



北京大学

硕士研究生学位论文

题目： 基于在线用户评论的
产品设计优化研究

姓 名：	钟易衡
学 号：	1401210891
院 系：	软件与微电子学院
专 业：	计算机技术
研究方向：	电子商务与物流
导师姓名：	李杰 教授

二〇一七 年 六 月

版权声明

任何收存和保管本论文各种版本的单位和个人，未经本论文作者同意，不得将本论文转借他人，亦不得随意复制、抄录、拍照或以任何方式传播。否则，引起有碍作者著作权之问题，将可能承担法律责任。

摘要

国内电子商务的发展已逐渐走向成熟市场，各种商业形态的拓展及完善，使国内商业市场更加活跃，现今文化创意产业也随着互联网时代的大潮复苏，设计端在各个企业的重要度更加提升，伴随着现今竞争市场对互联网思维的重视，不断聚焦客户需求，消费者的需求被更加细化，企业如何在复杂度不断提升的市场中去了解消费者呢？现今在线调查方法有别于传统调研，在线评论自由式的文本内容，更能彰显消费者心理的真实想法，尤其是带有强烈感情色彩的评论更具有分析价值，这些真实信息是企业可挖掘的金矿，本论文针对难以量化的艺术、工艺类商品作为分析对象，主要分为三个部分。

首先，探讨国内的电子商务背景概况，并介绍在电子商务日渐成熟的背景下国内文化产业现况，并对国内外有关在线评论的研究现况作的探讨，进而提出改善方案确立本论文之框架。

其次，对产品设计优化的相关理论与方法进行梳理，并针对电子商务环境下如何基于用户评论进行产品优化设计的方法研究，对此基于无监督学习文本挖掘和 KANO 模型特征需求，并结合专家方法，针对难以量化的商品，提出一个新的分析模型框架，将用户需求特征转化为消费者偏好属性，与产品设计的**技术指标**作相关度分析，进而得出各产品设计**技术流程**的优化顺序。

最后，本论文以在线用户评论为基础，结合以上所提出的分析模型，对紫砂壶产品进行了优化设计分析，给予设计师产品制作流程的建议，为产品优化设计提供参考和依据。

关键词：产品设计优化，QFD，Kano 模型，在线评论，无监督学习

The Research on optimization of product design based on the online review

Chung Yiheng(Computer Technology)

Directed by Professor Jie Li

ABSTRACT

The development of domestic e-commerce has become a mature market. Different kinds of business models and various types of mechanisms to make the domestic business market more active. Along with the current competitive market, the cultural and creative industries with the Internet era tide recovery. The importance of Internet thinking and design area in each enterprise got more important.

Focus on the customer needs to continuous refinement, by reading the customer online reviews can intuitive understanding of product and service experience. Different from the traditional way of research online reviews can clearly express customer real thoughts of especially with strong feelings of the comments. Real reviews is more intuitive and have more analytical value. These real information is the enterprise gold mine

This paper have three parts:

First, introduces the domestic background of e-commerce, and understands the current situation of domestic cultural industry under this business background. Discusses the research of online reviews at home and abroad, and puts forward the improvement plan.

Second, according the theory and method of product design optimization put forward the method of product optimization design and the analysis of product characteristic attribute based on unsupervised learning text mining. The user demand characteristics of KANO model turn into consumer preference attributes to connect the product design technical indicators, and then analyze the optimal order. This method combined with the expert method. It is a new analysis model framework,

Finally, based on the online user reviews and the analysis model. This paper gives the optimal design plan of the teapot product, and suggestions of the designer's production process.

KEY WORDS: Kano model, QFD, Online review, Product design, Unsupervised learning

目录

摘要	错误!未定义书签。
ABSTRACT	错误!未定义书签。
目录	错误!未定义书签。
第一章 绪论	错误!未定义书签。
1.1 研究背景与意义.....	错误!未定义书签。
1.1.1 国内电子商务发展现况.....	错误!未定义书签。
1.1.2 国内文化产业政策现况.....	错误!未定义书签。
1.1.3 国内外研究现况.....	错误!未定义书签。
1.2 研究目标内容与方法.....	错误!未定义书签。
1.3 论文结构与组织.....	错误!未定义书签。
第二章 相关理论与方法	错误!未定义书签。
2.1 产品优化设计调研方法概述.....	错误!未定义书签。
2.2 无监督学习方法.....	错误!未定义书签。
2.2.1 中文分词.....	错误!未定义书签。
2.2.2 TF-IDF	错误!未定义书签。
2.2.3 PCA 与 KPCA.....	错误!未定义书签。
2.2.4 K-means 聚类.....	错误!未定义书签。
2.3 调研与专家法.....	错误!未定义书签。
2.3.1 德尔菲法.....	错误!未定义书签。
2.3.2 Kano 模型.....	错误!未定义书签。
2.3.2 QFD 质量功能展开.....	错误!未定义书签。
2.4. Kano 模型集成到 QFD.....	错误!未定义书签。
2.4.1 整合方法.....	错误!未定义书签。
第三章 优化设计方法	错误!未定义书签。
3.1 提取消费者偏好属性.....	错误!未定义书签。

3.2 调研与专家法分析.....	错误!未定义书签。
3.3 提出优化方案.....	错误!未定义书签。
第四章 基于用户评论的紫砂壶设计优化案例	错误!未定义书签。
4.1 数据准备与采集.....	错误!未定义书签。
4.2 数据分析.....	错误!未定义书签。
4.3 专家模型建模与分析.....	错误!未定义书签。
4.4 产品优化方案.....	错误!未定义书签。
第五章 结论	错误!未定义书签。
参考文献	错误!未定义书签。
附录	错误!未定义书签。
致谢	错误!未定义书签。
北京大学学位论文原创性声明和使用授权说明	错误!未定义书签。

第一章 绪论

1.1 研究背景与意义

1.1.1 国内电子商务发展现况

近几年，互联网的发展造就了国内整体经济产业的转型，线上与线下的交互配合也创造出各式各样的行业形态，进而方便了全体国民的生活便利性，电子商务产业在近几年发展趋势也持续向上，根据中国电子商务研究中心的数据截止 2016 年上半年，电子商务市场交易规模已达到 10.5 万亿元[1]（见图 1.1）而网络用户规模 2016 年上半年已达 4.8 亿人，相当于中国人口的 38%，在发展渐趋成熟的产业，对于如此庞大的用户数量，也逐渐重视消费者体验且做出各方面的研究。

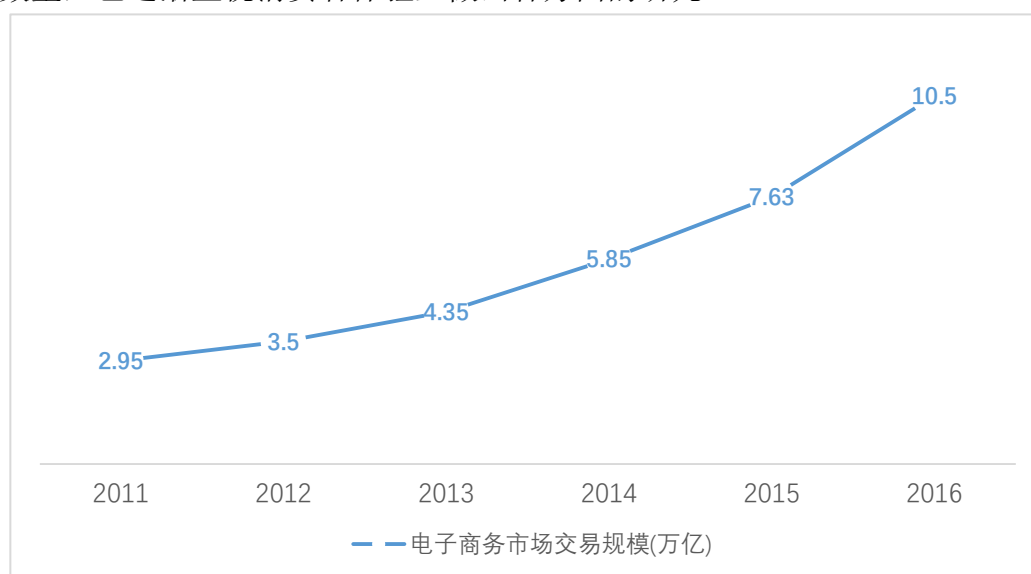


图 1.1 历年电子商务市场交易规模

数据源：中国电子商务研究中心 www.100ec.cn [1]

根据中国电子商务投诉与维权公共服务平台监测数据显示，2016 年上半年电子商务一共分为以下投诉领域（见图 1.2），网络购物投诉占全部投诉 52.75%，比例最高，生活服务 O2O，占据 21.03%，物流快递占 4.20%，与生活相关的数据就大约占了 78%。

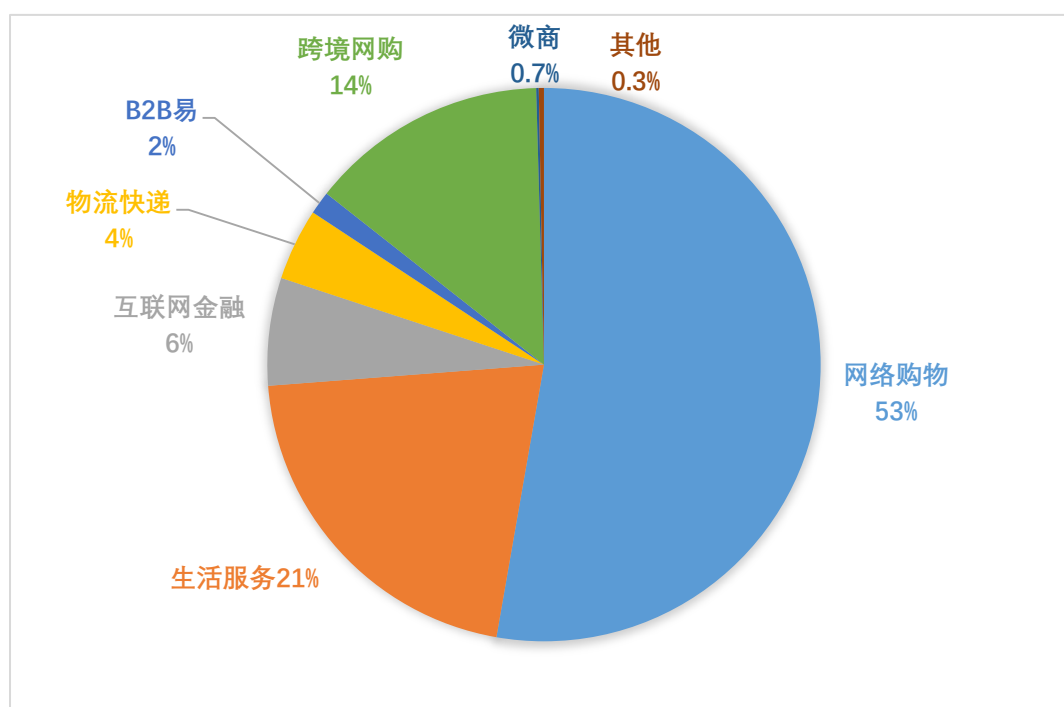
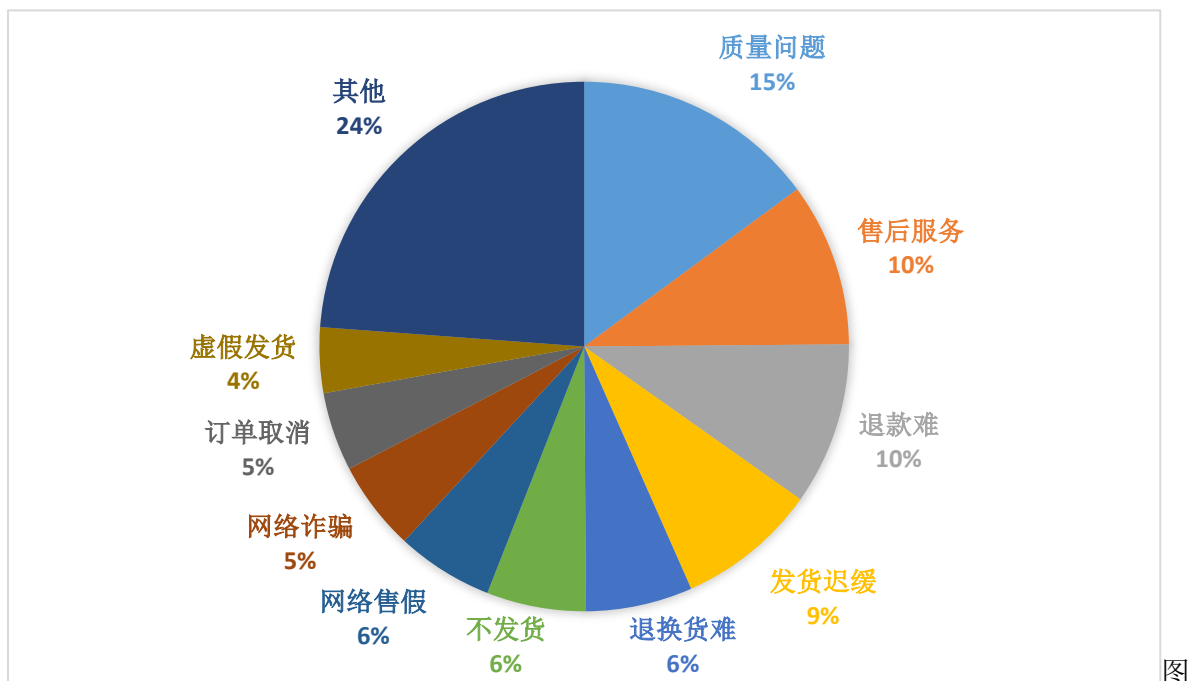


图 1.2 电子商务投诉领域分布

数据来源：中国电子商务用户体验与投诉监测报告[1]

电子商务带给消费者方便且价格优惠，从 2008 年起网上零售的交易规模也日渐攀升(见图 1.4)，但在快速交易的过程中也产生了很多的问题和困扰，而造成消费体验下降。根据中国电子商务投诉与维权公共服务平台监测数据显示，2016 年上半年通过多种不同的投诉管道，全国网络消费用户电子商务投诉数量比去年同期增长 4.16%，以下为投诉问题分布(见图 1.3)：质量问题(14.88%)、售后服务(9.99%)、退款难(9.92%)、发货迟缓(8.59%)、退换货难(6.51%)、不发货(6.07%)、网络售假(5.92%)、网络诈骗(5.48%)、订单取消(4.81%)、虚假发货(4.00%)[1]，成为网络零售十大被投诉问题。质量问题为影响消费者最重要的因素，这些投诉意见忠实地反映消费者对产品的不满，消费者负面情绪影响因素相关调查可以藉此分析，但消费者满意度是非常多维度所构成，并非只有单一维度所能概括，除了网站投诉的数据分析外，在线用户评论借着定量的文本分析方法也能提炼不少价值，透过在线用户评论分析，企业的业务部门、产品设计部门…生产链上的各部门也能够更加了解顾客对产品和服务的真实体验与需求，进而改进产品与服务。

现阶段国内对电子商务研究中，大多数的文献都是『质量』来概括所有的设计属性，并没有真正深入设计制作流程中去讨论各个环节的问题，且藉由分析在线评论影响产品本质（设计）的国内文献更是少之又少。



1.3 网络零售十大热点被投诉问题

数据源：中国电子商务用户体验与投诉监测报告[1]

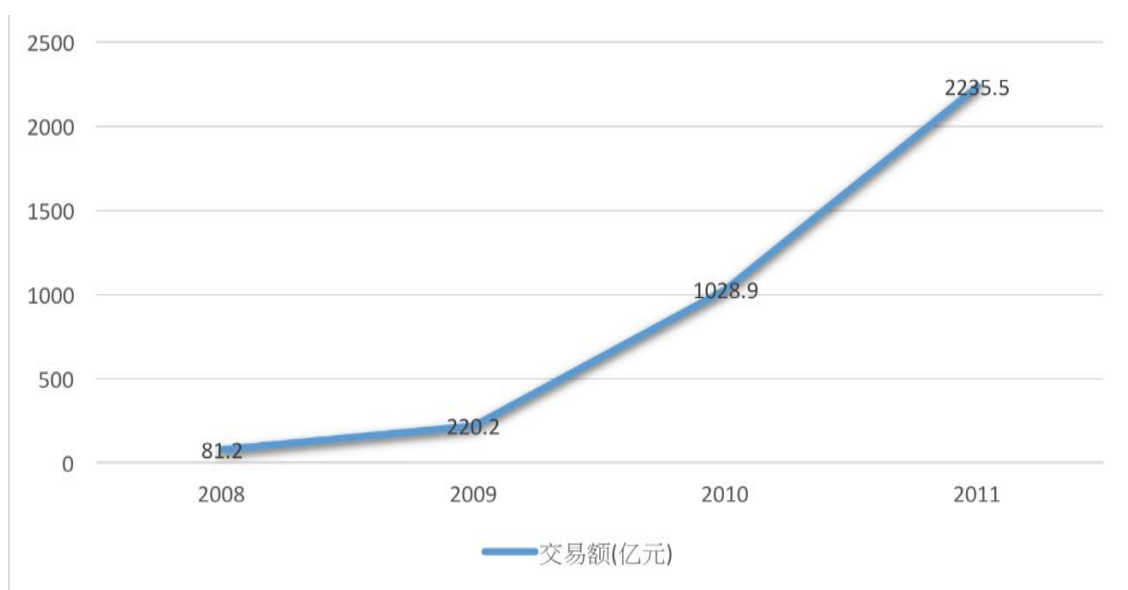


图 1.4 中国网络零售 B2C 交易规模

数据来源:易观智库,中国网上零售市场企业年度盘点 2015.1

1.1.2 国内文化产业政策现况

近两年根据政策面而言,十三五计划中的文化相关政策取向,出现了从”特惠型政策”转型”普惠型政策”[2],也就是说从”局限于文化主管部门的管辖范围”,扩大到”所

有受文化产业带动的相关领域”，根据 2016 中国文化产业报告(见图 1.5)指出国人对文化消费的支出日益提高，文化产业渐渐复苏，因此此类消费者与设计端的研究更应该被重视。

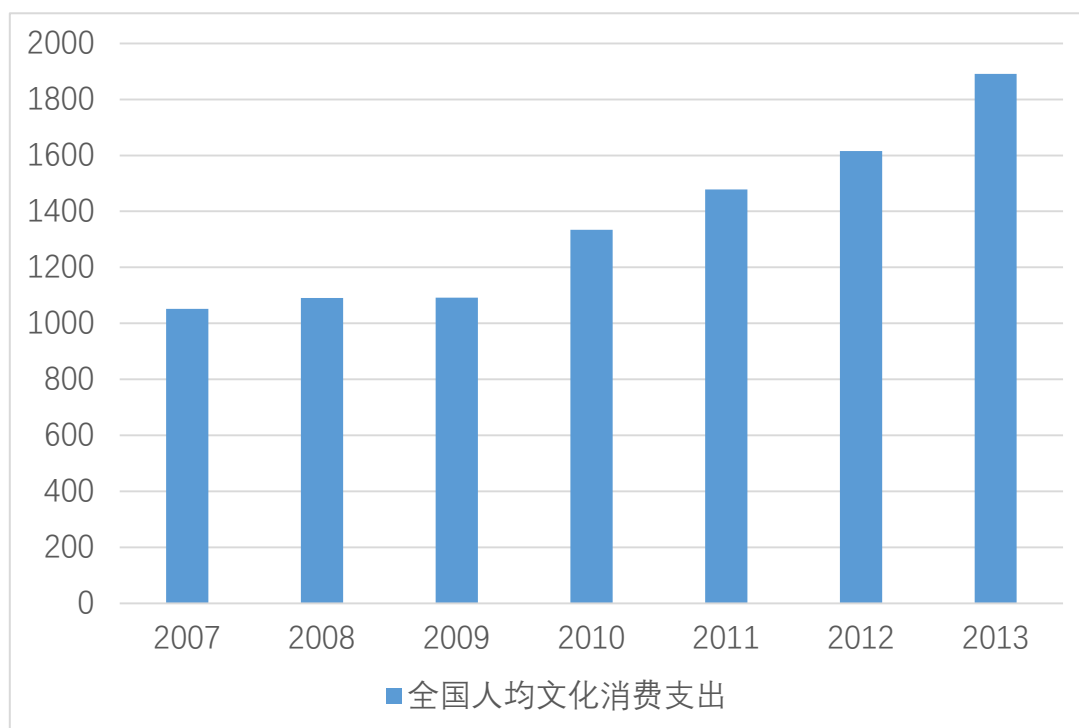


图 1.5 中国人均文化消费

数据源：中国文化产业报告（2015~2016）[2]

当前政策环境上也是重点支持文化产业发展,李克强总理在 2015 年 3 月 5 日上午十二届全国人大三次会议上,在中央工作报告中首次提出“互联网+”行动计划。在十三五计划中,先后出台了一系列扶持、促进实体产业互联网化的发展的相关政策,江西省政府已于近几年将电子商务列入全省战略性新兴产业目录,并正式出台了《江西省人民政府关于加快电子商务产业发展的若干意见》。根据政府政策支持所支持资本市场都相继完善[31],对文化产业有了非常大的支撑作用,国内电子商务在此对于陶瓷产业走上了规模化、专业化,形成了一系列可持续发展的轨道。

景德镇市是一座渊源悠久的特色产业城市,各类陶艺人才、环境物资高度具备,自古以来营造出一个完整的工艺行业体系,这个工艺古城的材料、技术、设计上的品牌的优势之中能够将陶瓷行业纳入电子商务,景德镇优异的资源禀赋使其行业品牌优势明显可见。

互联网+的产业形态发展趋势逐渐上升,各式各样的商品类别与互联网相结合,不同行业别不断将产品营销业务伸向互联网,互联网+并不会改变商业的本质,更会

提升该产业的核心竞争力，透过技术的手段解决信息不对称的问题，重新解构了商业价值链，传统企业不断互联网化，企业商品、服务可被消费者快速地看见，随之而来所要面对的就是产品的升级与迭代，如何利用互联网上的在线评论资源，撷取出对企业有用的金矿，成为大多数企业中后期所要面临的问题。

1.1.3 国内外研究现况

从 2003-2004 年开始，出现了一系列在线评论者的研究工作[6][7]，顾客满意源自于对商品的体验，现今电子商务在线评论中看到消费者的各种情绪，这些文字来自于消费者对产品的使用与期望，这些数据对于企业主而言具有分析价值，如表 1.1 可看到在线评论有各式各样的研究目标，这些学术研究利用挖掘这些文本的过程中，找到产品问题、企业运营的症结。

表 1.1 国内外电子商务顾客满意文献

文献作者	讨论议题
Chun-Chun Lin, Hsueh-Ying Wu, Yong-Fu Chang[8]	调查近 400 名有网络购物经验的台湾大学生，分析在线用户的满意度，利用回归方法验证。其网络用户满意度的影响因素有以下：消费者认知价格，客户服务质量，产质量量，供应链的质量
Srinivasan, Anderson and Ponnarolu [11]	研究(B2C)背景下客户忠诚度的前因和后果。提出了可能影响电子忠诚度的八个因素，从 1,211 个在线客户收集的数据表明，除便利之外，所有这些因素都会影响电子忠诚度。数据还显示，电子忠诚度对口碑宣传和支付意愿有影响。
Kim Hye-Ban [12]	从管理信息系统(MIS)，营销和电子商务三个方面综合运用。利用电商数据研究客户满意度，这项研究定义了电子客户满意度，并使用加权模型开发了一个指数。从管理者和研究人员评估在线满意度该指数使用韩国样本进行测试。

Jumin Lee a., Do-Hyung Park , Ingoo Han [9]	此研究调查负面在线消费者评论对消费者产品态度的影响。这项研究中的实验参与对象为韩国 248 名大学生。提出的三种影响因素,并用协方差分析来测试,从信息处理的角度考察负面在线消费者评论的比例和质量。
周晶晶(浙江大学)[10]	利用在线客户评论来进行在线口碑影响因素和作用机理的研究。探讨在线客户评论对消费者购买决策过程和最终购买决策结果的影响。在线口碑具有非面对面性、异步性、弱连接性、集中性、可存储性等突出特点,其传播速度更快、传播范围更大、影响程度更深。共获取了 294 份有效问卷,使用 SPSS15.0 和 AMOS7.0 两种统计分析工具对实证数据进行处理和分析。
黄梦婷(北京大学)[57]	该论文从内容要素和文本特征上研究了评论特征与有用性之间的关系。选定京东、淘宝、亚马逊上的手机、鞋靴、图书等 9 类商品、功能型与体验型两种产品类型,爬取了 90 件商品的 2299 条评论进行内容分析,确定了现有在线评论涵盖的内容要素,在此基础上,展开问卷调查与用户访谈,进而提出各类指标
王家园, 徐薛艳(上海商学院)[39]	以途家自营公寓为研究对象,三方面把大量途家网真实在线评论量化分析,得出关于途家自营公寓服务质量关注度和满意度水平及全国在合理构建其服务质量评价体系,从各级指标评价情况、各具体指标与总体评价的相关性、评价指标的地区差异性,区域性差异特征的相关结论
张智超(北京大学)[58]	根据在线评论,指出不同类型酒店存在的服务瓶颈,依据服务失败模型给出建议,进而从避免服务失败为酒店的服务管理者,提供优化服务质量,提高顾客满意度。
李 健(上海大学管理学院)[33]	以手机的在线评论为研究主体,研究在线评论内容质量、评论的时效性等因素对产品销量的影响,结果利用交易成本经济学原理,并在此基础上确定各影响因素对产品销量的影响因子大小和销量的逻辑回归模型,为企业的产品销售策略提供参考

郭潇(华南理工大学)[34]	从实证角度出发,分别假设在线评论的内容质量、评论的倾向性和评论的形式等因素,对旅游预订意向具有显著影响。在本次研究中主要调查在线评论对商务人士的在线旅游预订意向的影响。对有效样本数据进行了信度分析、方差分析和回归分析等。在线评论对顾客旅游预订意向影响的理论依据,深入讨论了有关研究结果,对旅游电子商务发展提出了一些建议
皮常玲, 郑向敏(华侨大学)[35]	以厦门鼓浪屿民宿为例,根据携程网 1026 条有关民宿顾客抱怨的在线评论,运用内容分析法对民宿顾客抱怨信息进行分析。
杨铭, 祁巍, 闫相斌, 李一军(哈尔滨工业大学管理学院) [36]	提出对在线商品评论效用分析的最新研究进行评述,研究需要充分关注消费者的购买决策过程,效用评价是基于 CMC(computer mediated communication) 文本模式分析同时为电子商务网站的运营商调整营销沟通策略提供决策支持
严建援, 张丽, 张蕾(南开大学)[37]	从在线评论文本内容的视角出发,在 B2C 电子商务网站的 221 个有效样本的实证分析,利用回归模型分析在线评论内容对评论有用性的影响。
张艳辉、李宗伟、赵诣成(管理学报)[38]	根据淘宝网的海量在线评论数据构建模型,根据评论将文本、追评、上传图片分类。并分析其影响的程度,将产品分类成搜索品与体验品,并将文本分类量化进而分析文本取向,研究比较其有效评论数、中差评、上传图片对评论有用性具有不同的正向影响作用!

资料来源:北京大学学位论文数据库、中国知网 CNKI、Google scholar

根据以上数据概要综述归类后,其局限与不足有以下:

1. 国内外大多数的分析文本数量都在 2500 以内
2. 以上文献都是利用监督学习的方式人为进行归类与分析,且在文本分析时,过度依赖定性分析
3. 没有针对产品的设计制作流程进行任何讨论

1.2 研究目标内容与方法

在现今的电子商务市场中，有效识别顾客需求成了非常重要的议题，也是提升顾客体验的基本条件，用户生成的产品在线评论是一种越来越重要的内容类型，它们提供有价值的信息，并以此了解消费者与设计师在产品设计中的思维差异，帮助产品设计者更好地了解消费者，客户需求是驱动产品设计的关键因素，所有具有生产部门的公司都必需面对的[3]。

在日常生活中，并非所有产品规格皆可量化，且并非所有功能都有对映的硬件支持，本论文所要探讨的领域为难以量化的陶瓷工艺产品，在此面向此类产品提出新的优化设计模型，以无监督学习的方式用定量分析方法来将特征提取，并与专家法相结合，将顾客需求转换为偏好属性，利用量化方法将偏好属性、技术环节集成到 QFD 中作分析。

此新模型可使设计师、工艺师能更加快速的了解消费者对产品的认知，根据分析结果对整体设计流程提出优化方案，进而推出更多适合的商品来满足目标市场、提升消费者的顾客满意。研究内容涉及以下三个方面：

第一部分 分析在线评论中的消费者偏好属性。

第二部分 基于在线评论分析结果，做行业调研分析。

第三部分 基于以上两部分(在线评论分析、行业调研分析)，提出设计优化方案。

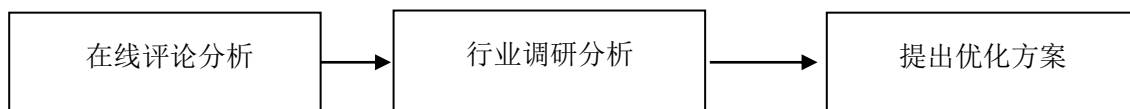


图 1.6 分析流程

1. 在线评论分析。将在线评论利用无监督学习的框架，用定量分析方法将特征提取。
2. 行业调研分析。将在线评论分析结果集成到专家方法中，利用德尔菲法(专家规定程序调查法)做调整，利用 Kano 模型中的需求框架将偏好属性设计问卷，并且与匠人探讨整体的制作流程与工艺难点，结合技术环节属性与消费者偏好属性，在不同的分析流程中了解匠人与消费者彼此之间的认知差异。
3. 提出优化方案。将在线评论分析与行业调研分析的结果集成，向设计端提出建设性的优化方案。此部分利用质量功能展开 (QFD; Quality function deployment) 中的质量屋(HoQ ;House of Quality)作为本架构下重要的分析模型，该模型被广泛用于将顾客需求转换为用于设计，开发和交付产品的技术属性，利用此模型分析消费者偏好属性与技术属性相互关系，进而向设计端提出优化方案。

本论文针对国人接受度最高的茶壶种类：紫砂壶 作为研究主题，以国内电子商务网站淘宝网、天猫网作为在线评论数据撷取的对象，利用 Python 编写爬虫采集了 96 种紫砂壶商品，28358 条文本数据，透过两大电子商务网站搜集的用户在线评论数据，探讨消费者对于该类产品的偏好属性，并且将此结果对比于行业调研分析的相关模型与理论，将之集成为优化方案。

1.3 论文结构与组织

第 1 章：绪论。本论文的研究背景、研究目的和意义，概述国内电子商务、文化产业发展现况，并针对本论文调研地区，概要介绍江西景德镇陶瓷行业发展现况，总结其现有设计发展的趋势，并根据国内外研究发展的现况，提出本论文的研究内容与方法。

第 2 章：相关理论与方法。介绍产品优化设计的调研现况研究，并综述本论文所应用的理论与研究方法。

第 3 章：优化设计方法。针对难以规格化的商品，提出一套新的分析模型，本章将详细论述此新模型的所有分析流程，以供设计师、工艺师排定工序的优先级。

第 4 章：基于用户评论的紫砂壶设计优化案例。利用提出的新模型，运用在现有的实例之中，并根据分析结果提出产品优化方案。

第 5 章：结论。

第二章 相关理论与方法

2.1 产品优化设计调研方法概述

现今国内外的文献对于在线客户的影响因素的相关文献，研究方向大多提出用户满意度的因素和各种消费者属性，对于产品设计流程的细化讨论是相对少的。在设计领域中只有少数研究关注在线评论的价值，并利用在线评论总结和分析产品设计的在线评论的主题结构[14] 传统设计人员获得大量调查数据的方式是非常耗时且耗费劳动力的。如今各大电商网站有大量的客户评论可以运用，有别于传统调查数据具有特定框架下去设计问卷，访问对象在实验设计或中选择出最适当的答案。

在线评论与传统调查问卷所收集的数据非常不同，这些评论完全是基于消费者的意愿，从他们的思维角度出发，他们的语句和评论都没有任何规定来引导，这个特点有别于调查数据包含很少个人意见，且回答的框架在实验之初就已被限制，而此种文本表达方式对于在线评论是至关重要，在线评论包含大量具有强、弱情绪的句子。这种情绪会激发设计师发现潜在的产品属性，这些情感表达的文本中包含客户提供所提供的情绪、需求、体验等有价值的输入，这样的模式有助于协助设计者在设计新产品或改进当前模型，但此类文本也又局限性，其局限为客户在线评论是自由文本。文字的多寡无法去限定，评论文本大小可差距数百至数千个字，并不是所有这些数据都包含了对产品设计师非常有价值的信息，过去的研究中显示，只有少数句子可能包含其特性[17]，因此设计师在概念设计的过程中能够多深刻的理解、正确的接收到客户的需求成为产品设计中的关键因素。

传统调研开发了各种模型与在线评论和调查数据之间的文本差异，两者数据采集方式、数据形式如此的不同，使得许多过去的常规方法在产品面临概念设计流程中不能在大量的被应用。新产品成功的可能性很大程度上取决于顾客满意度水平，但正向的在线评论不一定能为企业带来新的消费者，负面评论可能导致潜在客户的损失。因此，理解客户在开发新产品发挥着重要作用，对顾客需求有效地进行分析并立即做出有效响应，在竞争激烈的市场变得至关重要，从产品设计师的角度，QFD 评估和预测在线评论的有效性的关键步骤之一。因此，本项研究基于分析客户在线评论，有效的提取高质量的数据信息，并运用定量分析、专家法使设计师快速了解偏好属性，提出该产品的优化设计方案。

2.2 无监督学习方法

2.2.1 中文分词

本研究利用 Jieba 模块作为中文分词的方法, 该模块所应用的原理根据以下: Jieba 分词自带了一个词典, 叫做 dict.txt(此词典是中科院、清华大学等科研机构基于 2014 年人民日报语料资源训练, 词性为国标汉语词性), 词典中有 2 万以上的条词, 包含了词条出现的频率与词性, 将这 2 万多条词语, 生成至 trie 树中, 生成 DAG 有向无环图用于消除分词的歧异、提高切分准确度, 且利用动态规划查找最大概率路径, 找出基于词频的最大切分组合[53]。对于未登录的词采用隐马尔可夫链模型(Hidden Markov Model, HMM)和 viterbi 算法, 在隐马尔可夫链模型中分成 BEMS 四种状态, B(begin)开始位置, E(end)是结束位置 M (middle)中间位置, S (single)单独成词, 利用 viterbi 算法来得到这个最佳的 BEMS 隐藏状态序列。

一、DAG (Directed Acyclic Graph)有向无环图

有向无环图为有向图的一种, 被用来表示事件之间的驱动依赖关系, 管理任务之间的调度。常常根据其顶点用于拓扑排序(Topological Order)[54]

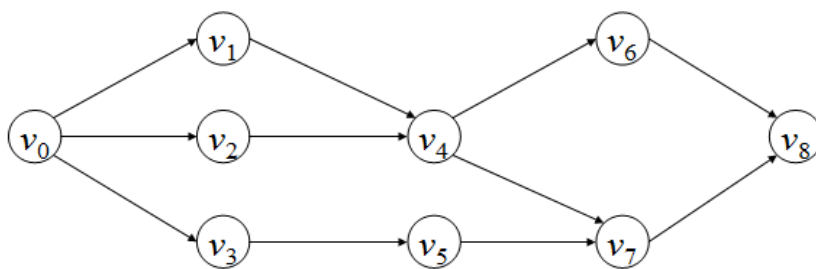


图 2.1 有向无环图

二、Trie 树

Trie 树, 又称键树、单词查找树, 为树形结构, 为哈希树的变型, 用于快速检索的多叉树结构。其的优点是: 核心思想利用空间换时间, 运用字符串的共前缀来减低查询时间花费, 最大限度地减少无谓的字符串的运算以提高效率。因为运用空间换取时间, 消耗内存非常大, 它的优点是: 最大限度地减少无谓的字符串比较, 查询效率比哈希表高, Trie 树有三个特性:

1. 每一个树状节点都只包含一个字符, 除了根节点不包含外。
2. 从根节点到一节点, 所有经过字符连接起来, 为该节点的字符串。
3. 所有节点字符都不相同。

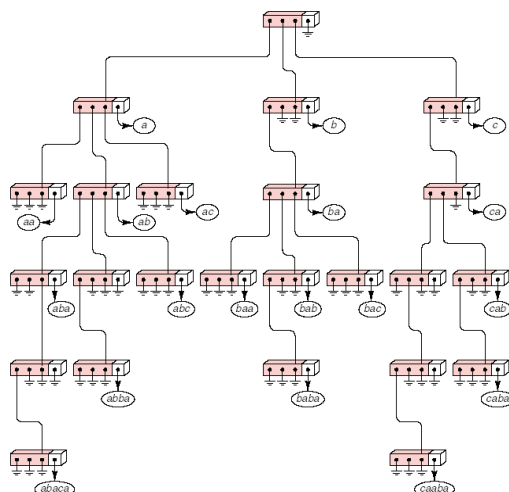


图 2.2 Trie 树状结构

三、 隐马尔可夫链模型(Hidden Markov Model, HMM) 与 viterbi 算法

隐马尔可夫模型是马尔可夫链的变形，它为一个抽象的概念无法直接观察，但能经过向量序列进行了解，每个向量都是透过一个相应的概率密度分布来呈现其状态，隐马尔可夫模型为双重随机过程，80 年代末，HMM 被应用于语音识别，取得重大成功。

维特比算法由安德鲁·维特比(Andrew Viterbi)于 1967 年提出，此算法为最广泛被应用的动态规划算法，用于求解 HMM 上的最大概率的算法。而该算法是针对篱笆网络的有向图 (Lattice) 的最短路径问题所提出[55]维特比算法利用动态规划算法用于寻找最有可能产生事件。凡是使用隐含尔可夫模型描述的问题都可以用它来解码，包括今天的数字通信、语音识别、机器翻译、拼音转汉字、分词等。[55]

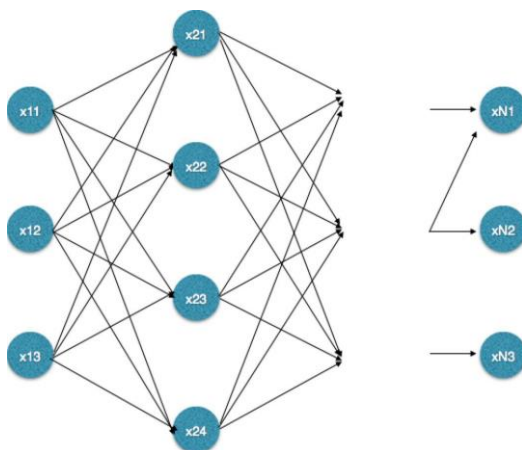


图 2.3 篱笆网络

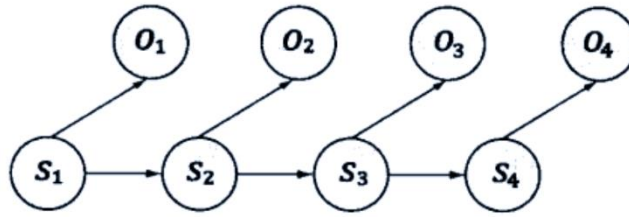


图 2.4 隐马尔可夫链模型

2.2.2 TF-IDF

TF-IDF (Term Frequency - Inverse Document Frequency) 是一种基于信息检索与文本挖掘的常用加权技术。TF-IDF 为一种统计方法，作为评估字词对于文件集中某份文件的重要程度。字词的重要性随着它在文件中出现的次数成正比增加[40]。一个词在特定的文档中出现的频率越高,说明它在区分该文档内容属性方面的能力越强(TF) ; 一个词在文档中出现的范围越广,说明它区分文档内容的属性越低(IDF)[41], TF-IDF 在各种文本分析、搜索引擎被应用在搜索引擎中，作为相关程度的度量或评级。其公式如下:

词频(Term Frequency: TF):

$$TF_{ij} = \frac{n_{ij}}{\sum_k n_{kj}} \quad \text{式(2.1)}$$

该词 $n_{i,j}$ 在文件 d_j 中的出现次数，而分母则是在文件 d_j 中所有字词的的出现次数之和。

逆向文件频率 (inverse document frequency, IDF) :

$$IDF_i = \log \frac{|D|}{1 + |\{j: t_i \in d_j\}|} \quad \text{式(2.2)}$$

$|D|$: 语料库中的文件总数; 若该词语 t_i 在文件中没有出现就会导致分母为零，一般情况下以此 $1 + |\{j: t_i \in d_j\}|$ 表示包含词语 t_i 的文件数目[40]

$$TF - IDF_{ij} = TF_{ij} \cdot IDF_i = \frac{n_{ij}}{\sum_k n_{kj}} \cdot \log \frac{|D|}{1 + |\{j: t_i \in d_j\}|} \quad \text{式(2.3)}$$

2.2.3 PCA 与 KPCA

一、主成分分析 (Principal Component Analysis, PCA)

主成分分析是最常见的降维方法，是一种简化数据集线性变换技术。将数据变换到一个新的坐标系统中，使得任何数据投影的第一大方差在第一个坐标(称为第一主成分)上，第二大方差在第二个坐标(第二主成分)上...。主成分分析经常被应用在减少数据集的维数，同时保持数据集的方差贡献最大的特征[42]。

PCA 具有以下两种特性[43]：最大重构性：样本点到超平面的距离都足够近。最大可分性：样本点在这个超平面的点都尽可能分开。主成分分析有以下的作用[42]：

1. 主成分分析能降低所研究的数据空间的维数。
2. 多维数据的一种图形表示方法。根据主成分的得分，画出 n 个样品在二维平面上的分布状况，由此数据可视化可直观地了解各样品在主分量中的权重，可以由图形发现远离大多数样本点的离群点，但当数据维数大于三时，无法将该研究的问题用图形表示。
3. 由主成分分析法构造回归模型。即把各主成分作为新自变量代替原来自变量 x 做回归分析。

二、核主成分分析 (Kernel Principal Component Analysis, KPCA)

核主成分分析，为一种非线性降维，线性降维的方法是假定是从高维空间到低维空间的函数映像是线性的，但在现实生活中某些数据是需要非线性映像才能找到低维嵌入[43] KPCA 内核方法可以应用于任何可以根据点产品专门制定的算法，包括例如 k-mans 和独立成分分析等[45]。

KPCA 的核可以根据产品专门制定算法，包括例如 k-mans 和独立成分分析[45]，KPCA 面对非线性主成分识别比线性主成分识别的提供更好的识别率，通过在线性情况下使用，可以进一步提高非线性组件的性能[45]，KPCA 不同于 PCA 为了以较佳的方法处理非线性数据，引入隐性的非线性映射函数 Φ ，将原空间中的数据映射到高维空间，以下推导进行演示 PCA 与 KPCA[43]：

从 PCA 的最近重构性推导，假定数据进行了中心化，即： $\sum_i x_i = 0$ 投影后的坐标为 $\{w_1, w_2, w_3, \dots, w_d\}$ ， w_i 为标准正交基向量 $\|w_i\| = 1$ ， $w_i^T w_j = 0$ ($i \neq j$)，将维度降低到 $d' < d$ 则样本点 x_i 在低维的投影是 $z_i = (z_1, z_2, z_3, \dots, z_{d'})$ 其中 $z_{ij} = w_j^T x_i$ 是 x_i 在低维坐标下第 j 维的坐标，基于 z_i 重构 x_i 则会得到 $x_i' = \sum_{j=1}^{d'} z_{ij} w_j$ ，其原样本点 x_i 与投影样本点 x_i' 之间的距离：

$$\sum_{i=1}^m \left\| \sum_{j=1}^{d'} z_{ij} w_j - x_i \right\|_2^2 = \sum_{i=1}^m z_i^T z_i - 2 \sum_{i=1}^m z_i^T W^T x_i + \text{const} - \text{tr} \left(W^T \left(\sum_{i=1}^m x_i x_i^T \right) W \right) \quad \text{式(2.4)}$$

根据最近重构性(上式)应该被最小化，考虑到 w_j 是标准正交基 $\sum_i x_i x_i^T$ 为协方差矩阵，这为从最近重构性推导的优化目标

$$\min_w -tr(W^T X X^T W) \quad \text{式}$$

(2.5)

$$\text{s.t. } W^T W = I \quad \text{式(2.6)}$$

从最大可分性推导，样本点 x_i ，在新空间中超平面的投影 $W^T x$ ，投影后样本点的方差 $\sum_i W x_i x_i^T W$ 优化目标为

$$\max_w tr(W^T X X^T W) \quad \text{式(2.7)}$$

$$\text{s.t. } W^T W = I \quad \text{式(2.8)}$$

其式(2.5)、式(2.7)等价，使用拉格朗日乘子法相乘可得

$$X X^T W = \lambda W \quad \text{式(2.9)}$$

KPCA：将高维空间数据投影到由 W 超平面

$$(\sum_{i=1}^m z_i z_i^T) W = \lambda W \quad \text{式}$$

(2.10)

z_i 为 x_i 在高维空间的像，推导出：

$$W = \frac{1}{\lambda} (\sum_{i=1}^m z_i z_i^T) W = \sum_{i=1}^m z_i \frac{z_i^T W}{\lambda} = \sum_{i=1}^m z_i \alpha_i \quad \text{式(2.11)}$$

其中 $\alpha = \frac{1}{\lambda} z_i^T W$ ， z_i 由原始空间 x_i 透过映射 ϕ 产生的 $z_i = \phi(x_i)$ ，将之映像到高维空间后再实施 PCA 即可

$$(\sum_{i=1}^m \phi(x_i) \phi(x_i)^T) W = \lambda W \quad \text{式}$$

(2.12)

$$= W = \sum_{i=1}^m \phi(x_i) \alpha_i \quad \text{式}$$

(2.13)

由于无法判定 ϕ 的具体形式，因此引入核函数

$$\kappa(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad \text{式}$$

(2.14)

整合上式后可得

$$K A = \lambda A \quad \text{式}$$

(2.15)

K 为对应 κ 的核矩阵 $(K)_{ij} = \kappa(x_i, x_j)$ ； $A = (\alpha_1; \alpha_2; \dots \alpha_m)$ 取 K 最大的 d' 个特征值所对应的特征向量，对新样本 x 投影后第 $j(j = 1, 2, 3 \dots d')$ 维坐标：

$$z_j = w_j^T \phi(x) = \sum_{i=1}^m \alpha_i^j \phi(x_i)^T \phi(x) = \sum_{i=1}^m \alpha_i^j \kappa(x_i, x) \quad \text{式}$$

(2.16)

经过规范化 α_i^j 为 α_i 的第 j 个分量 经过以上推导与参考文献所指出[45], KPAC 需对样本求和, 其计算成本相对较高。

2.2.4 K-means 聚类

K-means 算法也被称作 K-平均、K-均值、Lloyd's 算法, 此算法一开始作为讯号处理的向量量化方法, 术语「k-means」于 1967 年才被 James MacQueen 首次使用[44], 现在多被数据探勘领域作为一种聚类方法, K-means 算法是很典型的基于距离的聚类算法, 采用欧氏距离作为相似性的评价指标, 两个对象的距离越近, 其相似度就越大。但欧式距离的缺点为距离与各变量的量纲有关, 其变差大的变量在距离中的作用就会变大[56], 消除这方差所带来的影响可以将数据归一化、将各变量加权, 两个 n 维向量间的欧氏距离:

$$d_{12} = \sqrt{\sum_{k=1}^n (x_{1k} - x_{2k})^2} \quad \text{式(2.17)}$$

距离靠近的点组成簇组成的, 因此把得到紧凑且独立的簇作为最终目标。此算法为无监督学习(unsupervised classification)架构下的经典的原型聚类方法(prototype-based clustering)。此类算法假定聚类结构能透过一组原型刻划[43]。

K-means 第一步中是随机的选取任意 k 个对象作为初始聚类的中心, 初始地代表一个簇。从 N 个数值中随机选取 K 个文档作为质心, 剩余的每个数值测量其到每个质心的距离, 并把它归到最近的质心的类, 不断迭代至新的质心与原质心相等或是小于指定阈值, K-means 聚类: 给定样本集: $D = (x_1, x_2, x_3, \dots, x_m)$, 针对聚类所得到的簇划分 $C = (C_1, C_2, C_3, \dots, C_m)$ 点的最小平方误差

$$E = \sum_{i=1}^k \sum_{x \in C_i} \|x - u_i\|_2^2 \quad \text{式(2.18)}$$

其中 $u_i = \frac{1}{|C_i|} \sum_{x \in C_i} x$ 为簇 C_i 的均值向量, 上式描述了样本的紧密程度, E 值越小, 其样本相似度越高, 使得每个点都属于离他最近的均值为簇, 以之作为聚类的标准, k-man 的主要特征欧几里德距离、方差作为度量。簇 k 值为输入的参数, 当 k 的不适当时可能产生差的结果, 此二类特性使该算法的效率降低, 但此算法有部分参考书目提出改善方案[46]。

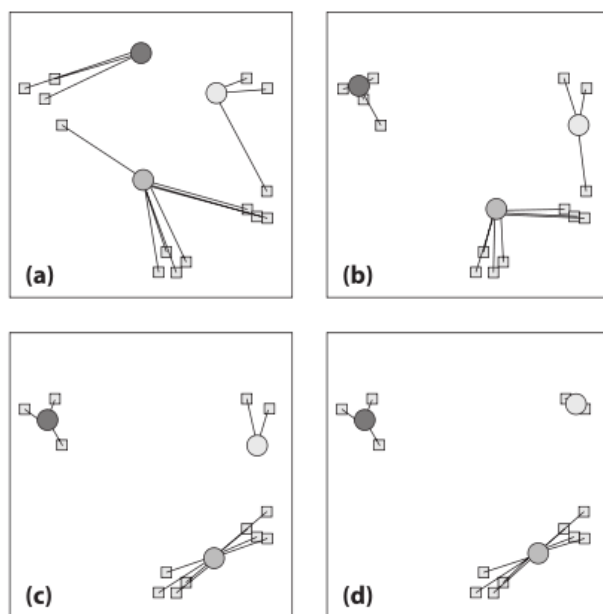


图 2.5 k-means 聚类顺序

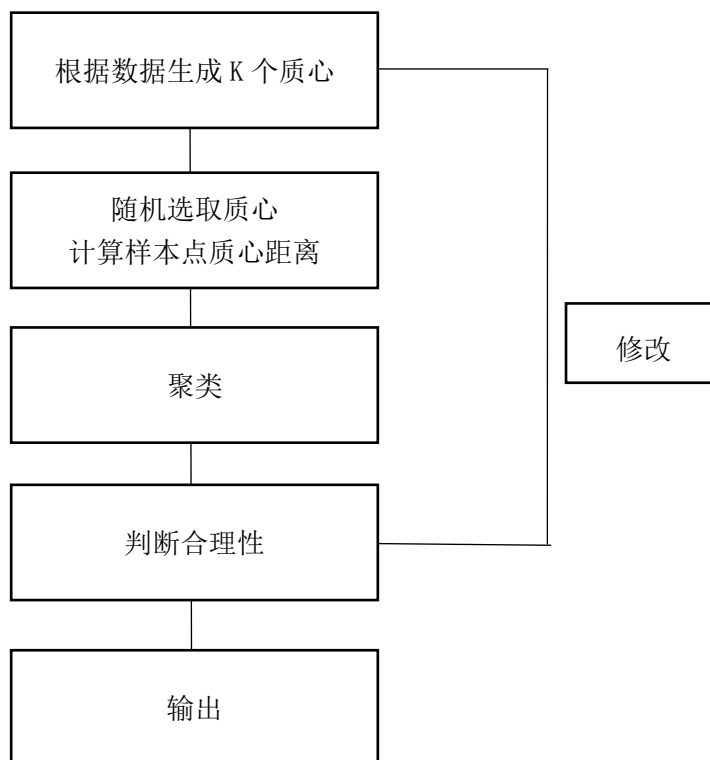


图 2.6 k-means 聚类流程

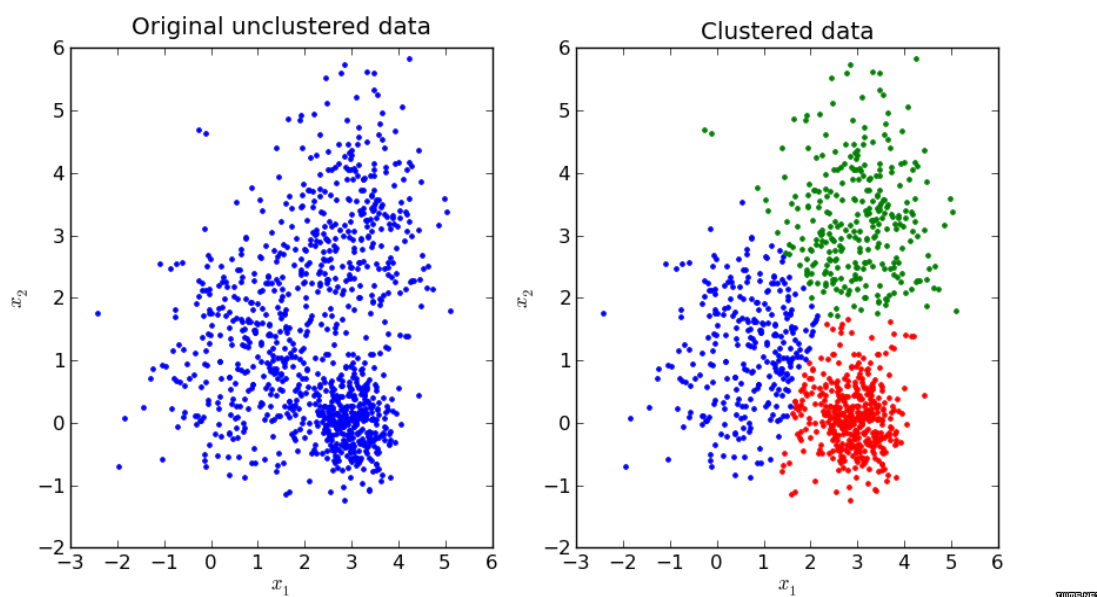


图 2.7 k-means 聚类效果图

2.3 调研与专家法

2.3.1 德尔菲法

德尔菲法 (Delphi method)，又称专家意见法，采用个别征询专家的模式，经过回答问题、反馈、归纳统计几次循环，使专家意见最后趋近集中，本方法为全匿名的方式进行，专家成员之间不能相互讨论、联系，所有对问题的了解都是根据调查人员、前一次的归类完的结果，流程图可见图 2.8 此方法可以规避团队意见领袖的影响和羊群效应的产生，德尔菲法有以下特性：

1. 匿名性

匿名是德尔菲法的极其重要的特点，所有专家组成员不能做横向的联系，只能透过联系人员交流，这样可以将每一个专家意见的权重都一致，使其不会被彼此权威影响。

2. 反馈性

此方法需经过多轮的信息统计与反馈，在每次反馈中调查人员都可以与专家进行探讨与研究，使每一次归纳的统计结果可以客观且充分反映专家看法。

3. 统计性

小组讨论的结果大多数是反映多数人的观点，这也是弊病之一，德尔菲法可以更客观地去统计这些意见，避免了专家会议法只反映多数人观点的缺点。

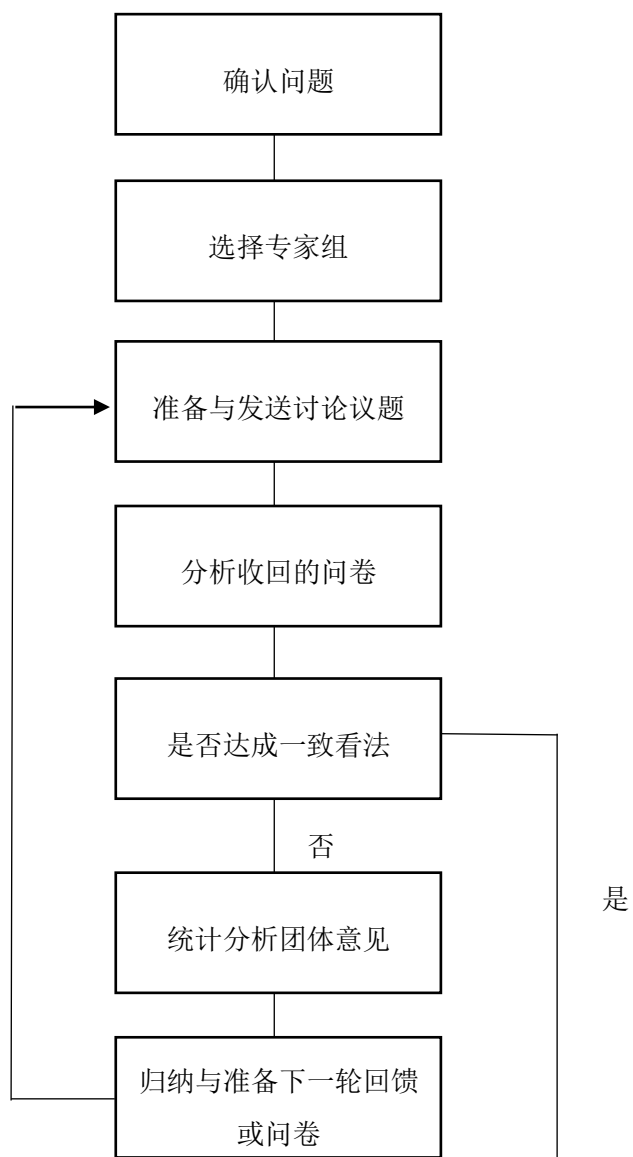


图 2.8 德尔菲法流程图

2.3.2 Kano 模型

Kano 是受行为科学家赫兹伯格的双因素理论启发，东京理工大学教授狩野纪昭 (Noriaki Kano) 在首次提出满意度的二维模式，构建出 KANO 模型。当时日本提高产品和服务一直都是难题。他提出的这个模型有效的解决了这个问题。[18] Kano 模型用于理解消费者需求和消费者满意度的影响[19]一般的满意度应用中，满意度假定为一维，即某个产品，提供更多功能，用户就会感到满意，当功能不充足时，用户会感到不满。因此添加新功能可能成为提升用户的满意度途径。但此种满意度评定方式可能会损害用户体验，满意度理论研究发现，并非所有的因素对用户满意度产生的影响都是一维的影响，二维模式指出当提供某些功能、服务时，未必会获得用户的满意，有时可能会造成不满意，有时提供的服务用户认为根本无差异，将所有需求做分类。Kano 模型能以消费者满意度区分五种类型的消费者需求。

1. 基本(必备)型需求—Must-be Quality/Basic Quality: 必备型需求的实现被消费者认为是理所应当的，因此该需求若是多提供也不会提高太多的满意度。然而，如果企业端不能满足此需求，客户将表现出非常高的不满意度。
2. 期望（一元）型需求—One-dimensional Quality/ Performance Quality: 期望型需求与客户满意度呈现正相关和线性相关，若此需求表现得不好，消费者也不会有太大的满意度下降，一般而言消费者对此需求也是模糊的，所以对于此类需求是企业可着眼的，是提升消费者体验很好的切入点。
3. 魅力（兴奋）型需求—Attractive Quality/ Excitement Quality: 魅力型需求是不会被客户过度期待的需求，一旦提供则可以提升非常高的满意度。若是缺乏这些此项也不造成满意度下降。
4. 无差异型需求—Indifferent Quality/Neutral Quality: 无差异型需求是不论提供与否对客户皆无影响。

5. 反向（逆向）型需求—Reverse Quality: 指引起强烈不满的和导致低水平满意值，提供后用户满意度反而会下降，而且提供的程度与用户满意程度成反比。 [20]

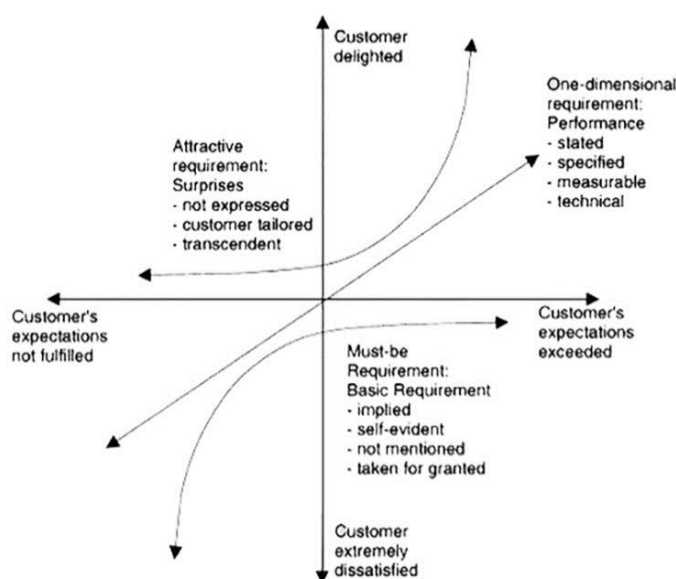


图 2.9 kano 模型

近年来有许多研究试图以不同的方式改进 Kano 的模型。通过考虑客户所定义的产品属性之重要性将 Kano 的模型扩展到八种类别[21]，另一方面，使用'Better'和'Worse'的模型将顾客满意度和顾客不满意程度，集成到卡诺模型做定量分析，以表示一个顾客需求对顾客满意度或不满意度的平均影响 [22]然后由其推断一个顾客需求有多强有力地影响满意度，Kano 模型中的关键问题，需要更多的定量研究才能广泛应用在各个领域之中。

2.3.2 QFD 质量功能展开

质量功能展开是由日本质量管理大师赤尾洋二(Yoji Akao)和水野滋(Shigeru Mizuno)提出，在所有产品和服务流程中，时刻确保产品设计满足顾客需求和价值。此方法现已成为一种重要的质量设计技术，被业界实证可以增进企业各种流程效益、满足顾客要求，为企业在消费竞争市场中能更加贴近消费者，以赢得市场。

质量功能展开(QFD)是一种非常系统性的决策技术，在每一个环节中皆可以将顾客需求转换成偏好属性，在概念设计阶段，此模型可将顾客需求转换成产品定义；在产品生产阶段，可将顾客要求的产品定义转换为整体工艺的制造过程。质量功能展开技术可以在整个产品制造的过程中，顾客的要求不会被曲解，也可以避免产品功能冗余出现很多不必要的功能与流程，使产品的整体设计、制造、所有使用过程中的维

修和运行消耗都可以有效地减少各种成本。以上这些特点，使得质量功能展开逐渐被世界各国所重视，让制造者耗费最短的时间、成本生产出功能上满足顾客要求的高质量产品。

传统生产品质控制是通过用测试与观察的方法达成，这种流程通常也被归于检验产品品质的方法。但质量功能开展则帮助产品改善设计[48] 所有设计阶段，QFD 在产品概念设计时就引入许多不同的概念与要素，使质量融入生产和服务及其工程设计之中。QFD 主要的优点为减少设计变动、设计和制造成本、提高产品质量、顾客满意度，传统设计流程集中于工程技术性能而较少关注顾客需求，QFD 的优点可寻求顾客明确告知的需求，也能发掘潜在的顾客需求，使得无形需求能和公司的战略提供改进方案，进而使得公司进行优先考虑，使其企业增强市场优势，并为消费者带来更好的质量，QFD 架构下的典型质量管理工具介绍：

1. 亲和图（Affinity Diagrams）。又称 KJ 法。KJ 法的应用范围很广，能迅速掌握未知领域的实际情况，找出解决问题的途径，提出明确的方针和见解。用以征别亲和图的缺失和遗漏。
2. 关系图（Relations Diagrams）。此方法用以发现优先需求、造成产品质量流程问题的根本原因。影响质量的因素之间存在着大量的因果关系，这些因果有纵横向关系。关联图法是根据事物之间横向因果逻辑关系找出主要问题的最合适的方法。[48]

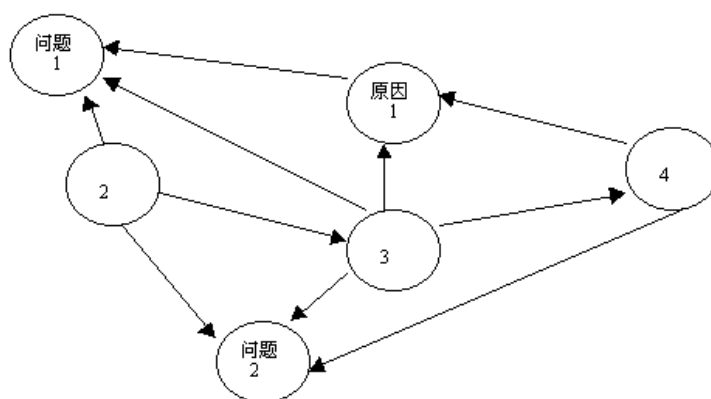


图 2.10 关系图

3. 流程决策程序图（Process Decision Program Diagrams, PDPC）。又称过程决策程序图法、重大事故预测图法，用以分析可能造成新产品或服务失败的潜在因素。在制定设计计划阶段时，事先预测可能发生的障碍，藉此模型的框架设计出一系列对策措施以引向最终目标。由于一些突发性的原因，可能会导致工作出现障碍该法可用于防止重大事故的发生。

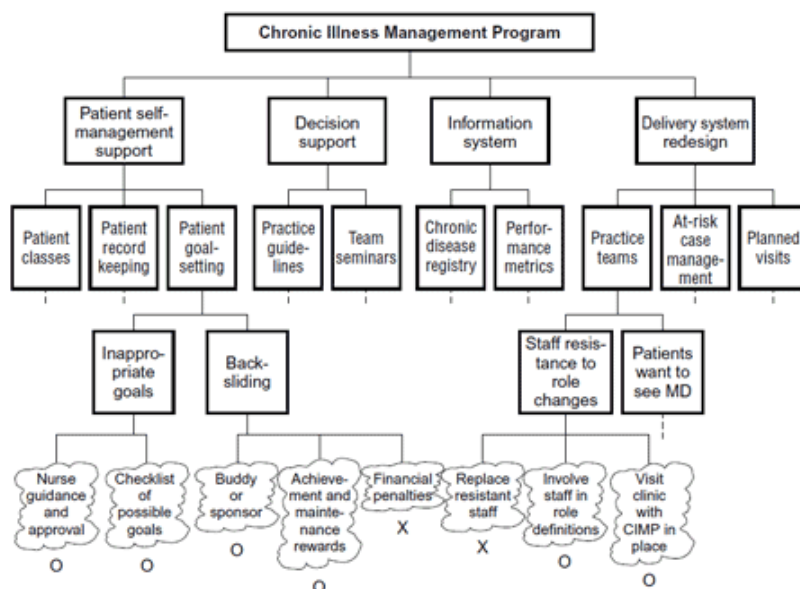
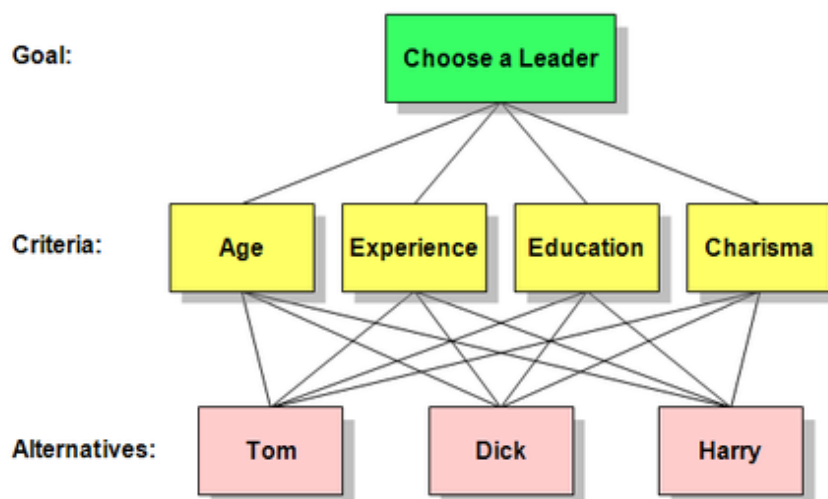


图 2.11 流程决策程序图

4. 层级分析法 (Analytic Hierarchy Process)。在 20 世纪中期由托马斯·塞 (T.L.Satty) 正式提出, 此分析方法结合定量和定性分析, 对一系列的顾客需求进行优先排列, 它在处理问题复杂度的决策上, 非常具有实用性, 由它此应用已普遍至经济计划和管理学界等领域。层次概念的提出, 使得衡量指标变得清晰, 层次结构模型最上层



为目标层, 整体架构将各个因素按照不同属性自上而下分类, 除了底层其每一层同时又支配下一层的因素或受到下层的影响。

图 2.12 层级分析法结构图

5. 质量屋 (the House of Quality) [30]。质量屋为新产品开发中在产品属性与用户需求中扮演很重要的连接角色, 连接用户需求与产品属性的经典工具。透过不同部门之间的相互协作(如市场人员与设计人员), 确立不同需求的可行性方案, 此模块在企业里非常实用, 结合了很多不同部门、不同类型的领域的专业知识, 利用此也排出很多技术无法达到的目标, 完整的质量屋模块, 包括技术各权重之间的关系、竞争对手表现、技术重要性等信息。质量物在产品开发阶段被用来连接产品属性与用户需求。质量功能配置分为七个部分:
- (1) 产品特性 (product features)。
 - (2) 顾客需求 (customer requirements)。
 - (3) 顾客需求重要度 (importance of customer requirements)。
 - (4) 顾客需求与产品特性之间的关系: 顾客偏好属性与产品技术属性的相互关系
 - (5) 产品特性之间的关系
 - (6) 计划矩阵 (planning matrix)
 - (7) 分析结果

QFD 是非常结构化的矩阵驱动 (matrix-driven) 模型其包括 4 个阶段:

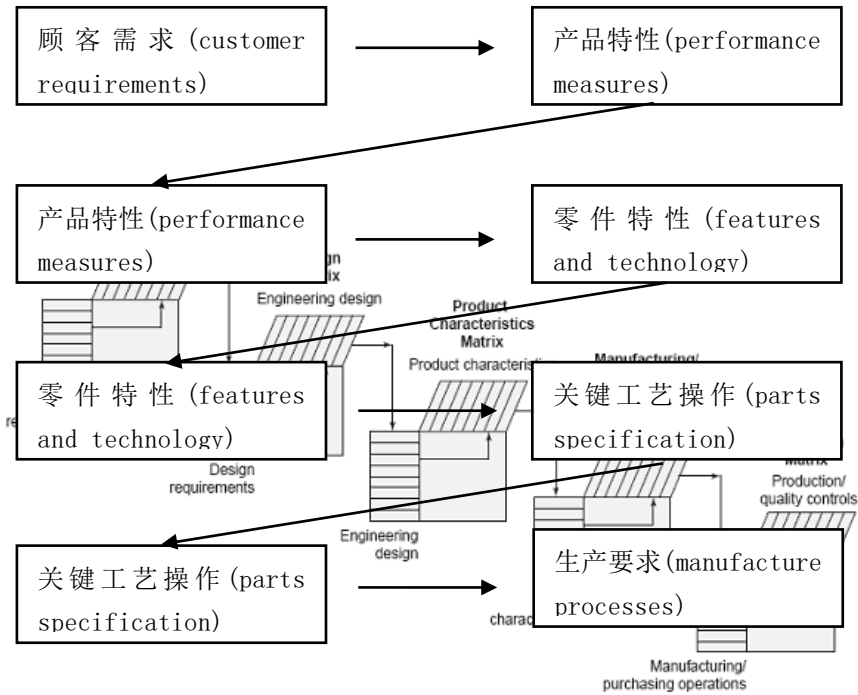


图 2.14 QFD 四阶段流程图

Computer Server Product Planning Matrix																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
Interactions: + Strong Negative + Moderate Negative • Strong Positive ○ Moderate Negative																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
Goal	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O

图 2.15 质量屋的结构[30]

2.4. Kano 模型集成到 QFD

基于 Kano 模型提出一种新的方法,通过计算顾客满意度(CS)和顾客不满意度(DS)来评估其重要性[22],并将 Kano 模型集成到 QFD,用于分析顾客需求[25],不同类型的顾客需求对满意度的贡献可能是多样化的。此外,在更大程度上实现客户期望不一定导致更高的满意度,将 Kano 模型集成到 QFD 的规划矩阵中。进一步改进了旧有的研究方法,包括评分方法和统计检验, Kano 的模型和 QFD 的集成类似的模型常被应用于产品生命周期管理[27]。

本方法的目的是通过考虑各种产品特性来提高消费者满意。研究的重点是如何将 Kano 的模型集成到一个健全的设计方法中,并通过计算顾客满意度和顾客不满意度值来评估其战略重要性[28]。在以上引用文献中, Kano 的模型可用于确定各种产品标准的优先级。过去的实验证明集成的可能性和效用,此方法仍然是 QFD 和 Kano 的模型的定性描述,利用少量的定量分析。因为在问卷中选择不同 Kano 类别的参数基本上取决于 QFD 从业者的经验和知识[49]目的是通过考虑多个标准特性来提高满意度和产品设计质量。

2.4.1 整合方法

QFD 矩阵,为了更好地表示 CR 和 EC 之间的关系,质量屋将中的关系矩阵归一化。因此,在本研究中参考了[50]方法,如下式

$$R_{ij}^{norm} = \frac{\sum_{k=1}^n R_{ik} \gamma_{kj}}{\sum_{j=1}^n \sum_{k=1}^n R_{ik} \gamma_{jk}} \quad i = 1, 2, 3, \dots, m \quad j = 1, 2, 3, \dots, n \quad \text{式(2.19)}$$

R_{ik} 和 γ_{jk} 为质量屋矩阵中的元素

大多数 QFD 优化模型中的一个缺点认为技术属性值被认为落入连续范围内,并且该范围中的任何值可以是可能的解决方案[51][52],然而某些技术属性只能采用离散值。在此情况下所获得的最优解可能不适用于产品开发和制造操作。然而连续和离散技术属性都应该包括在优化模型中的情况,为解决此问题,引用标准化方法将具体的技术值转移到模型中。连续技术属性可分为两组:正面和负面。正面组,技术属性的分数与其技术价值成正比,技术属性的负面则相反。可以通过式(2.20),即电池来归一化技术属性的正面组,负向组可以通过式(2.21)来归一化。

正向组由下式归一化

$$x_j = (X_j^* - X_{\min}) / (X_{\max} - X_{\min}) \quad \text{式(2.20)}$$

负向组由下式归一化

$$x_j = (X_{\max} - X_j^*) / (X_{\max} - X_{\min}) \quad \text{式(2.21)}$$

其中 X_j^* 表示的具体技术值, X_{\max} , X_{\min} 分别表示 EC 的最大和最小可行技术值。离散技术属性归一化通过添加二进制变量 x_{jk} 来处理, 对于每个离散技术属性, 存在用于选择的一组可行的离散值, 每个具有满足等级 d_{kj} , 如果为产品设计选择 EC_j 的值为 k , 则 $x_{jk} = 1$ 。因此, 满足水平等于如式(2.27)中所示的技术属性值的满足等级。然后将离散技术属性值映像到满意度以构建数学编程模型。

$$x_j = \sum_{k=1}^P x_{jk} d_{kj} \quad \text{式(2.22)}$$

QFD 优化模型的主要关注点是集成 Kano 模型的定量结果。顾客满意度和顾客需求的三种类型的关系已经通过 Kano 模型中的 S-CR 关系函数来量化。QFD 优化模型的共同目标是使用有限的资源来最大化顾客满意度。因此一种合适的集成方式是使用 Kano 的模型来建立 QFD 优化模型中的目标函数, 如下式所示, 对于将被设计的乘积 $S(y_1, y_2, \dots, y_m)$ 的总满意度的程度是对于各个顾客需求 s_i 的满意度的程度的加权和。

$$S(y_1, y_2, \dots, y_m) = \sum_{i=1}^m w_i s_i \quad \text{式(2.28)}$$

单个顾客需求的满意程度可以通过表 2.1 中所示的不同类型 CR 的 Kano 模型获得。通过式(2.23)中的积分方程, 在式(2.24)中建立目标函数。模型所有数值在式(2.25)~式(2.32)中建立。

表 2.3 QFD 模型所有数值

x_j	EC_j 的达成程度, $j = 1, 2, 3 \dots n$
x_{jk}	EC_j 的 k 值, 则 $j = 1, 2, 3 \dots n$, $k = 1, 2, 3 \dots p$
s_i	CR_i 的客户满意度, $i = 1, 2, 3 \dots m$
y_i	CR_i 的履行水平, $i = 1, 2, 3 \dots m$

d_{kj}	EC_j ($d_{kj} \neq 0$) 的离散值 k 的满足等级, $j = 1, 2, 3, \dots, n$, $k = 1, 2, 3, \dots, p$
$Comp_i^t$	竞争对手 t 的 CR_i 的值, $i = 1, 2, 3, \dots, m$; $t = 1, 2, 3, \dots, l$
R_{ij}^{norm}	CR_i 和 EC_j 之间的归一化关系,
w_i	CR_i , $i = 1, 2, 3, \dots, m$ 的重要性权重
c_j	EC_j , $j = 1, 2, 3, \dots, n$, 的单位改进成本
C	产品设计总预算
ECL_j	EC_j 的技术约束的下限,
ECH_j	EC_j 的技术约束上限

$$\text{MAX} \sum_{i=1}^m w_i s_i \quad \text{式(2.24)}$$

$$x_j = \sum_{k=1}^p x_{jk} d_{kj} \quad \text{式(2.25)}$$

$$\sum_{k=1}^p x_{jk} = 1 \quad \text{式(2.26)}$$

$$\sum_{j=1}^n c_j x_j \leq C \quad \text{式(2.27)}$$

$$y_i = \sum_{j=1}^n R_{ij}^{norm} x_j \quad \text{式(2.28)}$$

$$y_i \geq \frac{\sum_{t=1}^l Comp_t^i}{l} \quad \text{式(2.29)}$$

$$0 \leq x_j \leq 1 \quad j=1,2,\dots,n. \quad \text{式(2.30)}$$

$$ECL_j < x_j < ECH_j \quad \text{式(2.31)}$$

$$x_{jk} \in (0,1) \quad \text{式(2.32)}$$

式(2.25)是有关离散的技术数性的归一化。式(2.26)施加对于离散 EC 的选择约束。式(2.27)对产品预算限制,将每个 EC 的成本总和分配预算控制范围内。式(2.28)将 EC 的满足水平转换为 CR 的满足水平。式(2.29)是给定顾客偏好数性的最小满意度的约束。每个 CR 的履行水平设置为行业主要竞争对手的平均绩效的基准。式(2.30)将 x_j 定义为从 0 到 1。然而,对于某些 EC,某些技术限制可能限制改进的量。因此式(2.31)规定了某些 EC 的满足水平的额外边界条件。式(32)将 x_{jk} 定义为二进制变量。

<div>CRs \ ECs</div> <div>Relative Importance (%)</div>		1. CPU	2. RAM	3. Hard disk	4. Sound card	5. Graphic card	6. LCD display	7. Battery				
									$Comp^1$	$Comp^2$	$Comp^3$	$Comp^4$
1. Stylish design	14.6	0.1172	0.1172	0.1176	0.0415	0.1176	0.2725	0.2163	0.85	0.74	0.70	0.76
2. Mobility	15.62	0.1816	0.1683	0.1604	0.0923	0.1654	0.0283	0.2036	0.60	0.71	0.45	0.60
3. High computing speed	18.57	0.1876	0.1895	0.1686	0.1356	0.1628	0.0078	0.1481	0.72	0.55	0.40	0.47
4. Powerful graphics solution	16.24	0.1753	0.1702	0.1445	0.1101	0.1685	0.0829	0.1485	0.90	0.84	0.75	0.71
5. Solid audio capability	17.92	0.1909	0.1956	0.1603	0.1807	0.1458	0.0038	0.1230	0.54	0.55	0.61	0.60
6. Large storage	17.05	0.1664	0.2081	0.2280	0.1119	0.1256	0.0060	0.1540	0.83	0.75	0.50	0.66
Cost Index (units)		19.8	17.5	15.5	14	16.5	14.5	13				
Technical Constraints (ECL)		0.65	-	-	-	-	-	-				
Technical Constraints (ECH)		-	-	0.85	-	-	-	-				

图 2.17 QFD 整合参考矩阵[27]

第三章 优化设计方法

3.1 提取消费者偏好属性

在确立分析目标后，消费者偏好属性的提取一共分为 7 个步骤(图 3.1)。

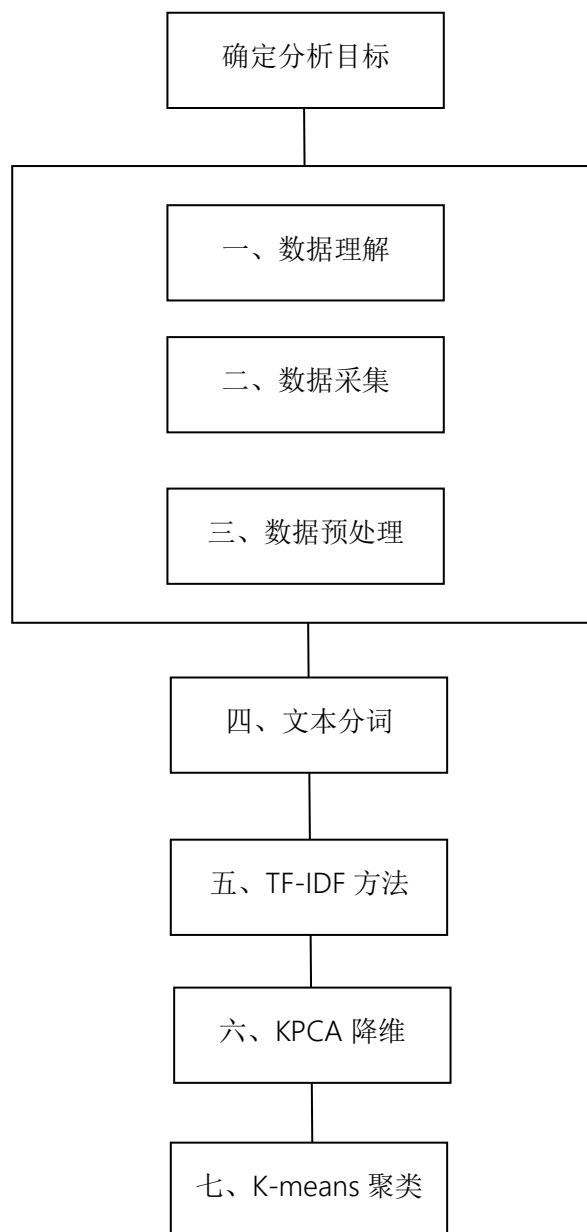


图 3.1 文本数据分析架构图

一、数据理解:

在数据理解阶段一共分为两步骤,在采集目标产品在线评论前,并作一次预采集,全网抽样各价位的商品的“商品价格”与“购买人数”,藉此判断消费者的价格区间落点,将采集数据分类,且计算各区间比例,正式采集时可参考预采集之结果来进行操作。

二、数据采集:

根据预采集结果,依照不同区段的比例来做数据采集,其采集工具可利用 Python 或其他语言编写脚本。

三、数据预处理:

将文本中所有空格去除,以免影响分词,且将所有重复的文本删除,以规避网络刷单的影响。

四、文本分词:

利用现有的 jieba 分词模块进行分词。

五、TFIDF 方法:

利用 TFIDF 生成文本词汇矩阵,根据此矩阵可看出所有文本中所富含的词分布,由于此矩阵维度非常的高,会造成,数量非常庞大,面对如此高维度的词是难以判读其重要程度,在此引入信息熵来做初次的筛选,信息熵又称香农熵,为 1948 年香农 (Claude Shannon) 在其论文“A Mathematic Theory of Communication”中所提出[55],此概念解决了信息量化的问题。信息量和该信息的不确定性有直接的关系,变量不确定性越大,熵也越大,其信息量也就越大[55],其定义如下:

$$H(X) = - \sum_{x \in X} P(x) \log P(x) \quad \text{式}$$

(3.1)

将信息熵的概念套用运用文本词汇矩阵,利用信息论的基础与概率论结合,此方法常用在关键词搜索中,其将每一个词的信息量做为权重 w , 以下为推导:

$$I(w) = -P(w) \log P(w) \quad \text{式}$$

(3.2)

$$= - \frac{\sum_{w \in W} P(w)}{N} \log \frac{\sum_{w \in W} P(w)}{N} \quad \text{式}$$

(3.3)

$$= \frac{TF(w)}{N} \log \frac{N}{TF(w)} \quad \text{式(3.4)}$$

N 为整个与料库的大小为可以省略的常数，但此公式无法反映出词的分辨率，在此做假定：

每个文献大小相同均为 M 个词， D 所有文本数。

$$M = \frac{N}{D} = \frac{\sum_w TF(w)}{D} \quad \text{式(3.5)}$$

一个词汇一但出现不论次数多少，贡献度都相等。

$$c(w) = \frac{TF(w)}{D(w)}, \quad c(w) < M \quad \text{式(3.6)}$$

根据其两个假定可得到以下公式：

$$TF(w) \log \frac{N}{TF(w)} = TF(w) \log \frac{MD}{c(w)D(w)} \quad \text{式(3.7)}$$

$$= TF(w) \log \left(\frac{D}{D(w)} \frac{M}{c(w)} \right) \quad \text{式(3.8)}$$

从以下公式我们可得知 TFIDF 与信息量的差异项在于 $TF(w) \log \frac{M}{c(w)}$ ，在下式可以看到一个词的信息量 $I(w)$ 越多其 TFIDF 的词越大，其差异项越小 TFIDF 越大，此结果与信息论完全相符，其 $c(w)$ 为递减函数此项大于 0：

$$TF - IDF(w) = I(w) - TF(w) \log \frac{M}{c(w)} \quad \text{式(3.9)}$$

六、 KPCA 降维:

降维是将矩阵转置后，将文本数量降维，其方法可见 2.2.3 之推导。

$$L = \begin{bmatrix} a_{11} & a_{1j} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \dots & \dots & a_{ij} & \dots \\ a_{M1} & \dots & \dots & a_{MN} \end{bmatrix} \xrightarrow{T} L^T = \begin{bmatrix} a_{11} & a_{1i} & \dots & a_{1M} \\ a_{21} & a_{22} & \dots & a_{2M} \\ \dots & \dots & a_{ji} & \dots \\ a_{N1} & \dots & \dots & a_{NM} \end{bmatrix} \quad \text{式(3.10)}$$

七、 聚类:

根据转置降维后的矩阵式(3.10)，将词在文本中的 TFIDF 数值作为词的向量，每一个词的向量皆不同，也代表其在高维度的点，藉此利用 K-means 聚类所有的词(方法可见 2.2.4)，根据聚类结果，运用于以下的专家方法。

3.2 调研与专家法分析

得到聚类结果后，调研方法一共分为五个步骤(图 3.2)

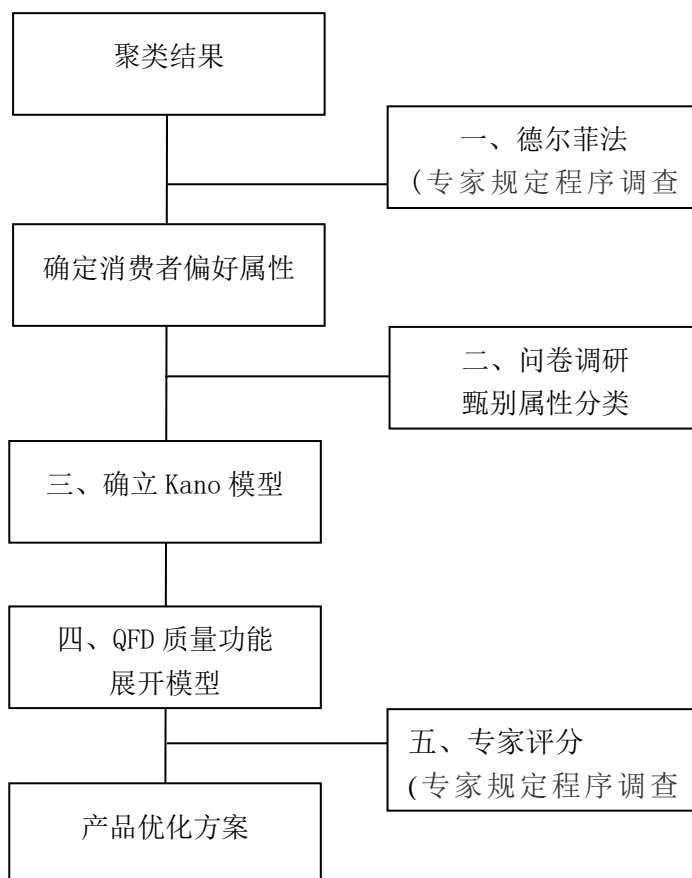


图 3.2 调研与专家法框架

一、德尔菲法:

根据数据分析的聚类结果，汇集该领域的专家利用此方法来做属性分类的调适。

二、问卷调查:

将德尔菲法所得出消费者偏好属性与 Kano 模型各类需求作结合设计问卷，问卷目标为了解各属性的 Kano 需求为何，并统计消费者对各偏好属性的心理归类为何，并藉此做以下的定量分析。

三、确立 Kano 模型:

根据以上问卷可发现，多样化的顾客类型对每一样的产品偏好程度会有所不同，本研究参考[27]方法，将 Kano 问卷结果用定量方法作运算，此方法为三个步骤:

1. 将 Kano 模型的定量分析从计算顾客满意度(CS)与不满意度(DS)

将顾客需求对满意度(CS)与不满意度(DS)对应出一个平均值，并予以百分比表示，为计算顾客需求的满意度数值，在此定义：顾客满意度 CS_i 、顾客不满意度 DS_i 、期望属性 f_o 、必备属性 f_M 、魅力属性 f_A 、无关属性 f_I

顾客满意度，将所有积极属性相加，并除以属性的总数

顾客不满意，度将所有消极属性相加，并除以属性的总数

$$CS_i = \frac{f_A + f_o + f_M}{f_A + f_o + f_I + f_M} \quad \text{式(3.11)}$$

$$DS_i = \frac{f_M + f_I}{f_A + f_o + f_I + f_M} \quad \text{式(3.12)}$$

2. 确立满意度与不满意度的点

若某一产品所提供的顾客需求完全被满足，则顾客需求满足水平为 1，若未能完全满足则为 0，根据此假定，顾客满意度和顾客不满意点。

每一个顾客需求所对应的顾客满意度点：表示为 $(1, CS_i)$

每一个顾客需求所对应的顾客不满意度点：表示为 $(1, -DS_i)$

将顾客满意和顾客不满意的点集成到 Kano 的模型中，CS 和顾客需求水平之间的关系以精确的方式被量化。

表 3.1 KANO 属性分类

属性 类别	A	O	M	I	Total	CS	DS
属性							
X
Y
.....

3. 绘制关系曲线

给定满意度和不满意度的各点, 下一个任务是确定 CS 和客户需求的满足水平(S-CR) 之间的关系的图。水平轴为顾客需求值, 其范围从 0 到 1, 垂直轴为顾客满意度, 范围从-1 到 1。S-CR 关系函数可表示为 $s_i = f(y_i, a, b)$; 其中 s_i 为 CR_i (偏好属性)所实践的满意度水平, S-CR 的函数可以以一般式 $s_i = af(y_i) + b$ 表示 a, b 为调整参数为不同 Kano 类别所提供。

(1) 期望型属性与客户满意度呈现正相关和线性相关可以明确地量化, 函数为 $s_i = a_1 y_i + b_1$, 其中 a_1 为 CS_i 、 DS_i 差, 并且 b_1 是顾客需求值等于 0 时 DS 的值。因此, 给定 $(1, CS_i)$ 和 $(0, DS_i)$

$$a_1 = CS_i - DS_i, \quad b_1 = DS_i \quad \text{式(3.13)}$$

(2) 魅力型属性的 S-CR 曲线不能仅由 CS 和 DS 点确定。该曲线指数函数 $s_i = a_2 e^{y_i} + b_2$ 。 a_2 与此函数近似。 a_2 是调整斜率的参数, b_2 是垂直 DS 值。给定 CS 和 DS 点, 两个参数可以估计为

$$a_2 = \frac{CS_i - DS_i}{e - 1}, \quad b_2 = -\frac{CS_i - e DS_i}{e - 1} \quad \text{式(3.14)}$$

(3) 必须属性的 S-CR 曲线, 能透过指数函数来估计 $s_i = a_3(-e^{-y_i}) + b_3$, 给定 CS 和 DS 点 $(1, CS_i)$ 和 $(0, DS_i)$

$$a_3 = \frac{e(CS_i - DS_i)}{e - 1}, \quad b_3 = \frac{e CS_i - DS_i}{e - 1} \quad \text{式(3.15)}$$

表 3.2 顾客满意度函数表

	a	b	$f(y_i)$	$s_i = af(y_i) + b$
A	$\frac{CS_i - DS_i}{e - 1}$	$-\frac{CS_i - eDS_i}{e - 1}$	e^{y_i}	$s_i = \frac{CS_i - DS_i}{e - 1}e^{y_i} - \frac{CS_i - eDS_i}{e - 1}$
O	$CS_i - DS_i$	DS_i	y_i	$(CS_i - DS_i)y_i + DS_i$
M	$\frac{e(CS_i - DS_i)}{e - 1}$	$\frac{eCS_i - DS_i}{e - 1}$	e^{-y_i}	$-\frac{e(CS_i - DS_i)}{e - 1}e^{-y_i} + \frac{eCS_i - DS_i}{e - 1}$

Kano 的模型将所有消费者偏好根据必备需求、期望需求、魅力需求再做一次分类,在此用于分辨各个消费者偏好属性,其中 $s_i = af(y_i) + b$ 表示通过 Kano 分类后不同偏好属性所表示的 S-CR 函数所获得, s_i 为不同顾客偏好属性(CR)的个别满意度(CS),不同的需求分类也决定了各属性的优先级。

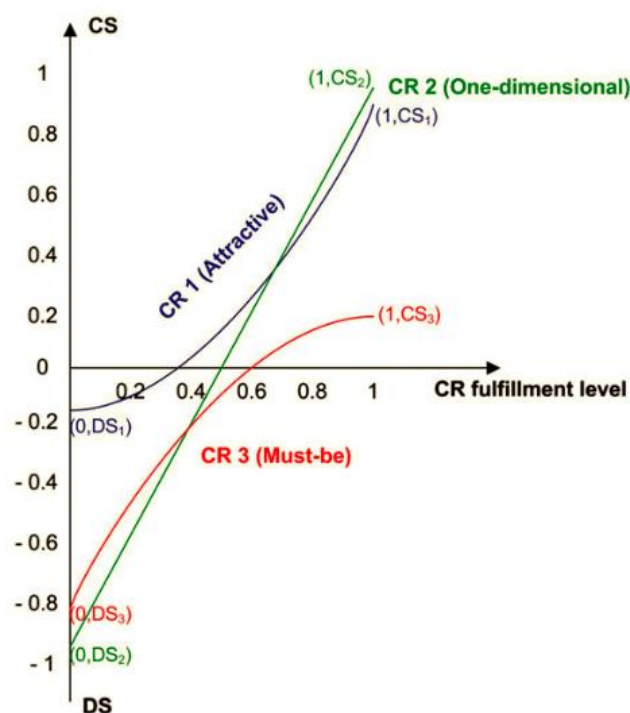


图 3.3 KANO 模型各类满意度函数图

四、 QFD 质量功能展开模型:

此 QFD 模型集成了消费者偏好属性、利用信息熵所求的属性权重与技术环节属性。各偏好属性,由各式各样的词汇所集成,根据各个属性所有集成词的信息熵加总,信息熵越大,其信息量也就越大,进而算出所有词的权重分布)。

五、 专家评分:

利用专家知识,将各个消费者偏好属性与不同技术环节属性之间的相关度评分。

3.3 提出优化方案

根据专家评分建立优化模型，得出完整的 QFD 矩阵，藉此矩阵与 Kano 模型我们将得出各消费者偏好属性与技术环节的优先级。

Kano 模型，在此将所有消费者偏好属性分成三类“期望需求”、“必备需求”、“魅力需求”，根据 Kano 模型之定义，其需求的优先级根据消费者满意度满足的优先次序做为参考，其首要需求为“必备需求”、次要需求为“期望需求”、三级需求为“魅力需求”，再根据不同属性的 S-CR 函数，利用其特性分辨出优先级，“必备需求”利用积分计算出每单位需求达成所需的成本，进而排序其优先级“期望需求”根据其线性地特征，用斜率判断其优先级，“魅力需求”，利用积分计算出每单位正向满意度上升所需的成本。

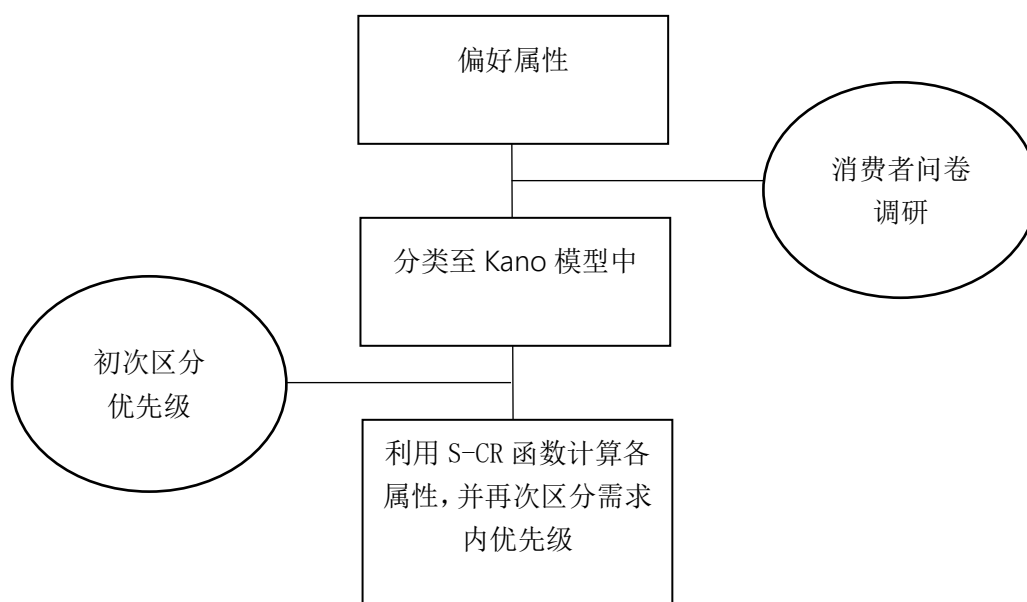


图 3.4 偏好优先级排序

QFD 关系矩阵，根据消费者偏好属性与技术环节属性所集成，矩阵中所有的数值接由 n 个专家评分且算出平均值所得出， T_{11} 为专家判定偏好数性 y_1 与技术环节属性 x_1 之相关性， T_{21} 为专家判定偏好数性 y_2 与技术环节属性 x_1 之相关性， T_{12} 为专家判定偏好数性 y_1 与技术环节属性 x_2 之相关性以此类推，根据每一个技术环节属性数值进行加总，此数值为专家参考偏好数性该技术环节属性的得分，并将所有属性的权重 w_j 乘以 x_{ij} 技术环节属性的相关性得分的总和， $i = \{1, 2, 3, \dots, m\}$ 可算出消费者认为此技术环节属性的重要度，其公式如下：

$$x_j = \sum_{j=1}^n x_{ij} w_j$$

式(3.13)

算出所有技术属性得分与加权(偏好数性权重)后的分相比，可看出设计师与消费者之间的认知差异。

		技术环节属性									
消 费 者 偏 好 属 性		x_1	x_2	x_3						x_n
	y_1	T_{11}	T_{12}
	y_2	T_{21}									
	y_3	T_{31}									
	...										
	y_m										

图 3.5 QFD 矩阵

第四章 基于用户评论的紫砂壶设计优化案例

本研究以国内现有的电子商务网站淘宝网、天猫网作为在线评论数据撷取的对象，并利用 Python 采集了 96 种紫砂壶商品，28358 条文本数据，结构分为三个部分：

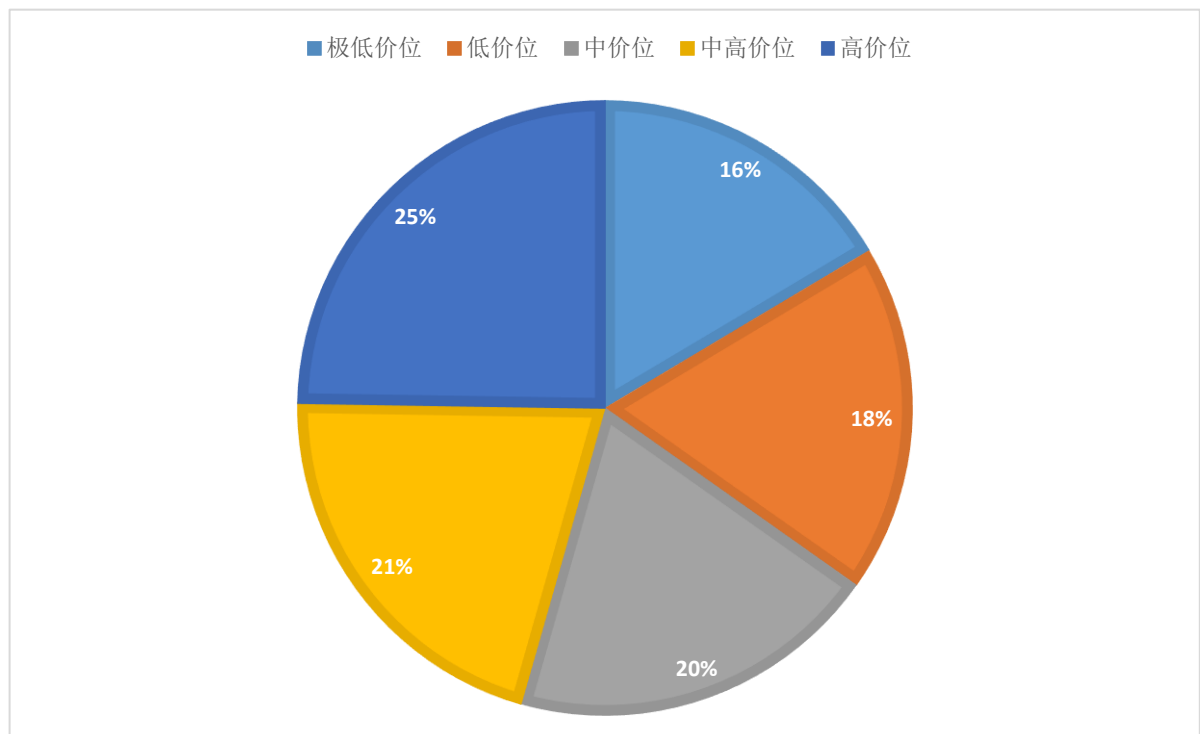
第一部分 分析在线评论中的消费者偏好属性

第二部分 基于在线评论分析，做行业调研分析

第三部分 基于以上两部分(在线评论分析、行业调研分析)，提出设计优化方案

4.1 数据准备与采集

本研究的数据来源为国内电子商务网站淘宝网、天猫网，在两个平台上以关键词“紫砂壶”搜寻，数据准备准备阶段，利用 Python 撰写爬虫爬取全网 2638 样紫砂壶商品，



一共 26948 条次数其销售数量的分布, 在此依照数据平均分布将所有数据分为五个价格区间, 价格 0~50 元分为"极低价位", 价格 50~100 元"低价位", 价格 100~160 元"中价位", 价格 160~800 元"中高价位", 价格 800 元以上"高价位" 其分布的比例为, 极低价位 16.4% , 低价位 18.3%, 中价位 19.6% , 中高价位 20.8% , 高价位 24.7%

图 4.1 紫砂壶全网销售分布

根据以上数据分布, 分价位等比例采集全网"紫砂壶"在线评论数据一共 96 个产品 28358 条在线评论。

表 4.1 数据采集环境

硬件配置	Macpro book 2010 虚拟机: windows10(64bit) 内存: 6G
语言选择	Python: 编写脚本与数据计算
科学计算	Scikit-learn 机器学习数据包

4.2 数据分析

一、分析预处理

数据预处理阶段:根据全网的数据分布, 追评的比例相对少, 因此本研究仅针对首次评论作为研究主要数据, 因此将所有系统默认评论"此用户没有填写评价"、"评价方未及时做出评价,系统默认好评"都先行过滤, 在电子商务市场的激烈竞争下, 国内时常出现不肖商人雇用刷单公司, 在其网店评论刷单、伪造评论, 根据调研市场了解到刷单公司有一整套评论模板, 以方便业务的进行, 为避免伪造评论影响到各词汇的权重判读, 在此将所有一模一样的评论数都只算为一种, 大幅降低伪造评论出现的概率, 根据此方法在 28358 条在线评论数据中, 一共检测出 5218 条相同评论, 剩下 23140 条有效评论数据以供本研究分析。

二、分词与文本向量化

利用 Python 语言 jieba 模块将所有在线评论文本做分词, 程序代码如下:

```
import jieba
import jieba.analyse
import pandas as pd
import numpy as np
import sys
reload(sys)
sys.setdefaultencoding('utf-8')
```

引入 jieba
分词模块

设置系统编码

```

file = open(r'C:\Users\chungyiheng\Desktop\db\data\all.txt', 'rb')
pls = file.readlines()
wset = []
n=0
for i in pls:
    s = jieba.analyse.extract_tags(i, topK=6, withWeight=False, allowPOS=())
    wset += list(s)
    wset = list(set(wset))
    for i in wset:
        print i
    dic = {wset[i]:i for i in range(len(wset))}
    df = np.zeros((len(pls), len(wset)))
    print df.shape
    df1 = pd.DataFrame(df, index=pls)
    for i in pls:
        print n
        s = jieba.analyse.extract_tags(i, topK=6, withWeight=True, allowPOS=())
        for j in s:
            df1.loc[i, dic[j[0]]]=j[1]
        n+=1
    df1.to_csv(r'C:\Users\chungyiheng\Desktop\db\data\all-TFIDF(T).csv')
    a = open(r'C:\Users\chungyiheng\Desktop\db\data\A\all-word.txt', 'w')
    for i in wset:
        a.write(i)
        a.write('\n')
    a.close()

```

将文本转化成
TF-IDF 词频向量

将文本词汇矩阵
储存成 CSV 档

将所有评论文本生成在 txt 文件中，用 Python 脚本去调用运行，并生成 TFIDF 的 csv 文本，以矩阵方式写入，其该矩阵的形式为：

$$L = \begin{bmatrix} a_{11} & a_{1j} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \dots & \dots & a_{ij} & \dots \\ a_{M1} & \dots & \dots & a_{MN} \end{bmatrix} \quad \text{式(4.1)}$$

L 为文本词汇矩阵，其每一个数值代表出现在某一在线评论文本中，所对应词汇的 TFIDF 值，其 a_{ij} 为词汇中的第 j 个词，在第 i 个文本中所出现的 TFIDF 值，就本论文有效文体数据而言，本研究一共有 23140 个评论， $M = 23140$ ，生成 10508 个词汇， $N = 10508$ ， L 矩阵含有 $(M \cdot N)$ 一共含有 243,155,120 个元素。

本研究此高维矩阵能够将词所分类有不只一种，最常用于文本词汇矩阵分析的方法为余弦相似度定理，将每一行或每一列的所有数值视为一向量，将所有向量内的数值，导入到向量空间内，求得两向量之间的夹角，并推算出夹角对应的余弦值，当两个向量的夹角愈接近 0 度(越小)，余弦值越贴近 1。其代表它们的方向更加吻合，更为相似。此方法在 n 维空间表示：

假定 A, B 为 n 维向量， $A = \{A_1, A_2, A_3 \dots A_n\}$ ， $B = \{B_1, B_2, B_3 \dots B_n\}$

$$\cos \theta = \frac{\sum_{i=1}^n (A_i \cdot B_i)}{\sqrt{\sum_{i=1}^n (A_i)^2} \cdot \sqrt{\sum_{i=1}^n (B_i)^2}} \quad \text{式(4.2)}$$

但此种方法仅仅能分辨两词汇之间的相关层度，庞大的数据去两两判读分类，并没有减低太多的属性撷取的工作量，且此方法需要非常多次的迭代，因此耗时非常的长，且上万维的运算在硬件上有一定的要求，本论文在硬件上受限可见表 3.1 数据采集环境的硬件配置，因此本论文不选用此方法，而利用以下方法用时间换取空间分析。

根据第三章内的证明可得出 TFIDF 正比于信息量(信息熵)，因此用信息熵来做第一阶段的词汇筛选，整个词库里的 TFIDF 值分布为线性下降(见图 4.2)，第一次实验利用帕累托定律，在 10000 条评论数据中取值 2000，但 2000 个词对后期的专家模型分析的执行，相对难度较高，此文本词汇矩阵过于稀疏，降维之后聚类效果非常不好(可见附录实验一)。因此在此取信息熵 3.7 以上的词，符合此范围的词汇有 325 个。

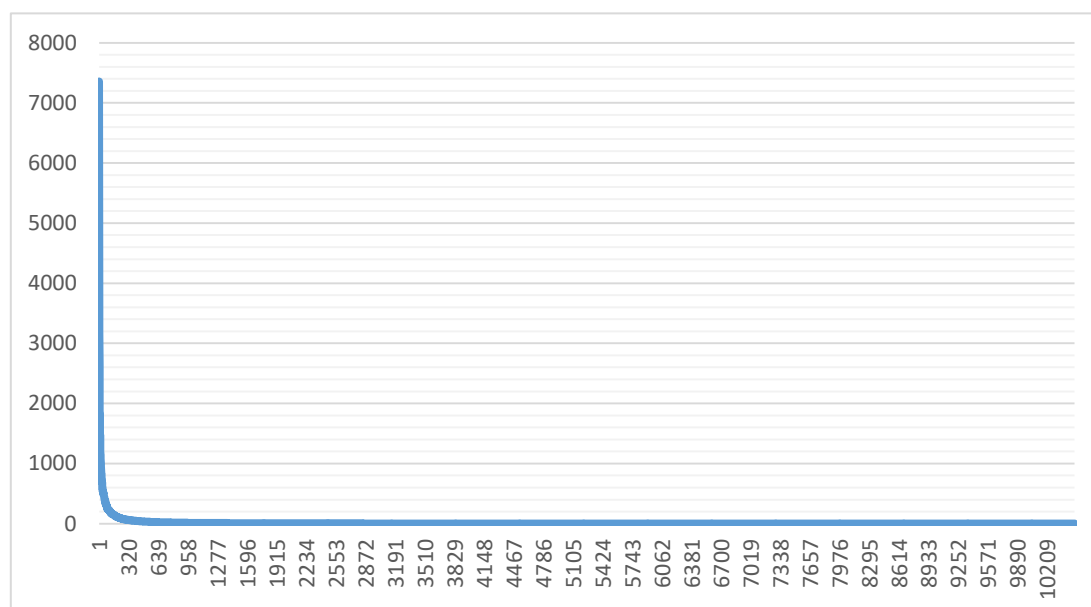


图 4.2 词库 TFIDF 值分布曲线

二、聚类

1. 聚类预处理

- (1) 初次筛选， L 文本词汇矩阵所，将横列从 10508 缩减为 325 个词，并且将此矩阵转置，将矩阵转置，利用文本中各个词在文本数据中所出现的次数作为高维坐标点。

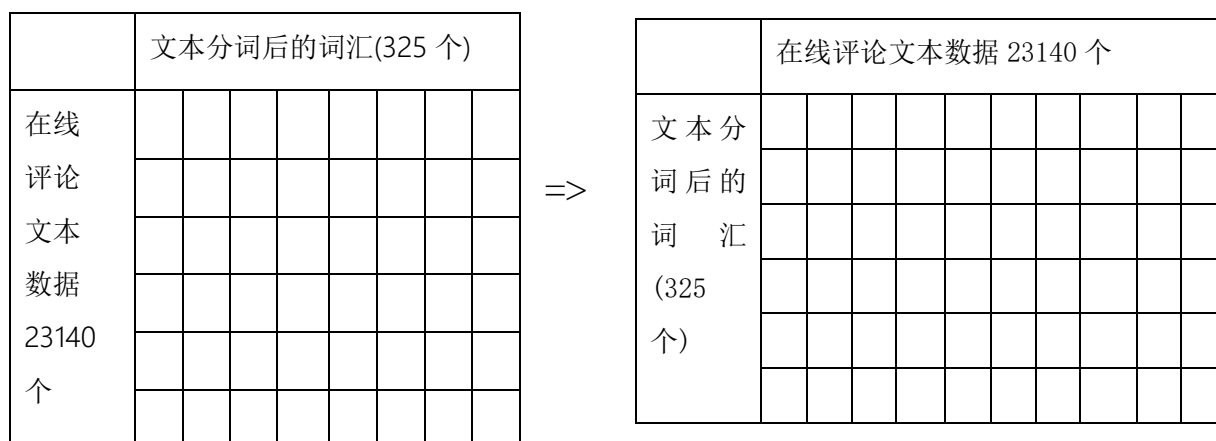


图 4.3 文本词汇矩阵

(2) 降维、归一化

由于此矩阵过于稀疏其聚类，聚类效果不彰，因此使用 KPCA 降维，但 K-means 聚类的缺点为利用欧氏距离作为衡量的目标，其各数值方差的不同将影响聚类效果，因此归一化数据成为首要工作，在此将量纲调整，在 Python 中写入 `x = preprocessing.MinMaxScaler().fit_transform(x)` 将所有的数值映像到 0~1 的区间中，根据 (附录实验一) 一共运用了 KPCA 的五个常用核 (Linear, rbf, sigmoid, poly, cosine) 作为应变变量，控制变量为各维度的在 30 簇的参数调整。

2. K-means 聚类

根据实验一结果，其最好的效果是在 cosine 核低维度时，因此实验二将利用 cosine 做为控制变量，一样设定在 30 簇的范围下，将不同维度的参数代入，并判读优劣。最后取值为 cosine 核 15 维，实验结果见附录中实验二。

4.3 专家模型建模与分析

一、德尔菲法

根据聚类结果，请三个专家进行德尔菲法，将聚类结果做为讨论的议题，经过三轮的反馈，将所有相近的词语归类到一个偏好属性的大类之中，专家们归类的意见如下：

表 4.2 消费者偏好属性集成(1)

属性 分类	功能设计									
偏好 属性	壶盖	壶口	壶把	壶内	壶嘴	手感	出水	密封	结实	原料
词汇 分布	盖子	壶口	壶把	壶内	壶嘴	手感	倒水	密封	结实	紫泥
	壶盖						出水	密封性	严实	泥料
							断水	倒立	漏水	
							流畅			

表 4.3 消费者偏好属性集成(2)

属性分类			造型设计		物流			
偏好属性	容量	表面	款式	颜色	包装	物流	手工	性价比
词汇分布	小巧	粗糙	造型	色差	包装	发货	手工	物有所值
	小巧玲珑	精致	款式	品相		快递	纯手工	实惠
	容量	精细	样式			补发	全手工	超值
	大小	细腻	器型			顺丰		性价比
	迷你	破损	西施			收货		便宜
	太小	细致				到货		物美价廉
	小点	外观				物流		价廉物美
	袖珍							值得
								划算
								很值
								价格便宜
								货真价实

每一个偏好属性，由各式各样的词汇所集成，根据其集成词的信息熵加总，信息熵越大，其信息量也就越大，进而算出所有词的权重分布。

表 4.4 消费者偏好权重

属性分类			造型设计		物流					
偏好属性	容量	表面	款式	颜色	包装	物流	手工	性价比		
	0.259	0.728	0.15	0.075	0.349	0.588	0.136	1		
属性分类	功能设计									
偏好属性	壶盖	壶口	壶把	壶内	壶 嘴	手感	出水	密封	结实	原料
	0.05	0.17	0.029	0.028	0.295	0.014	0.295	0.104	0.11	0.17

二、KANO 模型

本研究调研了 276 个人，收回了 270 个有效问卷，利用 KANO 模型分类且量化各个偏好属性，计算顾客满意度(CS)与不满意度(DS)，并将所有属性归类成期望属性 O、必备属性 M、魅力属性 A、无关属性 I，根据调研结果其各个偏好属性的 Kano 分类如下。

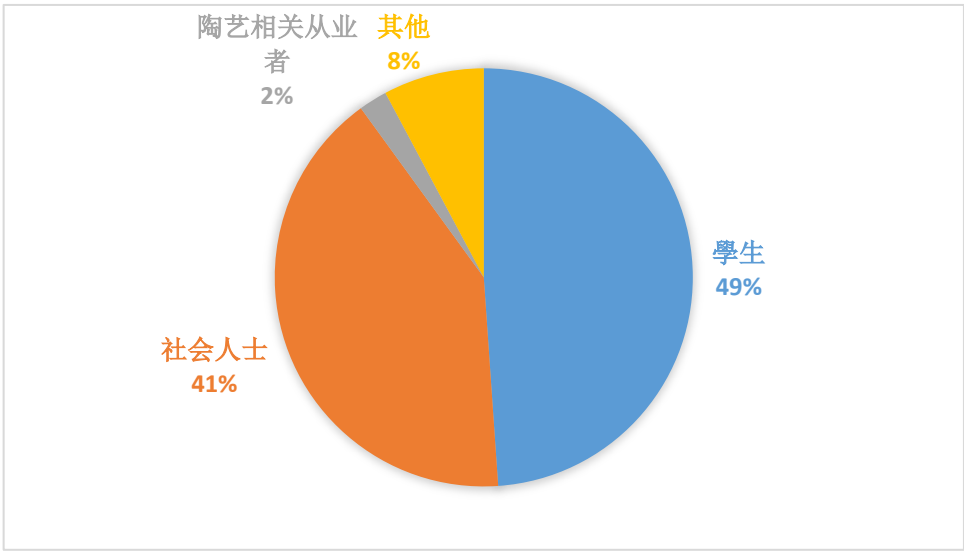


图 4.3 调研问卷分布

根据调研数据，将之代入在式(2.19)、式(2.20)，并根据表 2.2，将不同属性的满意度函数求出：

表 4.5 偏好属性的 Kano 分类

分类	属性	M	O	A	I	Total	Kano	CS	DS
	性价比	87	105	66	12	270	O	0.633333	-0.71111
	物流	120	96	36	18	270	M	0.488889	-0.8
	包装	66	105	69	30	270	O	0.644444	-0.63333
	手工	45	84	102	39	270	A	0.688889	-0.47778
造 型 设计	款式	117	54	90	9	270	M	0.533333	-0.63333
	颜色	91	108	63	8	270	O	0.633333	-0.73704
功 能 设计	出水	183	27	45	15	270	M	0.266667	-0.77778
	壶嘴	96	99	60	15	270	O	0.588889	-0.72222
	壶盖	93	102	70	5	270	O	0.637037	-0.72222
	壶口	102	93	45	30	270	M	0.511111	-0.72222
	壶把	114	90	51	15	270	M	0.522222	-0.75556
	壶内	120	72	48	30	270	M	0.444444	-0.71111
	手感	114	78	75	3	270	M	0.566667	-0.71111
	密封	95	100	70	5	270	O	0.62963	-0.72222
	结实	153	69	33	15	270	M	0.377778	-0.82222
	容量	103	92	48	27	270	M	0.518519	-0.72222
	表面	93	104	60	13	270	O	0.607407	-0.72963
	原料	123	75	57	15	270	M	0.488889	-0.73333

表 4.6 偏好属性的满意度函数

	需求	a	b	$f(y_i)$	$s_i = af(y_i) + b$
性价比	O	1.344444	-0.71111	y_i	$s1 = 1.344443y - 0.71111$
物流	M	-2.03899	1.238992	e^{-y_i}	$s2 = -2.038992e^{-y} + 1.238992$
包装	O	1.277778	-0.63333	y_i	$s3 = 1.277774y - 0.63333$
手工	A	0.678973	-1.15675	e^{y_i}	$s4 = 0.678973e^y - 1.15675$
款式	M	-1.84564	1.212306	e^{-y_i}	$s5 = -1.845634e^{-y} + 1.212306$
颜色	O	1.37037	-0.73704	y_i	$s6 = 1.370373y - 0.73704$
出水	M	-1.65229	0.874509	e^{-y_i}	$s7 = -1.65229e^{-y} + 0.874509$
壶嘴	O	1.311111	-0.72222	y_i	$s8 = 1.311109y - 0.72222$
壶盖	O	1.359259	-0.72222	y_i	$s9 = 1.359259y - 0.72222$
壶口	M	-1.9511	1.228882	e^{-y_i}	$s10 = -1.951101e^{-y} + 1.228882$
壶把	M	-2.02141	1.265859	e^{-y_i}	$s11 = -2.021421e^{-y} + 1.265859$
壶内	M	-1.82806	1.116951	e^{-y_i}	$s12 = -1.82806e^{-y} + 1.116951$
手感	M	-2.02141	1.310304	e^{-y_i}	$s13 = -2.02141e^{-y} + 1.310304$
密封	O	1.351852	-0.72222	y_i	$s14 = 1.35185y - 0.72222$
结实	M	-1.89837	1.07615	e^{-y_i}	$s15 = -1.898369e^{-y} + 1.07615$
容量	M	-1.96282	1.240601	e^{-y_i}	$s16 = -1.96282e^{-y} + 1.240601$
表面	O	1.337037	-0.72963	y_i	$s17 = 1.337037y - 0.72963$
原料	M	-1.93353	1.200194	e^{-y_i}	$s18 = -1.933522e^{-y} + 1.200194$

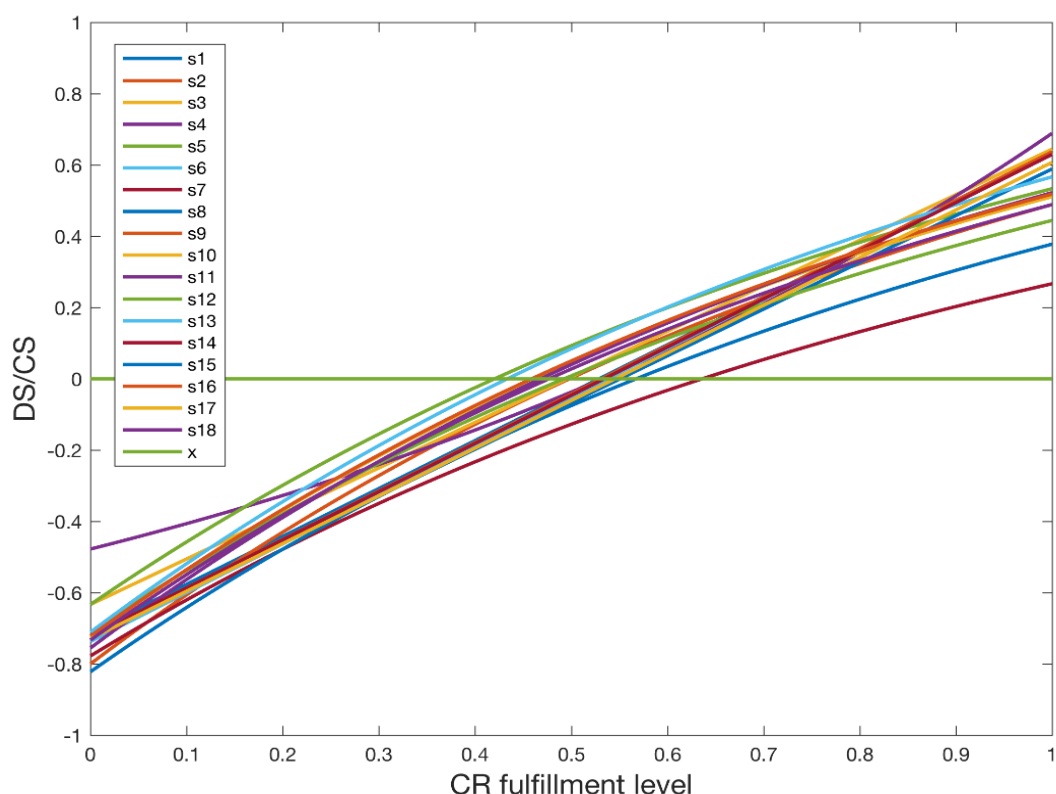


图 4.4 所有属性满意度函数

以上所得出的属性函数，根据 Kano 属性的不同，每个函数的分布情况有所不同

1. 必备需求(M)的实现被消费者认为是理所当然，因此该需求若是多提供也不会提高太多的满意度。所以在此属性只要不让消费者感到不满意即可，图 4.5 每一个必备属性都有自己的函数值，其每一个必备属性的优先级，根据每一条属性函数线与 x 轴所交的面积，即为该属性的必备需求成本，其每单位完成度需耗费的成本与其优先级成反比，其公式为：

$$= \int_0^{f^{-1}(-\frac{b_i}{a_i})} s_i / -\ln \frac{s_i - b_i}{a_i} \quad \text{式(4.3)}$$

表 4.7 必备需求优先级

	需求	函数线	CSWDS=0 耗费成本	CR 完成程度	每单位完成成本	优先级
款式	M	s5	0.1825	0.5199	0.35103	4
出水	M	s7	0.1156	0.8465	0.13656	1
壶口	M	s10	0.1827	0.5063	0.36085	6
壶把	M	s11	0.1825	0.4767	0.38284	8
壶内	M	s12	0.1763	0.6019	0.29291	3
手感	M	s13	0.1808	0.4422	0.40886	9
结实	M	s15	0.1710	0.6391	0.26756	2
容量	M	s16	0.1828	0.4969	0.36788	7
原料	M	s18	0.1822	0.51	0.35725	5

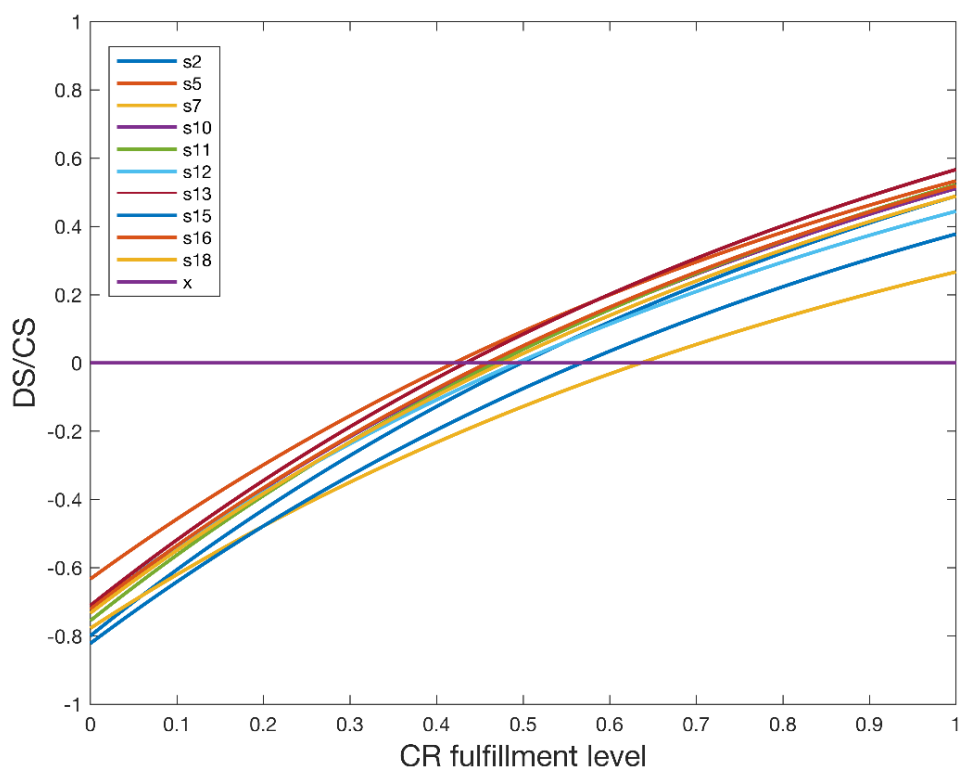


图 4.5 必备需求满意度函数

2. 期望（一元）需求与客户满意度呈现正相关和线性相关，若此需求表现得不好，消费者也不会有太大的满意度下降，根据期望型需求的定义，其需求优先级在必要需求之后，在这之中的消费者偏好属性函数呈现线性，因此其每单位的满意度达成与函数斜率相关。

表 4.8 期望需求优先级

	需求	函数线	函数斜率	优先级
颜色	O	s6	0.782609	1
壶嘴	O	s8	0.72973	5
壶盖	O	s9	0.762712	2
密封	O	s14	0.735695	4
表面	O	s17	0.739726	3

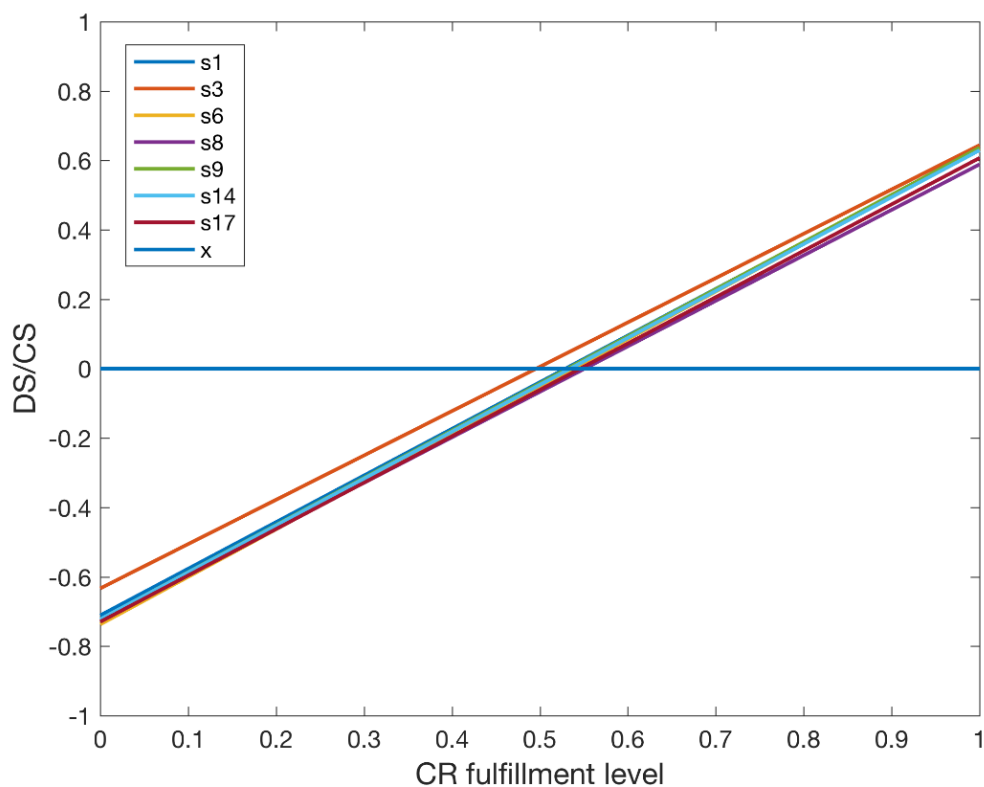


图 4.6 期望需求满意度函数

3. 魅力型需求：魅力型需求是不会被客户过度期待的需求，一但提供则可以提升非常高的满意度。若是缺乏此项也不造成满意度下降。魅力需求为最高层级的需求，因此其优先级在期望需求之后。

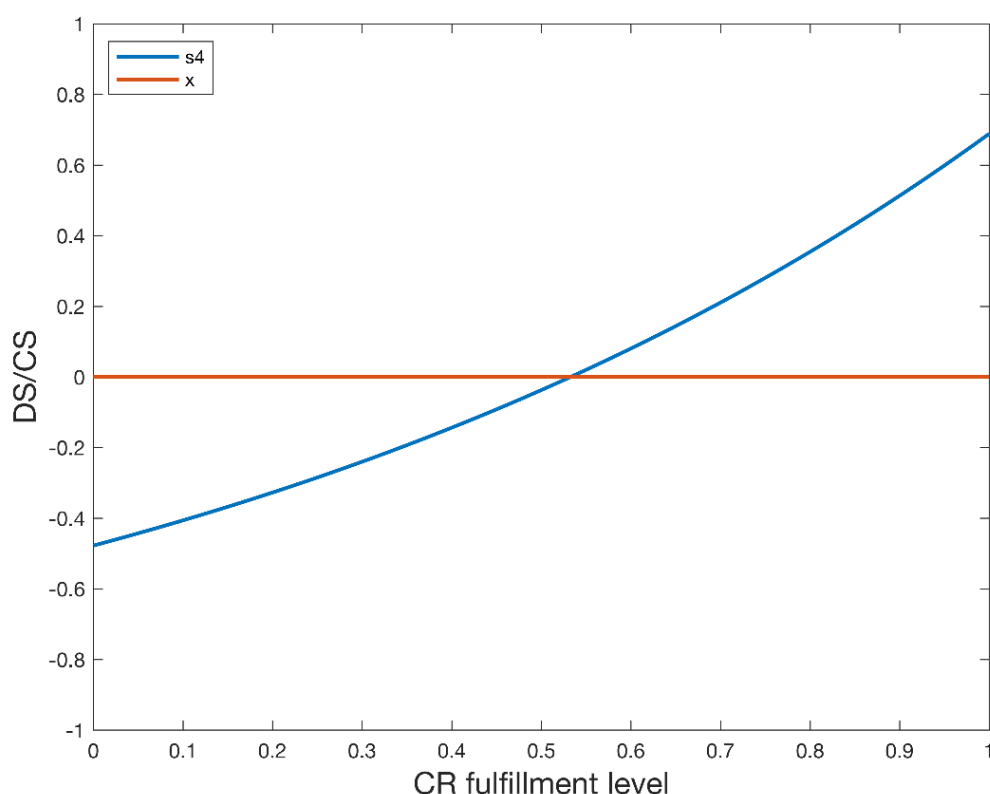


图 4.7 魅力需求满意度函数

一、 QFD 模型

本论文的 QFD 关系矩阵，组成的成分有消费者偏好属性与技术流程的关键环节，由于本研究案例为难以量化的艺术类陶艺产品，不同于一般的科技产品，所有功能都有相应的硬件去支持，茶具产品所有特性皆是环环相扣，且不确定的因素非常高，因此特性将技术流程的关键环节作为分析对象，由此技术属性去分析，作为提供给设计师的优化方案的主轴。

本关系矩阵(表 4.9)，由五位专家依照消费者偏好属性与技术环节之间的相关度大小打分，并计算其平均值，此评分为 6 级评分(0~5 分，完全相关为 5 分)，依照专家经验完成问卷(5 位专家；3 位，从事三年以上的陶艺工作者。2 位景德镇陶瓷大学研究生)。

表 4.9 偏好属性与生产流程矩阵

	权重	打泥片和泥条	打身筒	调脂泥	粘接口、颈、底部	整理粘接细部	做壶盖	做壶嘴	做壶把	钻嘴眼	校正口、嘴、把	光整壶体	窑烧
款式	0.15	3.2	4.4	2.2	2.4	3.6	4.4	4.6	4.8	3.2	4.4	3.2	3.2
颜色	0.075	4	3.6	4.2	2	1.8	3.2	2.8	4	1.2	1.8	2.8	4.2
出水	0.295	2	2.8	2	2.4	2	2.4	5	2.2	4.2	3.8	1.8	2.6
壶嘴	0.046	2	2.4	2.4	2.4	2.6	1.8	4.8	2.2	3.6	3	3.2	2.6
壶盖	0.172	1.8	2.2	2.4	2.4	2.6	4.8	1.8	2.2	1.6	3	3.2	2.6
壶口	0.029	1.8	3	2.2	2.6	3.4	3	1.8	1.6	1.8	3	3.2	2.6
壶把	0.028	2	2	2.4	2.6	2.6	2.2	1.8	4.4	1.6	3	3.2	2.6
壶内	0.014	3.8	3.4	2	1.8	1.8	2.6	1.8	1.2	2.4	1.8	3.4	2.6
手感	0.104	4.4	4.6	4	2.6	4	4.4	4.2	4.6	2	4.2	4.4	4
密封	0.112	2.6	3.6	4	4.4	3.6	4.8	3.4	2	2.8	3.8	2.8	2.4
结实	0.139	3.6	3.8	3.2	3	3	3.4	3	3	1.4	3.2	2	3.8
容量	0.259	1.4	4	1.4	1.8	1.8	1.6	1.6	1.8	1.8	1.6	2	2.2
表面	0.728	4.4	4.4	4	3.6	4.6	2.6	2.6	2.4	2.6	1.8	4.6	3.4
原料	0.16751	5	4.4	4.4	2.6	2	2.6	3.2	3.2	2	2	2	3.4

根据每一个偏好属性的权重，与专家打分后各个技术环节的相关性分数，可得出每一个环节的对于消费者属性偏好的重要度，技术环节重要度为 x_j ，将所有属性的权重 w_j 乘以 x_{ij} 技术环节的相关性的总和， $i = \{1, 2, 3, \dots, m\}$ 其公式如下：

$$x_j = \sum_{j=1}^n x_{ij} w_j \quad \text{式(3.13)}$$

4.4 产品优化方案

根据 QFD 关系矩阵,设计师根据消费者对应的偏好属性,给技术环节相关性打分,得出表 4.10,某些技术环节属性根据问卷中的统计,可清楚看出消费者与设计师在产品上的认知鸿沟,“整理粘接细部”、“做壶盖”、“做壶嘴”,其两者的重要度排名差距是非常大的,在表 4.11 中可看出“壶盖”、“壶嘴”在消费者的认知中都是属于期望属性,其优先级不高,但设计师却将之评为高度重要的选项,在成本有限的情况下,应先考虑必要属性有关的技术环节。

从表中可看出其消费者重要度打分中排列 1、2 名,“打身筒”、“打泥片和泥条”为壶的最重要的主结构,其两个技术环节影响的为壶的造型、泥料扎实程度、壶内外壁,其可对应表 4.11 中,优先级排列第 2、3、4 的“结实”、“壶内”,“款式”,“原料”,因此茶壶设计师可以根据此现象,了解消费者对于紫砂壶而言非常重视其泥料、结实程度,进而调整制作时的工序与思维,根据此二表可使匠人更加了解消费者心理,并且藉由调整制作工序,提升更多的消费者满意度。

表 4.10 消费者认为技术环节权重重要程度

	打泥片和泥条	打身筒	调脂泥	粘接口、颈、底部	整理粘接细部	做壶盖	做壶嘴	做壶把	钻嘴眼	校正口、嘴、把	光整壶体	窑烧
消费者重要度打分	7.59	8.82	7.24	6.62	7.42	6.91	7.11	6.21	5.79	6.13	7.40	7.14
消费者重要度	2	1	5	9	<u>3</u>	<u>8</u>	<u>7</u>	9	11	10	4	6
设计师重要度打分	42	48.6	40.8	36.6	39.4	43.8	42.4	39.6	32.2	40.4	41.8	42.2
设计师重要度	5	1	7	11	<u>10</u>	<u>2</u>	<u>3</u>	9	12	8	6	4

表 4.11 消费者偏好属性优先级

	款式	颜色	出水	壶嘴	壶盖	壶口	壶把	壶内	手感	密封	结实	容量	表面	原料
需求	M	O	M	O	O	M	M	M	M	O	M	M	O	M
优先级	4	10	1	14	11	6	8	3	9	13	2	7	12	5

第五章 结论与展望

5.1 结论

本论文基于国内 C2C、B2C 电商龙头淘宝网、天猫网作为采集数聚的平台，基于紫砂壶用户在线评论分析，以「让设计师、工艺师快速捕捉消费者的特征属性，并在有限的资源下加强优先的工序」此两点作为研究目标，在此基础之上安排研究路径，本研究根据无监督学习法提取消费者偏好属性，用定量方法与专家法相互配合，有别于国内外大多数研究(表 1.1)的分析文本数量小于 2500 以内，且绝大多数皆利用监督学习的方式人为进行归类与分析，在文本分析时过度依赖定性分析，没有产品对设计流程进行太多讨论，本论文根据以上的不足进行优化，分析的文本数量也达到 10 倍以上，定量方法与专家法配合，使结果更加客观且与业内对接的密切度更高，使本研究实践能力更强。

在我们日常生活中，不可能所有商品的规格都可以很明确的量化，茶壶就是一种非常难以量化的产品，所有的属性、结构、制作过程都是环环相扣非单独对应的，例如：“出水”，这一个消费者偏好属性，牵涉到各个层面的工序集结，“窑烧”这个制作工序也是牵涉到太多影响因素，例如：天气、湿度、窑内温度的掌控、用什么方式烧等，设计师与工艺师不可能做出完美的商品，但根据此论文的量化方法，利用更加客观的方法在设计过程中有意识去改进产品细部，本论文主要的研究成果体现在以下两方面：

第一，针对于设计工艺在线领域，对消费者鉴别特征属性提出解决方案，拓展了在线工艺品的设计端研究范围，并根据此建立模型仿真，融合业界专业人士之经验加以汇整，提出目标问题的新研究模型。

第二，在模型分析的过程中，将不同理论相互结合并将引用的 Kano 方法加以改良，用量化方法与信息论更客观的去分析消费者与设计制作者的不同，用这些数值逻辑推断其相互之间的关系，排定工序上消费者对特征属性的优先级，藉此捕捉二者之间思维之差异，对该领域提供更好的优化方案。

5.2 展望

此篇论文，为设计制作端提出了一个新的模型，但本研究的局限性主要有以下三个方面：

一、本研究的数据有以下两方面能作优化，在不同平台、数据数量上可以作拓展，采集数目可以增加，将国内外的其他平台数据导入数据集中，以让此研究结果有在现性、普遍性。

二、本研究所设计的研究路线皆基于线性假定，期望以后的研究人员可以对此作进一步的研究，利用非线性假定的方法(神经网络系统)对此问题作延续。

三、此次研究发现文本分析的方法非常多样，像奇异矩阵分解、余弦相似度分析等方法，模型分析时受硬件内存局限，此新模型可以拓展的地方还有非常多，例如短文本聚类方法的优化，利用不同维度、方法让此一枝干去开枝散叶，除了方法之外，在文本之中也有很多信息尚未被挖掘，像追评、字符长度与情感属性等，都是可以让后续的研究者进一步去探寻的。

参考文献

- [1] 2016 年(上)中国电子商务市场数据监测报告 http://www.100ec.cn/zt/upload_data/B2B/EC.pdf
- [2] 张晓明,王家新,章建刚.中国文化产业报告(2015~2016).社会科学文献出版社 2016
- [3] Chen, C. Y., L. C. Chen, and L. Lin. 2004. "Methods for Processing and Prioritizing Customer Demands in Variant Product Design." *IIE Transactions* 36 (3): 203–219.
- [4] Liu Y, Lu WF, Loh HT. Knowledge discovery and management for product design through text mining—a case study of online information integration for designers. In: *Proceedings of the 16th international conference on engineering design. ICED'07. Paris, France. 2007. 329 (1–12).*
- [5] Akao, Y. 1990. *Quality Function Deployment: Integrating Customer Requirements into Product Design*. Cambridge, MA: Product Press.2004
- [6] Ding X, Liu B. The utility of linguistic rules in opinion mining. In: *Proceedings of the 30th annual international ACM SIGIR conference on research and development in information retrieval. Amsterdam (The Netherlands): ACM; 2007.*
- [7] Hu M, Liu B. Mining and summarizing customer reviews. In: *Proceedings of the 10th ACM SIGKDD international conference on knowledge discovery and data mining. Seattle, WA. 2004. p. 168–77.*
- [8] Chun-Chun Lin, Hsueh-Ying Wu, Yong-Fu Chang. The critical factors impact on online customer. *Procedia Computer Science*, 2010 satisfaction[C]. *Procedia Computer Science*, 2011, 3: 276–281.
- [9] Jumin Lee a Do-Hyung Park , Ingoo Han The effect of negative online consumer reviews on product attitude: An information processing view 2007
- [10] 周晶晶. 在线客户评论对消费者购买决策的影响力研究[硕士学位论文] 浙江大学 2009
- [11] Srinivasan Srinii S., Anderson Rolph, Ponnnavolu Kishore. Customer loyalty in e-commerce: an exploration of its antecedents and consequences[J]. *Journal of Retailing*, 2002, 78 (1): 41-50.
- [12] Kim Hye-Ran. Developing an index of online customer satisfaction[J]. *Journal of Financial Services Marketing*, 2005, 10(1): 49-64.
- [13] 张丽娅.B2C 电子商务在线客户评论情感倾向影响因素研究[硕士学位论文] 河北工业大学 2012.12
- [14] Zhan, Jiaming, Han Tong Loh, and Ying Liu. 2009. "Gather Customer Concerns from Online Product Reviews – A Text Summarization Approach." *Expert Systems With Applications* 36 (2 Part 1): 2107–2115
- [15] Poel, I. 2007. "Methodological Problems in QFD and Directions for Future Development." *Research in Engineering Design* 18 (1): 21–36.
- [16] Xu, Q., R. Jiao, X. Yang, and M. Helander. 2009. "An Analytical Kano Model for Customer Need Analysis." *Design Studies* 30 (1):87–110.
- [17] Liu, Ying, Jian Jin, Ping Ji, Jenny A. Harding, and Richard Y. K. Fung. 2013. "Identifying Helpful Online Reviews: A Product Designer's Perspective." *Computer-Aided Design* 45 (2): 180–194.

- [18] MBA 智库百科 <http://wiki.mbalib.com/zh-tw/KANO%E6%A8%A1%E5%9E%8B>
- [19] Kano, N., N. Seraku, F. Takahashi, and S. Tsuji. 1984. "Attractive Quality and Must-Be Quality." *Hinshitsu: The Journal of the Japanese Society for Quality Control*, April, 39–48.
- [20] 魏丽坤. Kano 模型和服务质量差距模型的比较研究 [J]. 质量管理, 2006, (9): 10-12
- [21] Berger, C., R. Blauth, D. Boger, C. Bolster, G. Burchill, W. DuMouchel, F. Pouliot, et al. 1993. "Kano's Methods for Understanding Customer-defined Quality." *Center for Quality Management Journal*: 3–35.
- [22] Matzler, K., and H. H. Hinterhuber. 1998. "How to Make Product Development Projects More Successful by Integrating Kano's. Model of Customer Satisfaction into Qualify Function Deployment." *Technovation* 18 (1): 25–38. Poel, I. 2007. "Methodological Problems in QFD and Directions for Future Development." *Research in Engineering Design* 18 (1): 21–36.
- [23] MBA 智库百科 <http://wiki.mbalib.com/wiki/QFD%E6%B3%95>
- [24] Ping Ji, Jian Jin, Ting Wang & Yizeng Chen. Quantification and integration of Kano's model into QFD for optimising product design. *International Journal of Production Research*, 2014
- [25] Hashim, A., and S. Dawal. "Kano Model and QFD Integration Approach for Ergonomic Design Improvement." *Procedia - Social and Behavioral Sciences* 57 (9): 22–32. 2012.
- [26] Sireli, Y., P. Kauffmann, and E. Ozan. "Integration of Kano's Model into QFD for Multiple Product Design." *IEEE Transactions on Engineering Management* 54 (2): 380–390. 2007.
- [27] Lee, Y., L. Sheu, and Y. Tsou. "Quality Function Deployment Implementation Based on Fuzzy Kano Model: An Application in PLM System." *Computers & Industrial Engineering* 55 (1): 48–63. 2008.
- [28] Matzler, K., and H. H. Hinterhuber. "How to Make Product Development Projects More Successful by Integrating Kano's Model of Customer Satisfaction into Qualify Function Deployment." *Technovation* 18 (1): 25–38. 1998.
- [29] Mu L-F, Tang J-F, Chen Y-Z, Kwong C-K. A fuzzy multi-objective model of QFD product planning integrating Kano model. *Int J Uncertain Fuzziness Knowl -Based Syst* 2008;793–813.]
- [30] <http://wiki.mbalib.com/wiki/质量屋>
- [31] Lim SCJ, Liu Y, Lee WB. Faceted search and retrieval based on semantically annotated product family ontology. In: *Proceedings of the WSDM'09 workshop on exploiting semantic annotations in information retrieval*. Barcelona (Spain): ACM; 2009.
- [32] 王凯风. 浅议景德镇陶瓷电子商务产业集群的培育与发展[J]. 华南师范大学景德镇学院经济与管理系. 景德镇学院学报, 2016, (01): 26-30.
- [33] 李健. 在线商品评论对产品销量影响研究[J]. 上海大学管理学院. 现代情报, 2012, (01): 164-167.
- [34] 郭潇. 在线评论对旅游预订意向影响的实证分析[D]. 华南理工大学, 2010.
- [35] 皮常玲, 郑向敏. 基于在线评论的民宿顾客抱怨研究 ——以厦门鼓浪屿民宿为例[J]. 华侨大学 旅游学, 2017
- [36] 杨铭, 祁巍, 闫相斌, 李一军. 在线商品评论的效用分析研究[J]. 管理科学报, 2012, (05): 65-75.
- [37] 严建援, 张丽, 张蕾. 电子商务中在线评论内容对评论有用性影响的实证研究[J]. 情报科学, 2012, (05): 713-716+719.

- [38] 张艳辉,李宗伟,赵诣成. 基于淘宝网评论数据的信息质量对在线评论有用性的影响[J]. 管理学报,2017,(01):77-85.
- [39] 王佳园,徐薛艳. 基于网络文本分析的在线短租服务质量评价研究——以途家自营公寓为例[J]. 旅游论坛:1-15.
- [40] 维基百科-TFIDF <https://zh.wikipedia.org/wiki/Tf-idf>
- [41] 施聪莺,徐朝军,杨晓江. TFIDF 算法研究综述[J]. 计算机应用,2009,(S1):167-170+180.
- [42] MBALib <http://wiki.mbalib.com/wiki/主成分分析>
- [43] 周志华.机器学习.北京:清华大学出版社 2016
- [44] MacQueen, J. B. Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press: 281–297. 1967 [2009-04-07].
- [45] Bernhard Schölkopf, Alexander Smola, Klaus-Robert Müller; Kernel Principal Component Analysis 1999
- [46] Peter Harrington.机器学习实战.北京:人民出版社 2013
- [47] 百度百科-Kano model
http://baike.baidu.com/link?url=dtgz8y2er36uW5YuGf6njnilauVJG6xMQ-zEKEpXAcssxQtWEK8ussGbAYvwDTVQZdMPCa2TxYKpCQj3jnWd_g9_qAskpXKaROCok4jZcZZwWA1mtk3Pg8WK74FhLIHYnUFsg4MqaXsR9bvOBvqJCGE5IYhgEDGeSnEr8ipHcUwtFJZFILQsjZrqB4eRbfK
- [48] MBALib <http://wiki.mbalib.com/wiki/QFD%E6%B3%95>
- [49] Lai, X., M. Xie, and K. C. Tan. 2004. “Optimizing Product Design Using the Kano’s Model and QFD.” International Engineering Management Conference 3: 1085–1089.
- [50] Wasserman, G. S. 1993. “On How to Prioritize Design Requirements during the QFD Planning Process.” IIE Transactions 25 (3):59–65.
- [51] Fung, R. Y. K., J. Tang, Paul Yiliu Tu, and Y. Chen. 2003. “Modelling of Quality Function Deployment Planning with Resource Allocation.” Research in Engineering Design 14 (4): 247–255.
- [52] Lai, X., M. Xie, and K. C. Tan. 2005. “Dynamic Programming for QFD Optimization.” Quality and Reliability Engineering International 21 (8): 769–780.
- [53] <http://m.blog.csdn.net/article/details?id=52017275>
- [54] 百度百科-拓扑排序
<http://baike.baidu.com/link?url=8NGoVVuAr9lvepiCmnUlwnh2c1k0Dd5wigoNpt63zSkoDnKDEqAHiuDvCezGcwSu0OLSbcBBsoGRtCmcOvurrDLGaSpZbGLKYRdCMWixjhnvrnriPz12FGwQ1kBzCj1Z>
- [55] 吴军.数学之美.北京 人民邮电出版社 第2版 2014
- [56] 高惠璇.应用多元统计分析.北京 北京大学出版社 2012
- [57] 黄梦婷. 在线商品评论特征及有用性研究——以9类商品为例 [硕士学位论文] 北京: 北京大学信息管理系研究生 2014

- [58] 张智超. 基于在线评论的酒店客户需求特征分析及应用[硕士学位论文] 北京: 北京大学软件应用研究生,2013

附录

实验一：

信息熵取值 1.25 以上的词(一共 2035 个)，根据 KPCA 五个常用核降维，其维度选择维为 1000 维,500 维,250 维,125 维,62 维,31 维,15 维,7 维,3 维,1 维，来并用 K-means 聚类之结果。

维度 原 23000)	簇	方法	结果
1000	30	Linear	0: 2005, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 2
500	30	Linear	0: 68, 1: 1, 2: 1, 3: 1, 4: 2, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 2, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1937, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
250	30	Linear	0: 2006, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
125	30	Linear	0: 2004, 1: 1, 2: 1, 3: 1, 4: 2, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 2, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
62	30	linear	0: 2004, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 3, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
31	30	linear	0: 2002, 1: 2, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 2, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 2, 20: 1, 21: 1, 22: 1, 23: 2, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
15	30	linear	0: 1878, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 2, 20: 1, 21: 11, 22: 13, 23: 2, 24: 1, 25: 1, 26: 17, 27: 4, 28: 1, 29: 86
7	30	linear	0: 165, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 5, 11: 1, 12: 1, 13: 1751, 14: 1, 15: 1, 16: 1, 17: 1, 18: 20, 19: 3, 20: 45, 21: 3, 22:

			4, 23: 1, 24: 1, 25: 1, 26: 11, 27: 4, 28: 4, 29: 2
3	30	linear	0: 1660, 1: 1, 2: 1, 3: 1, 4: 1, 5: 3, 6: 2, 7: 29, 8: 1, 9: 2, 10: 1, 11: 15, 12: 1, 13: 2, 14: 1, 15: 53, 16: 1, 17: 1, 18: 1, 19: 16, 20: 1, 21: 10, 22: 210, 23: 2, 24: 1, 25: 6, 26: 4, 27: 5, 28: 2, 29: 1
1	30	linear	0: 36, 1: 1, 2: 1, 3: 1, 4: 180, 5: 8, 6: 1, 7: 2, 8: 6, 9: 3, 10: 2, 11: 27, 12: 54, 13: 1102, 14: 9, 15: 1, 16: 1, 17: 5, 18: 1, 19: 108, 20: 9, 21: 3, 22: 12, 23: 401, 24: 1, 25: 1, 26: 4, 27: 1, 28: 43, 29: 11

附录

维度 (原 23000)	簇	方法	结果
1000	30	Poly	0: 2005, 1: 1, 2: 1, 3: 1, 4: 2, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
500	30	poly	0: 2005, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 2, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
250	30	poly	0: 2006, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
125	30	poly	0: 2001, 1: 1, 2: 1, 3: 1, 4: 2, 5: 1, 6: 3, 7: 1, 8: 1, 9: 2, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
62	30	poly	0: 2001, 1: 1, 2: 1, 3: 1, 4: 2, 5: 1, 6: 3, 7: 1, 8: 1, 9: 2, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
31	30	poly	0: 1, 1: 2003, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 2, 8: 1, 9: 1, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 2, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
15	30	poly	0: 1918, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 2, 20: 61, 21: 1, 22: 10, 23: 13, 24: 2, 25: 1, 26: 1, 27: 1, 28: 1, 29: 6
7	30	poly	0: 16, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1784, 8: 1, 9: 1, 10: 3, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 7, 20: 33, 21: 156, 22: 4, 23: 1, 24: 4, 25: 1, 26: 3, 27: 1, 28: 2, 29: 4
3	30	poly	0: 1682, 1: 1, 2: 1, 3: 1, 4: 1, 5: 4, 6: 7, 7: 1, 8: 2, 9: 1, 10: 12, 11: 1, 12: 38, 13: 2, 14: 1, 15: 2, 16: 1, 17: 4, 18: 1, 19: 200, 20: 1, 21: 1, 22: 34, 23: 1, 24: 1, 25: 2, 26: 6, 27: 7, 28: 2, 29: 17
1	30	poly	0: 1160, 1: 1, 2: 1, 3: 4, 4: 1, 5: 23, 6: 2, 7: 1, 8: 7, 9: 58, 10: 1, 11: 4, 12: 18, 13: 179, 14: 1, 15: 10, 16: 43, 17: 1, 18: 1, 19: 101, 20: 7, 21: 1, 22: 363, 23: 30, 24: 1, 25: 1, 26: 1, 27: 5, 28: 8, 29: 1

维度 原 23000)	簇	方法	结果
1000	30	rbf	0: 2003, 1: 1, 2: 1, 3: 1, 4: 2, 5: 2, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
500	30	rbf	0: 34, 1: 1971, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 2, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 2, 27: 1, 28: 1, 29: 1
250	30	rbf	0: 1, 1: 2006, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
125	30	rbf	0: 33, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1970, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 3, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 3, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
62	30	rbf	0: 33, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1970, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 3, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 3, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
31	30	rbf	0: 2003, 1: 1, 2: 2, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 2
15	30	rbf	0: 3, 1: 1603, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 23, 19: 1, 20: 3, 21: 1, 22: 86, 23: 1, 24: 2, 25: 2, 26: 36, 27: 254, 28: 2, 29:
7	30	rbf	0: 1423, 1: 1, 2: 2, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 14, 11: 1, 12: 1, 13: 51, 14: 2, 15: 1, 16: 4, 17: 1, 18: 6, 19: 1, 20: 1, 21: 20, 22: 115, 23: 3, 24: 5, 25: 1, 26: 27, 27: 1, 28: 340, 29: 7
3	30	rbf	0: 375, 1: 1, 2: 2, 3: 1, 4: 16, 5: 1, 6: 1, 7: 33, 8: 1, 9: 1, 10: 2, 11: 206, 12: 7, 13: 16, 14: 57, 15: 2, 16: 87, 17: 3, 18: 1, 19: 22, 20: 2, 21: 1175, 22: 3, 23: 1, 24: 10, 25: 1, 26: 1, 27: 2, 28: 4, 29: 1
1	30	rbf	0: 369, 1: 1, 2: 2, 3: 1, 4: 12, 5: 2, 6: 47, 7: 11, 8: 1, 9: 3, 10: 1, 11: 4, 12: 78, 13: 23, 14: 4, 15: 12, 16: 977, 17: 64, 18: 1, 19: 2, 20: 16, 21: 1, 22: 23, 23: 141, 24: 5, 25: 3, 26: 214, 27: 1, 28: 13, 29: 3

附录

维度 原 23000)	簇	方法	结果
1000	30	sigmoid	0: 2005, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
500	30	sigmoid	{0: 1997, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 4, 8: 1, 9: 1, 10: 1, 11: 2, 12: 1, 13: 1, 14: 1, 15: 3, 16: 1, 17: 1, 18: 3, 19: 1, 20: 1, 21: 2, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
250	30	sigmoid	{0: 1997, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 4, 8: 1, 9: 1, 10: 1, 11: 2, 12: 1, 13: 1, 14: 1, 15: 3, 16: 1, 17: 1, 18: 3, 19: 1, 20: 1, 21: 2, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
125	30	sigmoid	0: 1997, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 4, 8: 1, 9: 1, 10: 1, 11: 2, 12: 1, 13: 1, 14: 1, 15: 3, 16: 1, 17: 1, 18: 3, 19: 1, 20: 1, 21: 2, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
62	30	sigmoid	0: 2002, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 2, 7: 2, 8: 2, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 2, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
31	30	sigmoid	0: 2002, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 2, 7: 2, 8: 2, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 2, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
15	30	sigmoid	0: 1911, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 3, 19: 1, 20: 1, 21: 41, 22: 3, 23: 1, 24: 1, 25: 2, 26: 9, 27: 1, 28: 4, 29: 40
7	30	sigmoid	0: 17, 1: 1, 2: 1821, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 5, 14: 1, 15: 1, 16: 1, 17: 4, 18: 1, 19: 1, 20: 110, 21: 1, 22: 1, 23: 4, 24: 1, 25: 1, 26: 3, 27: 31, 28: 7, 29: 13
3	30	sigmoid	0: 45, 1: 1, 2: 1, 3: 1, 4: 1655, 5: 1, 6: 4, 7: 1, 8: 9, 9: 1, 10: 3, 11: 1, 12: 2, 13: 1, 14: 1, 15: 9, 16: 1, 17: 41, 18: 2, 19: 2, 20: 17, 21: 1, 22: 1, 23: 215, 24: 1, 25: 7, 26: 1, 27: 1, 28: 5, 29: 4
1	30	sigmoid	0: 401, 1: 1, 2: 1, 3: 1, 4: 2, 5: 3, 6: 8, 7: 32, 8: 1, 9: 7, 10: 2, 11: 1, 12: 63, 13: 1, 14: 20, 15: 13, 16: 6, 17: 6, 18: 1038, 19: 12, 20: 40, 21: 186, 22: 1, 23: 6, 24: 22, 25: 1, 26: 1, 27: 1, 28: 109, 29: 49

维度 原 23000)	簇	方法	结果
1000	30	Cosine	0: 2005, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
500	30	Cosine	0: 7, 1: 27, 2: 3, 3: 2, 4: 8, 5: 14, 6: 2, 7: 9, 8: 3, 9: 2, 10: 10, 11: 2, 12: 3, 13: 4, 14: 26, 15: 37, 16: 3, 17: 3, 18: 2, 19: 2, 20: 2, 21: 40, 22: 2, 23: 12, 24: 12, 25: 61, 26: 2, 27: 4, 28: 1729, 29: 2
250	30	Cosine	0: 2005, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
125	30	Cosine	0: 2005, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
62	30	Cosine	0: 2005, 1: 1, 2: 1, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 2, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1
31	30	Cosine	0: 137, 1: 44, 2: 331, 3: 4, 4: 3, 5: 6, 6: 3, 7: 19, 8: 3, 9: 3, 10: 121, 11: 7, 12: 2, 13: 3, 14: 17, 15: 3, 16: 40, 17: 6, 18: 19, 19: 3, 20: 49, 21: 13, 22: 197, 23: 10, 24: 6, 25: 844, 26: 9, 27: 22, 28: 8, 29: 103
15	30	Cosine	0: 137, 1: 44, 2: 331, 3: 4, 4: 3, 5: 6, 6: 3, 7: 19, 8: 3, 9: 3, 10: 121, 11: 7, 12: 2, 13: 3, 14: 17, 15: 3, 16: 40, 17: 6, 18: 19, 19: 3, 20: 49, 21: 13, 22: 197, 23: 10, 24: 6, 25: 844, 26: 9, 27: 22, 28: 8, 29: 103
7	30	Cosine	0: 137, 1: 44, 2: 331, 3: 4, 4: 3, 5: 6, 6: 3, 7: 19, 8: 3, 9: 3, 10: 121, 11: 7, 12: 2, 13: 3, 14: 17, 15: 3, 16: 40, 17: 6, 18: 19, 19: 3, 20: 49, 21: 13, 22: 197, 23: 10, 24: 6, 25: 844, 26: 9, 27: 22, 28: 8, 29: 103
3	30	Cosine	0: 19, 1: 63, 2: 206, 3: 7, 4: 3, 5: 2, 6: 350, 7: 10, 8: 7, 9: 1, 10: 29, 11: 3, 12: 4, 13: 311, 14: 10, 15: 5, 16: 134, 17: 93, 18: 1, 19: 4, 20: 1, 21: 38, 22: 193, 23: 18, 24: 400, 25: 73, 26: 7, 27: 32, 28: 7, 29: 4
1	30	cosine	0: 29, 1: 224, 2: 5, 3: 114, 4: 23, 5: 75, 6: 7, 7: 3, 8: 196, 9: 8, 10: 150, 11: 55, 12: 3, 13: 12, 14: 30, 15: 17, 16: 204, 17: 5, 18: 9, 19: 230, 20: 140, 21: 2, 22: 17, 23: 54, 24: 62, 25: 114, 26: 216, 27: 3, 28: 8, 29: 20

一、 实验二

维度 原 23000)	簇	方法	结果
4628	30	Cosine	{0: 1, 1: 1, 2: 62, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 235, 27: 1, 28: 1, 29: 1}
2314	30	Cosine	{0: 1, 1: 1, 2: 62, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 235, 27: 1, 28: 1, 29: 1}
1000	30	Cosine	{0: 1, 1: 1, 2: 62, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 235, 27: 1, 28: 1, 29: 1}
500	30	Cosine	{0: 1, 1: 1, 2: 62, 3: 1, 4: 1, 5: 1, 6: 1, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 235, 27: 1, 28: 1, 29: 1}
250	30	Cosine	{0: 2, 1: 2, 2: 1, 3: 2, 4: 1, 5: 1, 6: 293, 7: 1, 8: 1, 9: 1, 10: 1, 11: 1, 12: 1, 13: 1, 14: 1, 15: 1, 16: 1, 17: 1, 18: 1, 19: 1, 20: 1, 21: 1, 22: 1, 23: 1, 24: 1, 25: 1, 26: 1, 27: 1, 28: 1, 29: 1}
125	30	Cosine	{0: 4, 1: 2, 2: 56, 3: 12, 4: 89, 5: 6, 6: 5, 7: 5, 8: 16, 9: 6, 10: 3, 11: 9, 12: 2, 13: 7, 14: 3, 15: 4, 16: 4, 17: 2, 18: 7, 19: 2, 20: 3, 21: 4, 22: 16, 23: 6, 24: 3, 25: 5, 26: 20, 27: 19, 28: 2, 29: 3}
62	30	Cosine	{0: 6, 1: 78, 2: 2, 3: 6, 4: 3, 5: 28, 6: 4, 7: 3, 8: 3, 9: 2, 10: 2, 11: 19, 12: 4, 13: 4, 14: 5, 15: 2, 16: 2, 17: 2, 18: 45, 19: 3, 20: 2, 21: 2, 22: 3, 23: 2, 24: 3, 25: 3, 26: 3, 27: 23, 28: 30, 29: 31}
31	30	Cosine	{0: 18, 1: 142, 2: 4, 3: 7, 4: 7, 5: 3, 6: 2, 7: 3, 8: 4, 9: 2, 10: 7, 11: 7, 12: 3, 13: 22, 14: 6, 15: 3, 16: 10, 17: 22, 18: 6, 19: 2, 20: 2, 21: 6, 22: 4, 23: 6, 24: 9, 25: 3, 26: 2, 27: 5, 28: 6, 29: 2}
15	30	Cosine	{0: 6, 1: 28, 2: 6, 3: 3, 4: 3, 5: 8, 6: 3, 7: 4, 8: 10, 9: 3, 10: 3, 11: 4, 12: 9, 13: 5, 14: 6, 15: 77, 16: 7, 17: 2, 18: 2, 19: 6, 20: 3, 21: 11, 22: 55, 23: 3, 24: 2, 25: 8, 26: 15, 27: 15, 28: 14, 29: 4}
7	30	Cosine	0: 9, 1: 9, 2: 2, 3: 3, 4: 10, 5: 1, 6: 3, 7: 70, 8: 3, 9: 19, 10: 2, 11: 3, 12: 24, 13: 34, 14: 6, 15: 17, 16: 11, 17: 1, 18: 4, 19: 4, 20: 4, 21: 4, 22: 33, 23: 2, 24: 39, 25: 1, 26: 2, 27: 1, 28: 3, 29: 1

紫砂壶 在线数据源

名称	价 位	网址	名称	价 位	网址
1.宜兴正品特价紫砂壶名家纯手工茶壶茶具套装朱泥现代艺术小天仙壶	50以下	https://item.taobao.com/item.htm?spm=a230r.1.14.209.OVI9aO&id=16262263238&ns=1&abucket=10#detail	48 宜兴全纯手工老段泥紫砂壶茶壶正宗小号功夫泡茶具家用套装金钟壶	800~	https://detail.tmall.com/item.htm?spm=a220m.1000858.1000725.31.1Bp2p2&id=44105319517&skuId=3433426422882&user_id=2450162590&catid=2&is_b=1&rn=3abf3129623dfc4a655e2bff6d3a5847
2.特价宜兴全手工朱泥紫砂壶西施壶过滤小泡茶壶陶瓷茶具套装花茶壶	50以下	https://item.taobao.com/item.htm?spm=a230r.1.14.16.uvS2Op&id=36254308057&ns=1&abucket=10#detail	49重霄堂18降坡泥吴国祥宜兴紫砂壶石瓢紫砂壶茶壶景舟石瓢壶145cc	800~	https://detail.tmall.com/item.htm?spm=a220m.1000858.1000725.106.1Bp2p2&id=43959797011&skuId=3102017817849&user_id=2106219105&catid=2&is_b=1&rn=3abf3129623dfc4a655e2bff6d3a5847
3.宜兴紫砂壶纯手工特价原矿名家全手工茶壶茶具套装黑金沙西施壶	50~100	https://detail.tmall.com/item.htm?spm=a230r.1.14.15.OVI9aO&id=524331033084&ns=1&abucket=10	50 唐家宜兴紫砂壶全纯手工名家大容量竹编竹段茶壶茶具套装家用石瓢	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.22.HFfzq4&id=536162350605&ns=1&abucket=16
4 宜兴紫砂壶名家全手工紫泥段泥球孔紫砂壶西施壶扁特价小茶壶茶具	50~100	https://item.taobao.com/item.htm?spm=a230r.1.14.43.HzdApJ&id=536755119904&ns=1&abucket=5#detail	51【湘谊】宜兴紫砂壶名家正品纯全手工茶壶原矿朱泥王品荣 西施壶	800~	https://detail.tmall.com/item.htm?spm=a220m.1000858.1000725.186.1Bp2p2&id=36424907829&skuId=3145329094598&user_id=773772503&catid=2

附录

					d=2&is_b=1&rn=3abf3129623dfc4a655e2bff6d3a5847
5 微瑕疵壶清仓 宜兴紫砂壶纯全 手工泡茶壶茶具 球孔 西施石瓢 仿古壶	50~ 100	https://item.taobao.com/item.htm?spm=a230r.1.14.159.HzdApJ&id=527422174909&ns=1&abucket=5#detail	52 藏壶天下宜 兴紫砂壶纯全 手工石瓢壶精 品茶壶稀有原 矿碧潭砂东坡 瓢	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.36.HFfzq4&id=520342096742&ns=1&abucket=16&skuId=3154901824329
6 宜兴紫砂壶 原 矿扁西施 黄金 段泥（大球孔出 水）200 毫升 有 证书	50~ 100	https://item.taobao.com/item.htm?spm=a230r.1.14.200.HzdApJ&id=528866322798&ns=1&abucket=5#detail	53 紫泥砂韵 宜兴原矿名家 纯全手工紫砂 壶朱泥西施茶 壶小品套装特 价	800~	https://item.taobao.com/item.htm?spm=a230r.1.14.43.HFfzq4&id=539386166271&ns=1&abucket=16#detail
7 宜兴紫砂壶 名 家正品 纯全手 工 原矿紫泥 188 球孔西施 茶 壶 特价	100 ~16 0	https://item.taobao.com/item.htm?spm=a230r.1.14.214.HzdApJ&id=520918694347&ns=1&abucket=5#detail	54 润砂 宜兴 紫砂壶纯全手 工茶具名家小 茶壶原矿紫砂 套装泡茶壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.57.HFfzq4&id=529568418998&ns=1&abucket=16
8 宜兴全手工定 制紫砂壶大小品 石瓢西施壶纯名 家刻字特价茶壶 茶具	100 ~16 0	https://item.taobao.com/item.htm?spm=a230r.1.14.172.HzdApJ&id=541086070268&ns=1&abucket=5#detail	55 名家 宜兴 紫砂壶纯全手 工朱泥大红袍 鱼化龙茶壶茶 具套装真正精 品	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.85.HFfzq4&id=534987174982&ns=1&abucket=16
9 宜兴紫砂壶纯 全手工名家正品 特价刻绘西施壶 茶壶茶具	50~ 100	https://item.taobao.com/item.htm?spm=a230r.1.14.287.HzdApJ&id=538878992587&ns=1&abucket=5#detail	56 壶光砂色 宜兴紫砂壶名 家徐鹏全纯手 工茶壶茶具大 红袍西施壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.78.HFfzq4&id=528778256066&ns=1&abucket=16

		cket=5#detail			
10 紫砂壶宜兴正 品大品大容量茶 壶紫砂名家清仓 纯全手工四方壶 1500cc	50~ 100	https://item.taobao.com/item.htm?spm=a230r.1.14.254.HzdApJ&id=39113746455&ns=1&abbucket=5#detail	57 藏壶天下宜 兴紫砂壶全手 工原矿茶壶茶 具 一厂老紫泥 岁寒松趣	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.71.HFfzq4&id=45187519675&ns=1&abbucket=16
11 宜兴紫砂壶名 家纯全手工原矿 龙血砂清仓特价 正品精品茶壶茶 具限量	50~ 100	https://item.taobao.com/item.htm?spm=a230r.1.14.1.HzdApJ&id=540997845256&ns=1&abbucket=5#detail	58 宜兴紫砂壶 功夫茶具套装 名家全手工纯 原矿紫泥底槽 清茄段茶壶	800~	https://item.taobao.com/item.htm?spm=a230r.1.14.64.HFfzq4&id=539043647408&ns=1&abbucket=16#detail
12 宜兴紫砂壶西 施壶原矿全纯手 工卡盖功夫过滤 陶瓷茶壶茶具套 装特价	50~ 100	https://item.taobao.com/item.htm?spm=a230r.1.14.21.HzdApJ&id=529795534257&ns=1&abbucket=5#detail	59 荆邑紫砂壶 纯全手工茶壶 名家茶具套装 杯周小明宜兴 石瓢龟龙珍藏	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.50.HFfzq4&id=526589330487&ns=1&abbucket=16
13 宜兴正品紫砂 壶名家纯全手工 茶具绿泥仿古石 瓢微瑕疵清仓捡 漏特价	100 ~16 0	https://item.taobao.com/item.htm?spm=a230r.1.14.81.etgn6f&id=525591303235&ns=1&abbucket=5#detail	60 名家 宜兴 紫砂壶纯全手 工原矿紫泥刻 绘心经大功夫 泡茶壶茶具正 品	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.41.HzdApJ&id=520458296506&ns=1&abbucket=5
14 宜兴紫砂壶名 家纯全手工石瓢 壶 原矿紫泥泡 茶壶茶具套装正 品	100 ~16 0	https://item.taobao.com/item.htm?spm=a230r.1.14.85.etgn6f&id=530668749757&ns=1&abbucket=5#detail	61【古悦堂】宜 兴紫砂壶纯全 手工名家底槽 清泥雄霸四方 茶壶茶具套装	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.6.HzdApJ&id=523014645671&ns=1&abbucket=5

附录

15 宜兴紫砂壶纯全手工紫泥底槽清西施壶正品名家茶具球孔泡茶壶特价	100~160	https://item.taobao.com/item.htm?spm=a230r.1.14.127.etgn6f&id=527253921666&ns=1&abbucket=5#detail	62 玖砂 宜兴名家紫砂壶原矿纯手工龙血砂大红袍茶壶茶具筋囊壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.257.HzdApJ&id=43765202918&ns=1&abbucket=5
16 宜兴紫砂壶纯全手工原矿大红袍朱泥西施壶泡茶壶茶具	100~160	https://item.taobao.com/item.htm?spm=a230r.1.14.134.etgn6f&id=38775689160&ns=1&abbucket=5#detail	63 玖砂 宜兴紫砂壶名家全手工茶壶名家徐老师描金仙鹤秋水壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.243.HzdApJ&id=532080271013&ns=1&abbucket=5
17 紫砂壶宜兴正品大茶壶纯全手工清仓名家扁玉西施壶石瓢壶捡漏茶具	100~160	https://item.taobao.com/item.htm?spm=a230r.1.14.189.etgn6f&id=35747216321&ns=1&abbucket=5#detail	64【畅陶】宜兴紫砂壶 名家陶建纯全手工茶壶 珍藏满天星西施瓢壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.195.HzdApJ&id=40720242759&ns=1&abbucket=5&skuId=3452081459490
18 嘉木宜兴正品纯全手工原矿紫泥卡盖球孔西施壶紫砂茶具泡茶壶套装	100~160	https://item.taobao.com/item.htm?spm=a230r.1.14.224.AsNHwV&id=536794027286&ns=1&abbucket=5#detail	65 国工泥人周冠华 名家碧玉紫砂壶宜兴原矿大红袍朱泥纯全手工茶壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.161.HzdApJ&id=537899637742&ns=1&abbucket=5
19 宜兴紫砂壶纯全手工原矿老段泥西施壶名家正品茶具球孔小茶壶	100~160	https://item.taobao.com/item.htm?spm=a230r.1.14.286.AsNHwV&id=526197494711&ns=1&abbucket=5#detail	66 名家 宜兴紫砂壶纯全手工原矿老朱泥仿古如意壶大泡茶壶茶具套装	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.154.HzdApJ&id=37442483128&ns=1&abbucket=5
20 降坡泥紫砂壶	100~160	https://item.taobao.com/item.htm?spm=a230r.1.14.286.AsNHwV&id=526197494711&ns=1&abbucket=5#detail	67 壶光砂色宜兴紫砂壶名家	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.154.HzdApJ&id=37442483128&ns=1&abbucket=5

宜兴原矿名家纯全手工内壁章刻绘莲花紫砂小茶壶特价	0	=a230r.1.14.88.AsNHwV&id=538661217328&ns=1&abucket=5#detail	纯全手工段泥鱼籽砂茶壶茶具套装特价卧禅		91.HzdApJ&id=531862036120&ns=1&abucket=5&skuId=3210813884603
21 宜兴正品紫砂壶原矿民国绿泥名家纯全手工打造一品竹段拍下送两杯	100~160	https://item.taobao.com/item.htm?spm=a230r.1.14.1.AsNHwV&id=536641924433&ns=1&abucket=5#detail	68【畅陶】紫砂名家工艺师 董大师全手工紫砂壶 原矿紫泥飞跃	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.29.HzdApJ&id=525184058740&ns=1&abucket=5
22 宜兴名家紫砂壶纯全手工实力一厂艺人华丽娟民国绿泥扁玉西施壶	100~160	https://item.taobao.com/item.htm?spm=a230r.1.14.124.AsNHwV&id=541521196550&ns=1&abucket=5#detail	69 金奖 宜兴紫砂壶时来运转纯全手工高档原矿茶壶茶具套装正品特价	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.8.HzdApJ&id=534434905412&ns=1&abucket=5
23 宜兴原矿老紫泥全手工紫砂壶套装礼品整套茶具汉云茶壶包邮特价壶		https://detail.tmall.com/item.htm?spm=a230r.1.14.29.AsNHwV&id=538349035543&ns=1&abucket=5	70 玖砂 宜兴紫砂壶全手工资深国工徐老师描 24K 真金梨皮老朱砂纳福壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.243.HzdApJ&id=528566134655&ns=1&abucket=5
24 名家全手工宜兴紫砂壶茶壶民国绿泥 紫泥 300cc 玉兰壶纯手工		https://detail.tmall.com/item.htm?spm=a230r.1.14.231.eddXHx&id=43680062075&ns=1&abucket=2	71 金奖 宜兴紫砂壶纯全手工原矿紫泥竹节段石瓢壶大功夫泡茶壶茶具	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.161.HzdApJ&id=19697322471&ns=1&abucket=5
25 宜兴紫砂壶纯手工特价原矿名家全手工茶壶茶具套装七彩砂西		https://detail.tmall.com/item.htm?spm=a230r.1.14.22.eddXHx&id=542635215	72 藏壶天下宜兴原矿紫砂壶名家纯全手工茶壶原矿底槽	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.22.HzdApJ&id=42107691393&ns=1&abucket=5

附录

施壶		045&ns=1&abbucket=2&skuId=3426004623291	青茶具古树系列		
26 悦目清心 宜兴名家纯全手工紫砂壶 功夫泡茶茶具套装 紫泥秦权壶		https://detail.tmall.com/item.htm?spm=a230r.1.14.57.eddXHx&id=543462899588&ns=1&abbucket=2	73 藏壶天下宜兴紫砂壶全手工正品茶壶戴晨光精品原矿绿泥紫泥星月	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.265.HzdApJ&id=38583865747&ns=1&abbucket=5&skuId=3416733206936
27 嘉誉紫砂 包邮手工 宜兴紫砂壶 原矿降坡泥 球孔嵌盖小玉乳半月壶		https://detail.tmall.com/item.htm?spm=a230r.1.14.286.eddXHx&id=44637398417&ns=1&abbucket=2	74 名家 宜兴紫砂壶纯全手工原矿老紫泥石瓢壶竹节大泡茶壶茶具套装	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.113.HzdApJ&id=523182442348&ns=1&abbucket=5
28. 藏壶天下宜兴紫砂壶纯全手工名家原矿老紫泥吉祥如意茶壶茶具套装	160~800	https://detail.tmall.com/item.htm?spm=a230r.1.14.34.HFfq4&id=45477459075&ns=1&abbucket=16&skuId=3457885022498	75 藏壶天下宜兴名家紫砂壶纯全手工倪新安高工原矿茶壶将军壶系列	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.106.HzdApJ&id=45379516112&ns=1&abbucket=5
29 【畅陶】宜兴紫砂壶纯全手工家用套装名家李晓璐黑金沙西施茶壶具	160~800	https://detail.tmall.com/item.htm?spm=a230r.1.14.6.HFfq4&id=524873703307	76 【不亦乐壶】宜兴紫砂壶名家全手工茶壶茶具套装 六方井栏紫砂壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.57.HzdApJ&id=41251960537&ns=1&abbucket=5
30 特价宜兴紫砂壶茶壶名家胡艳霞纯全手工正品原矿七星砂朱泥扁玉	160~800	https://item.taobao.com/item.htm?spm=a230r.1.14.239.HzdApJ&id=521470708075&ns=1&abbucket=5#detail	77 藏壶天下宜兴原矿紫砂壶十二生肖系列纯全手工茶壶原矿紫泥茶具套	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.22.HzdApJ&id=522032504235&ns=1&abbucket=5

31【畅陶】宜兴名家纯手工紫砂壶大红袍朱泥西施茶壶全手工李晓璐	160~800	https://detail.tmall.com/item.htm?spm=a230r.1.14.8.HzdApJ&id=542994153277&ns=1&abbucket=5	78【畅陶】宜兴全手工紫砂壶茶壶名家陶建真品满天星泥西施壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.286.HzdApJ&id=40623006308&ns=1&abbucket=5&skuId=3287960537798
32 藏壶天下宜兴紫砂壶套装纯全手工泡茶茶壶家用名家特价石瓢壶茶具	160~800	https://detail.tmall.com/item.htm?spm=a230r.1.14.15.HzdApJ&id=19405558977&ns=1&abbucket=5	79 藏壶天下宜兴紫砂壶全手工原矿茶壶戴晨光底槽青景舟石瓢	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.293.HzdApJ&id=523808746604&ns=1&abbucket=5
33 御壶宜兴全纯手工紫砂壶正品特价茶壶茶具套装原矿底槽清景舟石瓢	160~800	https://item.taobao.com/item.htm?spm=a230r.1.14.92.etgn6f&id=522803435807&ns=1&abbucket=5#detail	80 藏壶天下宜兴原矿紫砂壶十二生肖系列纯全手工茶壶原矿紫泥茶具套	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.112.x8eYkB&id=522032504235&rn=354ce0cec51221eacbc2d1e8daf375b1&abbucket=6
34 依云堂 宜兴紫砂壶名家纯全手工正品茶壶原矿紫泥汉瓦壶茶具特价	160~800	https://item.taobao.com/item.htm?spm=a230r.1.14.99.etgn6f&id=536260452117&ns=1&abbucket=5#detail	81 藏壶天下宜兴名家全手工紫砂壶茶具倪新安高工红泥玉兔	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.172.IMECjE&id=527410872552&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6
35 天天特价】张杰紫砂宜兴纯手工紫砂壶正品特价紫泥石瓢壶全	160~800	https://item.taobao.com/item.htm?spm=a230r.1.14.141.etgn6f&id=525144501306&ns=1&abbucket=5#detail	82 藏壶天下宜兴名家全手工紫砂壶茶具倪新安高工红泥玉兔	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.172.IMECjE&id=527410872552&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6

附录

36 【清风竹韵】 宜兴名家紫砂壶 纯全手工西施石 瓢泡茶壶茶具套 装家用	160 ~80 0	https://item.taobao.com/item.htm?spm=a230r.1.14.15.etgn6f&id=41637446470&ns=1&abbucket=5#detail	83 藏壶天下 宜兴名家全手 工紫砂壶高工 倪新安原矿紫 泥绿泥冰纹四 方	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.168.IMECjE&id=520358120820&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6
37 【陶源】名家 李晓璐 宜兴纯 全手工紫砂壶茶 壶套装 黑金砂 石瓢壶	160 ~80 0	https://item.taobao.com/item.htm?spm=a230r.1.14.267.etgn6f&id=43644536926&ns=1&abbucket=5#detail	84 藏壶天下宜 兴紫砂壶鱼化 龙名家全手工 纯原矿朱泥紫 泥大茶壶茶具	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.156.IMECjE&id=523279430278&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6
38 【天天特价】 宜兴名家全手工 紫砂壶卡盖西施 原矿黑星砂泡茶 壶茶壶	160 ~80 0	https://item.taobao.com/item.htm?spm=a230r.1.14.232.etgn6f&id=529684685189&ns=1&abbucket=5#detail	85 藏壶天下宜 兴原矿紫砂壶 名家纯全手工 茶壶功夫茶具 清水泥南瓜壶	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.140.IMECjE&id=521843865132&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6
39 天天特价 宜 兴名家紫砂壶 纯全手工茶壶茶 具朱泥大红袍扁 圆珠西施	160 ~80 0	https://item.taobao.com/item.htm?spm=a230r.1.14.204.etgn6f&id=537997695987&ns=1&abbucket=5#detail	86 藏壶天下宜 兴紫砂壶名家 全手工原矿茶 壶茶具 家藏老 紫泥 鲤鱼龙门	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.144.IMECjE&id=524862216289&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6&skuId=3281089933412
40 雨轩宜兴紫砂 壶纯全手工正品 特价名家原矿老 紫泥秦权茶壶泡	160 ~80 0	https://item.taobao.com/item.htm?spm=a230r.1.14.211.etgn6f&id=822910424	87 藏壶天下 宜兴紫砂壶全 手工原矿茶壶 戴晨光 原矿老	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.132.IMECjE&id=2276423

茶套装		0&ns=1&abbucket=5#detail	红泥 樱桃壶		6394&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6&skuId=3109721066101
41 陶渊 宜兴紫砂壶名家纯全手工仿古壶大红袍暗香茶壶茶具家用套装	160~800	https://item.taobao.com/item.htm?spm=a230r.1.14.148.etgn6f&id=522061559010&ns=1&abbucket=5#detail	88 鱼化龙 紫砂壶 藏壶天下 宜兴精品全手 89 工茶壶原矿 赵庄朱泥鱼跃龙门	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.128.IMECjE&id=527071320393&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6
42 【玉颐阁】 宜兴紫砂壶全手工茶壶原矿老紫泥名家刘浩 玉乳壶	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.29.HFfzq4&id=25892440433&ns=1&abbucket=16	90 藏壶天下宜兴紫砂壶名家倪新安正品全手工原矿底槽青龙腾壶	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.120.IMECjE&id=522563438318&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6
43 金奖 宜兴紫砂壶纯全手工石瓢西施仿古大功夫泡茶壶茶具套装特价	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.1.HFfzq4&id=19712013798&ns=1&abbucket=16	91 藏壶天下宜兴名家全手工紫砂壶 倪新安高工原矿老紫泥龙行天下	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.112.IMECjE&id=45724247509&rn=ce461373ff87960cd8b8fac1a7375108&abbucket=6
44 金奖 宜兴紫砂壶纯全手工名家朱泥大红袍鱼化龙壶泡茶壶茶具套装	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.8.HFfzq4&id=26192092563&ns=1&abbucket=16	92 藏壶天下 宜兴全手工紫砂壶收藏升值老戴 早期稀少天青砂 君玉	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.120.WJTVFD&id=38822929623&rn=56b140375279307d910d06ed0b7978a0&abbucket=6&skuId=3200910251797

附录

45 藏壶天下宜兴紫砂壶名家纯全手工茶壶专利原矿紫红泥紫砂茶具月影	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.15.HFfzq4&id=35342855734&ns=1&abbucket=16	93 藏壶天下宜兴紫砂壶全手工名家戴尔其收藏茶壶茶具老紫泥牡丹春色	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.16.WJTVFD&id=522643682351&rn=56b140375279307d910d06ed0b7978a0&abbucket=6&skuId=313630224151
46 名家 宜兴紫砂壶纯全手工原矿段泥大竹节提梁壶功夫茶壶茶具正品	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.78.X9IvpF&id=21986103004&ns=1&abbucket=16	94 藏壶天下宜兴紫砂壶全手工正品茶壶名家老戴早期天青砂天际	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.12.WJTVFD&id=38547598799&rn=56b140375279307d910d06ed0b7978a0&abbucket=6&skuId=3194750938660
47 润砂 正宗宜兴紫砂壶西施瓢如意纯全手工名家功夫老紫泥泡茶茶具	800~	https://detail.tmall.com/item.htm?spm=a230r.1.14.1.HFfzq4&id=532840878493&ns=1&abbucket=16	95 藏壶天下宜兴紫砂壶全手工功夫茶具茶壶原矿紫泥四方生肖龙马精神	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.108.WJTVFD&id=529754340530&rn=56b140375279307d910d06ed0b7978a0&abbucket=6
			96 宜兴原矿紫砂壶 全手工花器紫砂壶高工倪新安 金蟾茶壶茶具	800~	https://detail.tmall.com/item.htm?spm=a1z10.3-b-s.w4011-14745408145.104.WJTVFD&id=42164573063&rn=56b140375279307d910d06ed0b7978a0&abbucket=6

致谢

本论文历经数月，从题目的选择、框架的设置到分析方法，从毫无头绪到扎实地利用理论与逻辑对于议题分析，这一个过程着实艰难但充实，此篇论文是由李杰老师的悉心指导下完成的，李杰老师为经验丰富的教师，在整个论文框架和论文写作过程中，对我进行了耐心的指导、帮助和提出严格要求，非常感谢李杰老师不断引导我开阔思路，感谢李杰老师多次周末都抽空与我会谈，为我答疑解惑，在这一段艰辛但宝贵的论文撰写的过程中，短时间增长了大量的知识、且开阔了视野，对于心智的锻炼肯定对我往后的人生有所帮助。

在如此短的时间丰富了各领域的知识，很大一部分都是得益于朋友的帮助，「每一个朋友，都是一本书」在此感谢那些曾经被我追着问，并且一起共同讨论问题的同学们，感谢远在江西景德镇的朋友们，协助我做调研问卷，陪我讨论整体陶瓷工艺的流程，也科普了我很多对于工艺上的知识，也感谢父母这段期间时常打视讯来对我的加油打气，最感谢远在日本的女朋友玮婷，陪我历经这个富有挑战性的过程，予以我很大的空间并且不断地鼓励我。

最后再一次向曾经帮助过我的所有人，献上最诚挚的谢意！

北京大学学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明：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。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文作者签名： 日期： 年 月 日

学位论文使用授权说明

（必须装订在提交学校图书馆的印刷本）

本人完全了解北京大学关于收集、保存、使用学位论文的规定，即：

- 按照学校要求提交学位论文的印刷本和电子版本；
- 学校有权保存学位论文的印刷本和电子版，并提供目录检索与阅览服务，在校园网上提供服务；
- 学校可以采用影印、缩印、数字化或其它复制手段保存论文；
- 因某种特殊原因需要延迟发布学位论文电子版，授权学校☐一年/☐两年/☐三年以后，在校园网上全文发布。

（保密论文在解密后遵守此规定）

论文作者签名： 导师签名：

日期： 年 月 日