佳。微波炉产品中，三星、丹比、jx7227sfss 等品牌表现不佳。在电吹风产品中，andis、conair、xtava、mhd、remington 等品牌表现较差。

4.任务 3

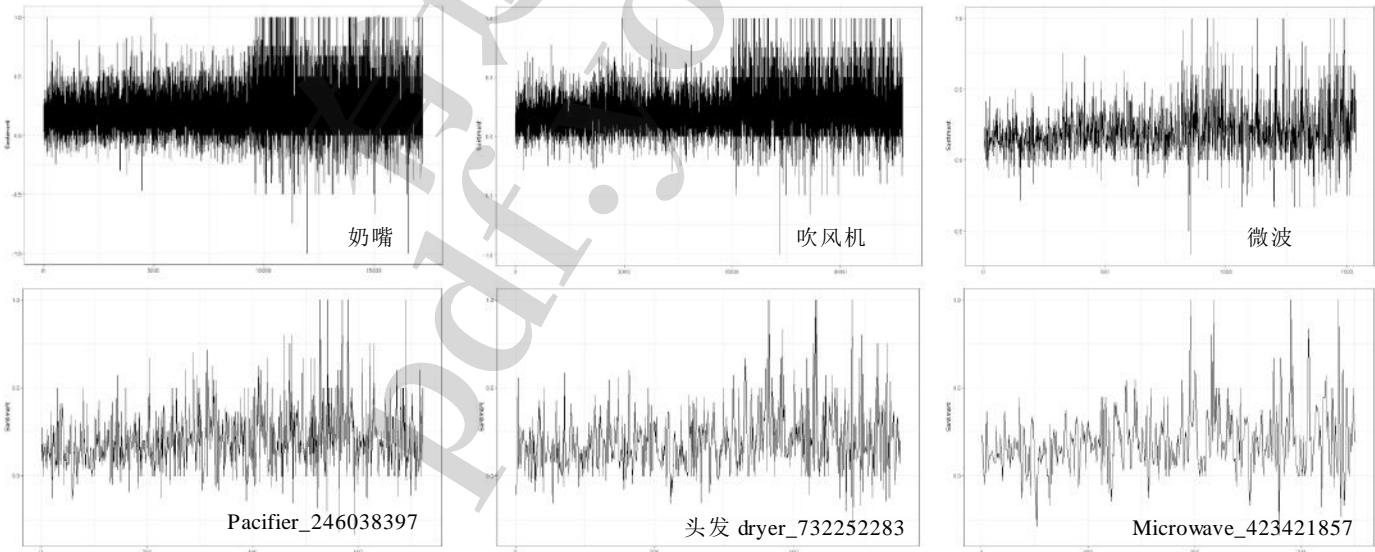
4.1 问题分析

为了探究客户评论随时间的变化，并给出产品在市场上的信誉度趋势。在求解 task2 时，首先利用以天为单位有断点的原始时间序列数据计算累积动态均值，获得以月为单位的连续时间序列数据，并结合时间序列 ARMA 模型对数据进行建模拟合和预测未来趋势。接下来，计算每个具体产品在每次评分前后的评论数量变化和评论情感变化，并使用非参数检验来确定星级评分对评论的影响。最后，找出所有情绪极端的评论，使用非参数测试确定是否对星级评分有影响，并结合 KMEANS 聚类给出客户评论的特征，讨论潜在的虚假评论。

4.2 问题解决

4.2.1 三个变量的时间序列可视化和预测

绘制折线图可以直观地展示情感得分的演变，可以看到商品评论情感得分在时间维度上的变化趋势。接下来的时间序列模型中的数据都是按月处理的，因为原始数据是按天计算的，但也有连续很多天没有评论的情况，所以我们按月整合后对所有数据进行分析，然后为了保证数据的有效性，我们过滤掉了前几个月有少量有效评论影响建模质量的月份。我们发现，极端情绪评论出现的概率较低，但随着时间的增加，商品销量的增加，极端情绪的数量也会相应增加。下图是安抚奶嘴、电吹风、微波炉这三种商品评论的情绪变化趋势。可以看出，商品评论中的情感得分随着时间的推移而增加，波动性增大。一般来说，评论中的积极(积极)情绪要高于消极(消极)情绪。产品模型



关注数学模型
获取更多资讯

从这三款产品中提取评论数量最多的，它们的 product_parent 代码分别为 246038397 (philips avent bpa free soothie 安抚奶嘴，0-3 个月，2 包，包装可能有所不同)，732252283 (remington ac2015 t | studio salon collection pearl 陶瓷吹风机，深紫)，423421857 (danby 0.7 cu.ft. countertop microwave)，三款产品的情绪得分均为正，但安抚奶嘴和吹风机的情绪得分呈下降趋势，而微波类产品则处于后期的上升阶段。

图 3 三款产品情感图

图中是安抚奶嘴、电吹风、微波炉三种产品评分随时间变化的折线图。从垂直线(2005 年 4 月)观察，星星呈上升趋势，最终星星分为 4.26 颗。点赞数上升较快，呈指数上升趋势;H 产品前期波动较大。从 2005 年到 2006 年，星评经历了大幅下滑，然后缓慢反弹。2015 年 8 月的最终评分为 4.08。作为耐用品，点赞数远高于评论数，即“看一看”更多是给购买者的。在经历了 2004 年下半年的大幅下滑后，M 产品的星级评分逐渐回升。2008 年初达到峰值后，逐渐回落。在 2012 年年中跌至谷底后，逐渐反弹。2015 年 3 月的最终得分为 3.39 分。还会继续缓慢上升。值得注意的是，M 作为耐用品的销量和评论都远低于两类产品，但它的点赞数却远高于评论数，累计点赞数为 8996，累计评论数为 1536，相差几乎 5 倍，而且在某些时间段点赞数是逐步上升的。2014 年底，随着葡萄树评论的突然增加，点赞数急剧上升。

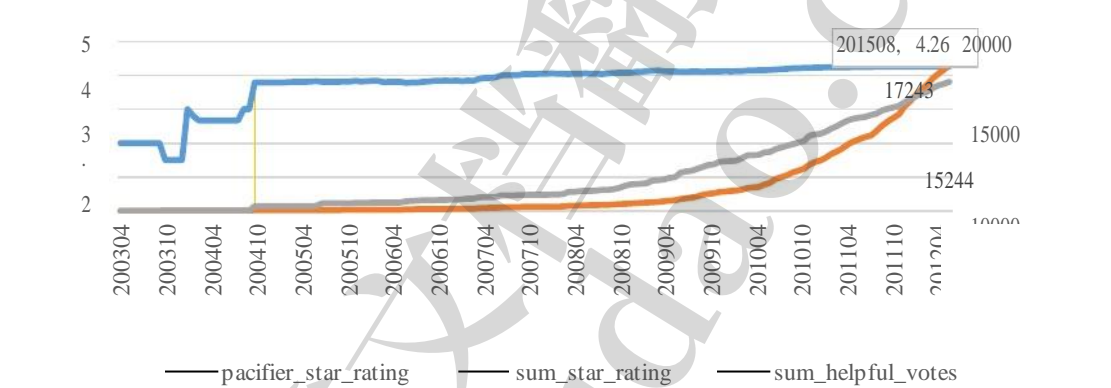


图 4 星级评分、评论和有用性评分累计数量的安抚者趋势图

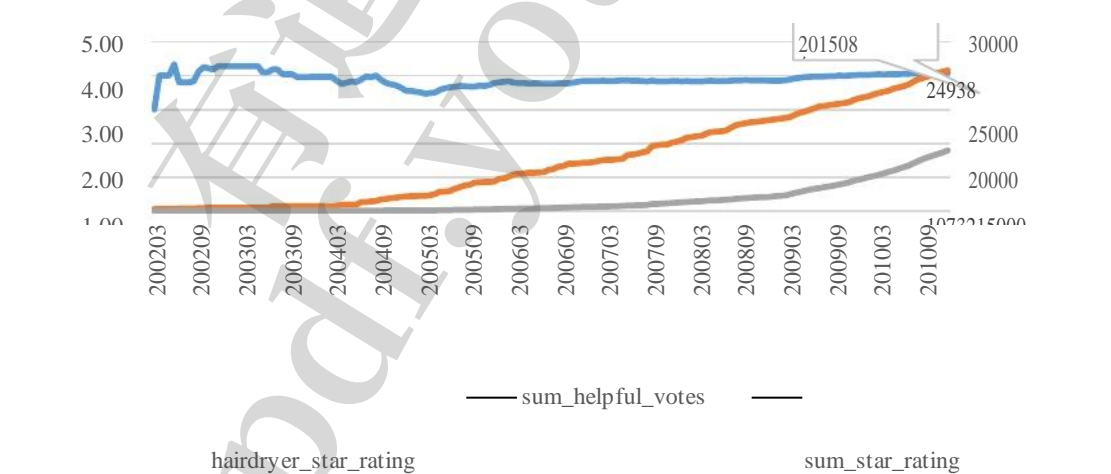


图 5 电吹风的星级评分、评论和有用性累计数量趋势图



关注数学模型
获取更多资讯

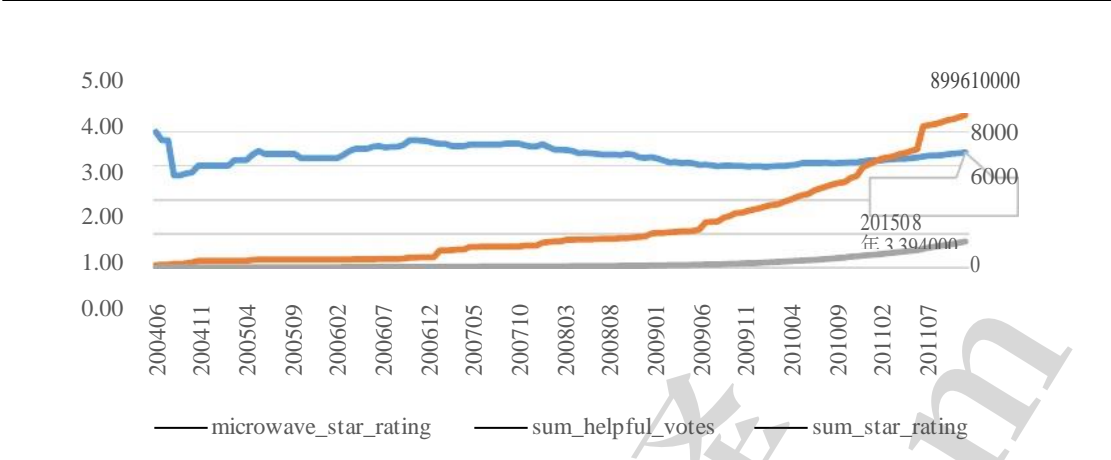


图 6 星级评分、评论和有用性评分累计数量的微波趋势图 4.2.2 星级评分预测

然后我们使用时间序列模型来预测星级评分。首先，我们使用上面描述的时间序列可视化来获取三款产品的主要趋势。然后我们检查时间序列的稳定性，并绘制 ACF / PACF 图来寻找最优参数。然后，建立差分自相关移动平均模型 (ARIMA) 模型，可以预测 2015 年 9 月和 10 月的星级评级如下：

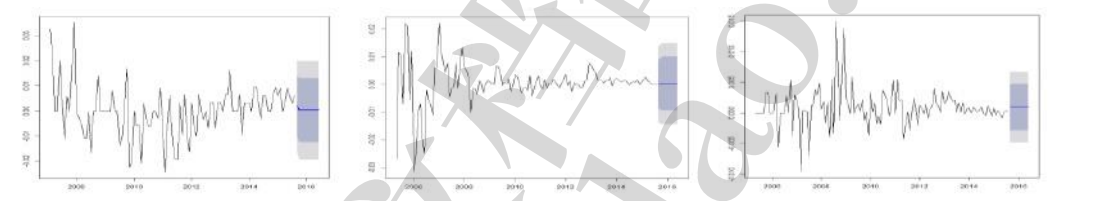


图 7 星级评分预测图

9 月微波炉星级评分预测:3.401379;10 月:3.395562。9 月吹风机星级评分预测:4.080642;10 月:4.081117。9 月安抚奶嘴星级评分预测:4.268717;10 月:4.268825。4.2.3 评论情绪预测

我们可以用同样的方法预测这三款产品未来三个月的情感趋势。结果如下表所示。与从情感趋势图中得到的结果类似，未来对 P 和 H 产品的评论的情感得分会下降。P 产品的负面情绪和正面情绪会先增加后减少，而 H 产品的负面情绪不会有太大变化，正面情绪的下降更严重，反映了客户积极性可能存在的问题；M 作为耐用品，销量和评论相比其他两种产品都非常少，相对较多的数据更符合正态分布，相对较少的数据波动更大。因此，M 产品未来评论的情绪趋势是上升、下降、上升的波动趋势，但总体是上升的。

表 6 未来 3 个月 3 个产品评论情感图

| date | sentiment | Aug-15 | Sep-15 | Oct-15 | Nov-15 | mini-Line chart |
|------------|-----------|--------|--------|--------|--------|-----------------|
| pacifier | negative | 0.0556 | 0.0759 | 0.0626 | 0.0595 | |
| | positive | 0.2552 | 0.3322 | 0.2838 | 0.2641 | |
| | sentiment | 0.1996 | 0.2564 | 0.2212 | 0.2047 | |
| hair dryer | negative | 0.0679 | 0.0679 | 0.0679 | 0.0679 | |
| | positive | 0.2654 | 0.2714 | 0.2666 | 0.2652 | |
| | sentiment | 0.1975 | 0.2035 | 0.1987 | 0.1973 | |
| microwave | negative | 0.0606 | 0.0893 | 0.0653 | 0.0636 | |
| | positive | 0.2501 | 0.2760 | 0.2415 | 0.2517 | |
| | sentiment | 0.1896 | 0.1867 | 0.1762 | 0.1881 | |

4.2.4 时间序列分析

首先，选择星级评分，next_past, next_past_sen，并使用多个独立样本非参数检验。检验方法为 J-T 检验 (Jonckheere-Terpstra):



关注数学模型
获取更多资讯

由结果可知， p_1 近似为 0，而 $p_2 = 0.531 > 0.05$ 。在 0.05 的置信度水平下，不同明星评论发生前后的评论数量存在显著差异，但前后评论的情感倾向没有显著差异。

查看一星评价是否会引起评论和评论情感上的显著变化。选取一星样本对应的前后评论数和评论情感，使用 Wilcoxon 对两个成对样本进行非参数检验。

表 7 检验统计量

| | Past_sen - next_sen 过去-下一步 | |
|------------------------|----------------------------|----------------------|
| Z | -39.067 ^b | -37.732 ^b |
| Asymp. Sig (2-tailed)。 | 组织 | 组织 |

a. Wilcoxon Signed Ranks Test

b.基于正秩。

P 近似为 0，且在显著性水平 0.001 下存在显著性差异，即一星评论会导致评论数量发生显著变化，评论情绪发生显著变化。

同理，检查五星评价是否会导致评论发生显著变化:

P 接近 0，且在显著性水平 0.001 存在显著性差异，即一篇五星评论会导致评论数量和评论情绪发生显著变化。

综合以上，不同星级评论前后的评论数量存在差异，因此不能判断特殊星级评论会影响评论数量，但一星、五星等特殊星级评论会引起评论情绪的变化。

表 8 最终的聚类中心

| | 集群 | | |
|---------------|------|-------|------|
| | 1 | 2 | 3. |
| 明星 | 3.55 | 3.90 | 4.17 |
| next_past_sen | 56 | -1.44 | 6.66 |

K-MEANS 聚类基于每个物品的星级评分和一星评论前后的情感变化。可以发现，表中有 1 组产品的星级评分最低，受一星评论影响最小;2 组发表的一星评论差评较多。这些产品的星评为 3.9，算是比较高的了。可能是客户的星级评分过高，与产品不匹配，导致了大量差评和一星评分。此类产品中的高、1 星评论会引起客户大量正面评论。可能出现的情况有:第一，当一个好产品失信时，顾客会更加正面地赞美它;影响，好评现象出现。

4.2.5 star_rating 与 review 的互动

根据 product_parent 和 date 对数据进行汇总，得到某类产品(不包括本次评论)评论前后的评论数量和评论情感。为了消除时间的影响，两条评论之间的时间越长，相互影响越小，因此时间权重为 $1 / [(date_after-date_before) + 1]$ ，其中+1 是处理在同一天分母为零时发生的评论。

评论发生后，对评论号 next 和评论情绪 next_sen 进行变量处理的伪码为:



关注数学模型
获取更多资讯

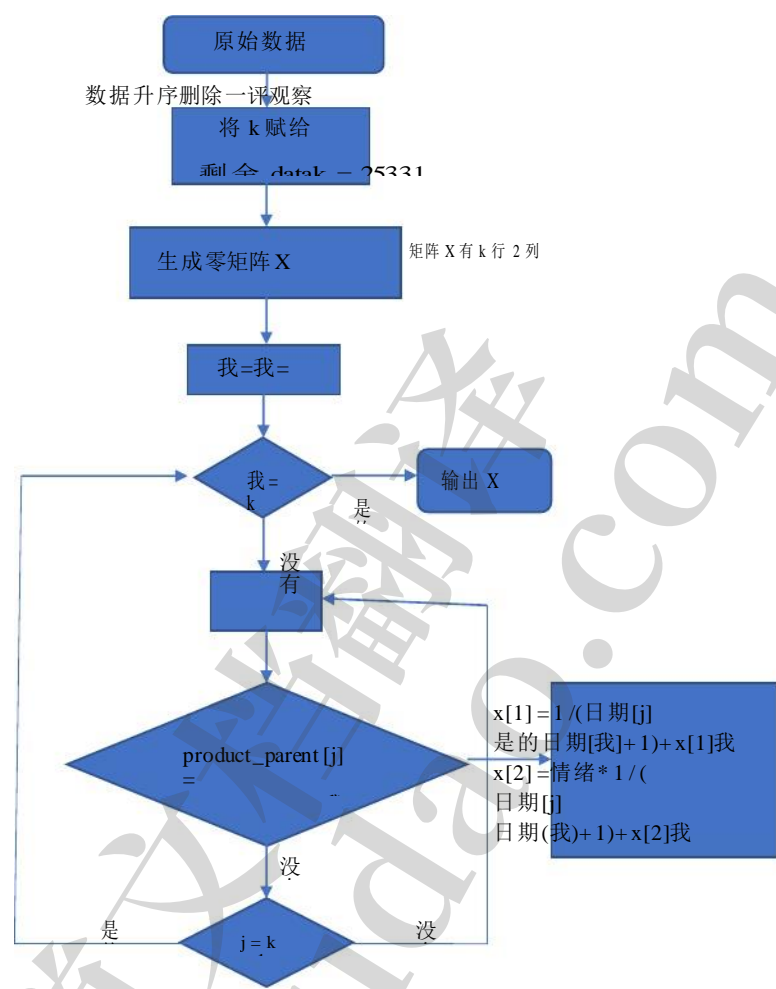


图 8 变量处理伪代码流程

将原始数据集按日期降序排列，根据上述伪代码计算过程，可以得到评论发生前的评论数，以及评论情感变化的 `past_sen`，不重复，然后可以得到某条评论发生前后评论数变化的 `next_past` 和评论情感变化的 `next_past_sen`。经过测试，两者都不符合正态分布，因此选择非参数测试进行后续测试。

我们选取了包括失望在内的 17 个极端消极词汇和包括热情在内的 17 个极端积极词汇来获得极端情感值 `NEG` 和 `POS`。

表 9 最终的聚类中心

| Cluster | 1 0.1283 | 2 0.1321 | 3 0.1699 | 4 0.1334 | 5 0.0399 | 6 0.0909 | 7 0.0982 | 8 0.1235 |
|-------------------|----------|----------|----------|----------|----------|----------|----------|----------|
| COM_S | 4 | 4 | 4 | 4 | 5 | 4 | 4 | 4 |
| star_rating | | | | | | | | |
| helpful_votes | 1 | 156 | 23 | 331 | 814 | 0 | 17 | 1 |
| 葡萄树 | 0 | 年 0 | 日 0 | 年 0 | 年 1 | 0 | 日 0 | 0 |
| verified_purchase | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| review_zishu | 51 | 115 | 291 | 67 | 233 | 38 | 254 | 44 |
| 底片 | 0 | 年 0 | 年 0 | 年 0 | 年 1 | 0 | 年 0 | 0 |
| POS | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 评论 | 46 | 205 | 17 | 518 | 43 | 722 | 619 | 379 |

通过石头图确定分类数量，利用 `Kmeans` 聚类将整个样本分为表中的 8 个类别。可以看到，极端否定词出现在第 5 类中，其特点是低



关注数学模型
获取更多资讯

有道文档翻译
pdf.youdao.com

产品满意度，很少评论，但星级。高评分，高有效点和 VINE 评论数，长词评论，虚假信息。这样的产品，要么是被恶意抹黑过的好产品，要么是被吹捧过的坏产品；

正面词汇出现在第 3 类和第 7 类中，其中第 3 类产品满意度高，评论和点赞较少，评论较长，vine 评论较少，验证购买认证较少，质疑评论的有效性和可信性，或者存在好评现象;第 7 类产品满意度可接受，评论数高，点赞数少。随着评论数量的增加，出现极端词汇的概率也会相应增加，所以在这个群体中并不存在异常评论行为。

表 10 否定词相关性表

| | | star_rating | 负的 |
|-------------|-----------------|-------------|--------|
| star_rating | 皮尔森相关 | 1 | -209** |
| | Sig (2-tailed)。 | | 组 |
| 负的 | N | 29511 | 29511 |
| | 皮尔森相关 | -.209** | 1 |
| | | | |
| | N | 29511 | 29511 |

**相关性在 0.01 水平上显著(2-tailed)。

| | | | |
|-------------|---------|-----------------|-------|
| star_rating | 皮尔森相关 | 1 | 本场.. |
| | Sig (2- | | 组 |
| POS | N | 29511 | 29511 |
| | 皮尔森相关 | 年本 场比 赛** | 年 1 |
| | Sig (2- | 组 | |
| | N | 29511 | 29511 |

**相关性在 0.01 水平上显著(2-tailed)。

可以看出，在 0.01 的置信水平下，极端情绪与星级评分之间存在相关性。极端负面情绪与星级之间存在负相关关系。极端积极情绪与星级评分呈正相关。系数均小于 0.3，相关性较低。考虑到极端情绪词汇出现的频率，进行非参数检验。结果如下：

表 12 检验统计量

| | | star_rating |
|-----------------------|--|---------------|
| Mann-Whitney U | | 5310158.500 |
| Wilcoxon W Z | | 5655854.500 |
| Asymp。Sig (2-tailed)。 | | -31.131 |
| | | 组织 |
| a.分组变量:NEG | | |
| | | star_rating |
| Mann-Whitney U | | 63602803.000 |
| Wilcoxon W Z | | 266080429.000 |
| Asymp。Sig (2-tailed)。 | | -43.030 |
| | | 组织 |

a.分组变量:POS

可以看出，对于带有极端情绪词汇的评论，星级评分与其他星级评分存在显著差异，可以说明极端情绪与星级评分确实存在关系。极端负面情绪与星级评分的关系为负相关，极端积极情绪与星级评分的关系为正相关。



关注数学模型
获取更多资讯

有道文档翻译
pdf.youdao.com

5.模型评价

5.1 模型局限性

时间序列分析:

- (1)模型存在预测误差的缺陷，忽略了其他随机变量的变化。当外界有较大变化时，往往会出现较大偏差。
- (2)时间序列预测方法可以有效预测短期内变化的影响。因为客观事物，特别是经济现象，在较长的时间内更容易改变外部因素，所以必然对市场经济现象产生重大影响。

Lexicon-based 方法:

- (1)评论的情感得分是其所有句子得分的平均值。这种方法并不符合实际情况。就像文章中的段落重要性不同一样。在一个句子中，句子前后的句子也有重要的区别。
- (2)有一类文本用贬义词来表示褒义。基于词典的方法无法识别这种情况。对于这个问题，需要深度学习的方法来有效解决。用普通的机器学习方法很难解决这个问题。
- (3)对于正负评论的判断，算法忽略了许多其他否定词、程度副词和情感词;对情绪强弱的判断过于简单。
- ④基于词典的方法准确率高(80%以上)，但人工工作量比较大。
- (5)忽略了表情符号的存在以及评论中表达的情绪。表情包往往带有一定的情感极性，更清晰地表达了顾客的情绪。但在情感得分上，表情符号被忽略了。

主成分分析缺点:

- (1)主成分分析适用于变量之间相关性较强的数据。如果原始数据弱相关，则不会起到很好的降维作用。
- (2)降维后有少量的信息损失，不可能包含 100% 的原始数据。
- (3)原始数据标准化后，含义会发生变化，主成分的解释含义会比原始数据更加模糊。

5.2 模型强度

时间序列分析:

时间序列分析预测方法突出了时间因素在预测中的作用，暂时不考虑具体外部因素的影响。时间序列分析进行的是成交量预测。事实上，所有的影响因素都归结于时间因素，只认可所有影响因素的综合效应，未来仍然会起到预测对象的作用。输入变量可以根据时间序列模型进行调整，使系统开发过程保持在目标值，即在预测过程偏离目标值时，可以进行必要的控制。通过分析时间序列，可以做出合理的预测，提前掌握未来的发展趋势，为业务决策提供依据。

Lexicon-based 方法:

基于词典的方法不仅在准确率、精度、召回率和 F-measure 方面优于监督机器学习方法，而且在使用的时间和精力的经济方面。

ARIMA 模型实现了更准确的预测，它是一种简单的建模技术。



关注数学模型
获取更多资讯

主成分分析的优点:

- (1)数据不要求服从正态分布。主成分是向数据最离散的方向旋转基组。
- (2)通过对原始变量的综合和简化，可以客观确定各个指标的权重，避免主观判断的随意性。最终的结果只与数据有关，与模型选择无关。
- (3) PCA 将原始指标转化为独立的主成分。实践表明，PCA 是一种丢失原始数据信息最少的线性降维方法。

6.结论

1.通过分析 star_ratings、reviews、helpfulness ratings、SentimentQDAP，初步得到三个产品的整体特征。

- (1)对于微波，星级评分集中在 1 星和 5 星，而电吹风和安抚奶嘴则集中在 5 星。此外，电吹风和安抚奶嘴提供的高评分客户评论，可以提供更有效的信息。可以假设，安抚奶嘴和吹风机是亚马逊比较成功的产品。
- (2)对于微波炉，公司要注意产品的性能和价格，并及时对产品进行改进。对于电吹风，公司要重点关注产品性能和外观问题，找出产品成功的因素。对于安抚奶嘴，评价普遍较高，公司要注意产品寿命。

2.利用时间序列可视化，在时间维度上刻画评论情感和星级的变化趋势、累计评论数、累计点赞数。

- (1)随着时间的流动和商品销量的增加，极端态度的数量也会相应增加。
 - (2)三篇商品评论中积极(积极)情绪高于消极(消极)情绪。但安抚奶嘴和吹风机的的情绪评分呈下降趋势，微波产品在后期处于上升阶段。
- 3.使用 ARIMA 模型预测情感得分。一旦产品在市场上销售，根据客户的评分和评论，可以推断出产品的评分会在一定的时间内。公司对未来发展趋势的把握，可以为商业决策提供依据。以下是三种商品未来三个月的情绪得分预测结果:

- (1)安抚奶嘴的负面情绪和正面情绪均呈现先升后降的趋势。
 - (2)电吹风的负面情绪变化不大，正面情绪下降比较严重，反映了客户积极性可能存在的问题。
 - (3)微波炉作为耐用品，销量和评论都少于其他两种产品。未来评论的情感趋势是上升、下降、上升的趋势，但总体呈上升趋势。
- 4.采用主成分分析法确定客户评论的决策指标。第一主成分是投票信息，第二主成分是评论内容，第三主成分是客户。身份信息。随着评审“年龄”的影响，得出了确定客户满意度(CS)的计算公式。综合满意度(COM_S)则源自同一产品的 CS 汇总。
- 5.利用非参数测试方法，我们得出结论，特殊的星级评论(一星、五星)可以引起评论中的情感变化。
- 6.使用非参数检验方法，极端态度评论与评分存在相关性:极端态度与星级评分存在相关性，极端消极态度与评分存在负相关，极端积极态度与评分为:正相关。



关注数学模型
获取更多资讯

7.参考文献

[1] Carolanne, M.(2017)。理解在线客户评论的价值

[2] Gamal, D., Alfonse, M., El-Horbaty, E., & Salem, A.(2019)。使用 N-Gram 特征实现阿拉伯语情感分析中的机器学习算法。《*procdia Computer Science*》，2019 年第 154 卷，第 332-340 页

[3] Sailunaz, K., & Alhajj, R.(2019)。来自推特文本的情绪和情绪分析。《*计算科学杂志*》，第 36 卷，2019 年 9 月，101003

[4]年代,魏,Z,王,Y.,与辽、t(2018)。基于扩展情感词典的中文微博文本情感分析。《*未来一代计算机系统*》，卷 81,2018 年 4 月，第 395-403 页

[5] Bernabé-Moreno, J., Tejeda-Lorente, A., Herce-Zelaya, J., Porcel, C., & Herrera-Viedma, E.(2020)。一种上下文感知嵌入支持的方法来提取模糊情感极性词典。《*基于知识的系统*》，卷 190,2020 年 2 月 29 日，105236

[6] Mukhtar, N., Khan, M. & Chiragh, N.(2018)。基于词典的方法在多个领域的乌尔都语情感分析中优于有监督的机器学习方法。《*Telematics and Informatics*》，卷 35，第 8 期，2018 年 12 月，页 2173-2183

[7] Matyjaszek, M., Fernández, P., Krzemień, A., Wodarski, K., & Valverde, G.(2019)。利用 ARIMA 模型和神经网络预测焦煤价格，考虑转基因时间序列理论。《*资源政策*》，第 61 卷，2019 年 6 月，页 283-292

[8] Nguyen, H., AsifNaeem, M., Wichitaksorn, N., & Pears, R.(2019)使用时间序列模型进行短期价格预测的智能系统。《*计算机与电子工程*》，卷 76,2019 年 6 月，339-352 页



关注数学模型
获取更多资讯

有道文档翻译
pdf.youdao.com