# 大数据可视化平台设计说明书

目录

[一、 背景 3](#_Toc521767873)

[1.1开发大数据平台有利因素 3](#_Toc521767874)

[1.2当前旅游行业的挑战 4](#_Toc521767875)

[二、 需求分析 5](#_Toc521767876)

[三、 开发环境 5](#_Toc521767877)

[3.1软件环境 5](#_Toc521767878)

[3.2硬件环境 6](#_Toc521767879)

[四、 平台设计 6](#_Toc521767880)

[五、 算法设计及描述 7](#_Toc521767881)

[5.1聚类算法 7](#_Toc521767882)

[5.2关键词提取分析算法 8](#_Toc521767883)

[5.3机器学习算法 8](#_Toc521767884)

[六、 数据库设计 9](#_Toc521767885)

[6.1用户信息表（tb\_userInfo） 9](#_Toc521767886)

[6.2管理员和代理信息表（tb\_managerAndAgent） 10](#_Toc521767887)

[6.3地理位置信息表（tb\_address） 10](#_Toc521767888)

[6.4产品信息表（tb\_produce） 10](#_Toc521767889)

[6.5出行方式表（tb\_ways） 11](#_Toc521767890)

[6.6用户访问以及购买记录表（tb\_accessRecord） 11](#_Toc521767891)

[七、 功能设计 12](#_Toc521767892)

[7.1系统功能图 13](#_Toc521767893)

[7.2系统功能列表 13](#_Toc521767894)

[7.7模块功能详细设计 15](#_Toc521767895)

# 背景

1.1开发大数据平台有利因素

（一）“旅游+”战略

2015年8月，国家旅游局研究部署实施“旅游+”战略，通过智慧旅游、乡村旅游、工业旅游、商务旅游、研学旅游、医疗旅游、养老旅游、健康旅游等领域，重点推进“旅游+”融合发展。“旅游+”具有天然的开放性、动态性，“+”的对象、内容、方式都能不断拓展丰富，融入各种文化创意，“+”的速度越快，“+”的内容越多，能够满足游客不同的消费需求，旅游业的发展就能够越快。

（二）全域旅游战略

2015年8月19日，全国旅游工作研讨班上，国家旅游局局长李金早提出推进全域旅游发展。随后，国家旅游局发布《关于开展“国家全域旅游示范区”创建工作的通知》，并于2016年1月全国旅游工作会议上全面提出从景点旅游走向全域旅游，开创中国“十三五”旅游发展新局面，将推进全域旅游作为新时期的旅游发展战略。景区旅游作为旅游产业链的核心，是全域旅游的一部分，全域旅游能缓解旅游旺季时景区最大承载量和旅游人数过多之间的矛盾，给景区带来正面的效应，但同时也对景区提出了更高的要求，景区必须解决产品单一、景区资源的文化内涵挖掘不深，特色吸引力不足等问题。

（三）“一带一路”带来的机遇

随着“一带一路”战略的实施推进，中国与周边国家的交通运输等基础建设的建成，中国与“一带一路”沿路国家的交流就会越来越方便，更低的交通成本与更多的文化交流，必定能吸引更多境外游客，给国内旅游业带来更多的人流量，有利于旅游的全产业链发展，提升景区总营业收入。

（四）“互联网+”及大数据的发展

现在人们更热衷于自驾游或是自由行，借助“互联网+”，人们可以更方便地在网上获取更多的信息，从酒店，出行方式到行程的规划，都可以按照自己的喜好在网上安排好，可以最大程度地满足自己个性化的需求。在线旅游交易的快速发展，同时也可以激活更多的旅游投资与消费热点领域，给游客提供更多的选择，提高旅游人均消费水平。

大数据的发展，可以更准确地了解客源、开发客源、锁定客源。了解和关注游客需求，最大程度开发和满足游客的潜在需求，重视游客的感受，增加和延长价值链，增强游客的旅游体验，提高旅游服务品质，提升游客的满意度和美誉度，增加游客对景区的粘度。

1.2当前旅游行业的挑战

（一）过度依赖门票经济，景区门票乱涨价现象严重

门票车票保险捆绑销售，学生票只在暑假销售，全价票40元半价票变30元，几个景点打包销售等现象多有发生。在景区的营业收入中，门票收入占了很重的比例，多数5A级景区门票收入占比超过40%，其中最高的超过80%，目前国内景区还无法摆脱“门票经济”的盈利模式，然而国外一些著名景点却通常不要门票，或将门票定价很低，以吸引众多游客前来参观。如意大利著名的古罗马斗兽场只需6欧元；日本自然景观不收门票，文化景点收费很低；法国旅游景点的门票价格基本在10欧元以下。与国外相比，国内景点票价过高的门票会降低游客在国内旅游的热情，降低国内旅游行业在国际中的竞争力。而且过度依赖门票收入，有可能就会忽视旅游服务质量，不注重第三产业链的延伸和挖掘，让游客找不到其他满意的旅游消费。

（二）出境旅游对国内旅游的冲击

随着中国与世界的交流越来越多，人们走出国门到世界其他地方旅游也越来越方便。中国是全球增长最快的客源输出国之一，2014年，中国大陆公民出境旅游人数突破1亿人次，达到了1.09亿人次，2015年上升为1.2亿人次，比1992年的298.87万人次增长了39倍。目前，中国公民出境旅游目的地已扩大到151个国家和地区，成为世界重要的旅游客源国。这就对国内旅游业造成了一定的冲击，分走了相当多一部分的人流量。要提高国内景区的竞争力，就必须要提高景区的服务质量与多样化。降低门票的票价，吸引更多的人流量，以“旅游+”提供更多的增值服务，开拓更多的收入增长点，同时也能更好的满足游客的需求。

（三）国内景区对自由行、自驾游等旅游方式不够重视

随着80、90后消费力的提高，他们开始成为旅游消费的主群体，游客群体年轻化成为趋势，而年轻人对旅游多样化的需求，使得自由行或私人订制游渐入主流。数据显示，2015年自由行市场已经赶超跟团游。然而大多数的景区运营方式和理念却没有大的改变，依然倚重于与旅行社合作的方式，给予散客的门票打折力度不大，在散客如何到达景区的出行指导及便捷性，景区接待等方面依然不够重视，没有足够重视散客的感受，满足游客的潜在需求，增强散客的旅游体验。

# 需求分析

旅游数据可视化信息平台中的主要功能是针对获取的旅游统计数据，对不同类型的数据进行统计、分析与展示。本文研究的主要数据来源包括用户的浏览记录、用户的购买记录、产品信息等在平台中会产生的数据类型。具体说来，旅游数据信息平台中各用户的主要业务有：

管理人员：主要指的是平台的管理人员。它是可视化系统的管理者，主要是通过可视化平台，查看某产品综合统计数据、游客大数据分析、查看各市级代理销售情况等功能。

用户：主要是通过可视化平台，搜索各条路线的实时销量信息、查看区域热力图等功能。

代理：主要是通过可视化平台，搜索各条路线的实时销量信息、查看区域热力图、获取产品关键词、本地用户大数据分析等功能。

根据以上对可视化系统用户的分析，可以得出大数据可视化平台所需要完成的主要功能。

# 开发环境

3.1软件环境

操作系统：Windows10；

界面开发工具：eclipse；

可视化组件 Highcharts；

缓存数据库：Redis；

大数据平台：Hadoop

3.2硬件环境

三台1核1G内存的Ubuntu云服务器，两台1核1G内存的Cent OS云服务器，开发使用两台双核8G内存PC机。

# 平台设计

大数据可视化平台的总体架构是基于分层思想来实现的，即将整个系统的功能进行拆分，每一层只需要实现对应的功能即可。对于整个系统来说各个层之间相互独立，层与层之间通过接口相互调用，这样分层的目的是为了实现低耦合高内聚的系统框架，本系统在设计上遵循了传统的三层架构模式，这样可以使软件的设计和开发更加标准化，同时方便了系统的管理和维护。

本系统采用的是经典的 B/S 架构，由数据访问层、业务逻辑层和用户表现层构成。其中用户表现层就是在浏览器端看到的内容，实现可视化的功能，它与业务逻辑层通过 HTTP 协议进行数据的交换与传输，而业务逻辑层则根据业务需要对数据层进行访问，这种分层结构使得系统相对灵活。对于整个系统来说业务逻辑层是核心，它通过解析用户通过浏览器发送的请求对数据进行分析并根据用户的需求对数据层进行相关操作，然后将数据层返回的数据进行解析后以网页的形式展示，

本系统的总体架构如图 4-1 所示：

图 4-1 平台总体架构

数据层主要负责对数据进行存储与管理，首先通过关系型数据库Mysql将旅游统计的相关数据进行存储和整理，同时针对数据量较大的情况，系统中使用 Hadoop框架实现对数据的处理，从而能够高效的完成对大数据进行分析和信息挖掘任务，通过开源工具Sqoop 在 Mysql和 Hadoop 进行数据的传递。

业务逻辑层在用户层与数据层之间，是页面可视化与数据库之间的一座桥梁，有数据管理、统计分析和可视化三个模块。

用户表现层就是展现在用户面前的页面，是用户与系统的交互接口，一方面向用户展示系统所包含的功能，另一方面又接收用户输入的请求并向下调用。

# 算法设计及描述

5.1聚类算法

K 均值（K-means）聚类算法是一种无监督的实时聚类算法。自从该算法提出之后，就被广泛地应用在不同的学科领域，很多研究人员基于各自的领域发展出大量的改进算法。尽管 K-means 聚类算法自提出以来已有近 60 年，由于其算法比较简单，容易操作且运行高效，目前仍然是较流行的经典聚类算法。

K-means 算法是聚类算法中一种典型的划分方法。其基本思想是把数据对象经过反复迭代划分到不同类中，使得同一类的数据对象间相似度很大，不同类间的数据对象相似度很小。该算法原理比较简单，通常采用距离作为相似性的评价指标，如果两个数据对象间的距离越近，就说明两数据对象的相似度越大。K-Means 聚类算法大体可以分为三个步骤：第一步是从 N 个数据中初始随机选择 K 个对象作为聚类中心；第二步是计算每个样本到这 K 个聚类中心的距离，按照最邻近原则，找出最小距离把样本归入到离它最近的聚类中去；第三步是计算每个聚类中所有样本的平均值，并将该平均值作为新的聚类中心；（4）对第二步和第三步反复迭代，直到聚类中心不再进行大范围移动或者移动距离小于某个给定的阈值为止。这样得到的各个类是独立的，且尽可能的不同。

5.2关键词提取分析算法

根据显示性偏好理论, 消费者在一定价格条件下的购买行为暴露了或显示了他内在的偏好。对用户的购买产品和搜索记录进行关键词分析算法，提取用户特征，以及产品卖点。对用户行为的关键词提取问题视作二分类问题处理。具体来说，首先对训练集进行标注， 即将关键词标注为正样本和将非关键词标注为负样本，然后根据正、负样本的特征集学习一个分类函数，最后用学习得到的分类函数去判别新的候选关键词是否为关键词。最后根据关键词对行为相似的用户做精准推送。

5.3机器学习算法

机器学习作为人类智力的延伸 ，作为人工智能的重要研究方向之一 ，它试图 从模拟人类的学习能力出发 ，运用一些最基本的统计方法 ，去探索客观世界，获得 各种知识和技能 ，在汁算机技术的帮助下建立相关的学习模型 ，最终可以让计算 机系统获得某些学习能力 。

支持向量机（ Support Vector Machine ， SVM ） 是由 Vapnik 于 1995 年根据统计学 理论 ，提出的一种新的机器学习方法 ，它以结构风险化最小原则为基础 ，其主要思 想是建立一个分类超平面作为决策曲面 ，使得正例和反例之间的隔离边缘最大 化。它在许多诸如车牌识别 、文本分类等实际应用中体现了其大有可为之处。另 外，支持向量机还有一个特点 ，就是可以很好地应用于高维数据 ，避免了“ 维数灾难”问题。

对游客的浏览、购买、评价记录做情感分析，将使用支持向量机的机器学习方法。首先对评论内容做预处理，需要进行中文分词 ，将由汉字序列组成的评价语句 ，通过一定的方法分 割成若干个有着单独意义的汉语词条，建立一个旅游情感词典，将经过预处理的文本数据需要转化成某种结构化的形式，让计算机“ 读懂” ，进而可以进行识别和处理。点评文本转化为向盘空间模型后 ，可以得到一个稀疏矩阵 ，通常是一个高维的空间，在机器学习的过程中，过高的特征空间维度，有造成“维数灾难”的可能。虽然理论和实践证明了采用支捋向量机可以很好地应用于高维数据，越免了“维数灾难”问题，但如果考虑算法的时间复杂度，最好还是把特征的个数控制在一个合理的范围内。这就要求采用某种特征选择算法对特征空间进行筛选，从而达到降低维数的目的。已经研究出多种方法可以进行特征选择，最常用的两种方法是：文档频率（OF）和卡方（CHI）统计。文档频率是一种简单的，但有较好性能的特征选择方法，该方法通过将文档频率设置在某个范围来进行特征的选取。卡方统计是通过分析特征和类别之间的依赖程度来进行特征的选取。

# 数据库设计

6.1用户信息表（tb\_userInfo）

记录用户旅客的基本信息。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 长度 | 是否可为空 | 简要说明 |
| id | int | 12 | 否 | 用户旅客id |
| username | varchar | 50 | 否 | 用户名 |
| name | varchar | 30 |  | 姓名 |
| phone | varchar | 20 |  | 电话号码 |
| addressId | int | 12 |  | 用户地址 |
| register\_time | datetime | 30 | 否 | 注册时间 |
| last\_time | datetime | 30 |  | 用户最近访问时间 |
| sex | int | 1 |  | 0男，1女 |
| age | int | 3 |  | 年龄 |
| idCard | varchar | 30 |  | 身份证信息 |

6.2代理信息表（tb\_agent）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 长度 | 是否可为空 | 简要说明 |
| id | int | 12 | 否 | 代理旅客id |
| username | varchar | 50 | 否 | 代理用户名 |
| name | varchar | 30 |  | 代理姓名 |
| phone | varchar | 20 |  | 电话号码 |
| addressId | int | 12 |  | 用户地址 |
| register\_time | datetime | 30 | 否 | 注册时间 |
| last\_time | datetime | 30 |  | 用户最近访问时间 |
| sex | int | 1 |  | 0男，1女 |
| age | int | 3 |  | 年龄 |
| idCard | varchar | 30 |  | 身份证信息 |
| Place | Varchar | 30 |  | 代理地区 |

6.3管理员和代理信息表（tb\_manager）

记录管理员和代理的基本信息，不与用户一个表是因为担心用户过多，导致查询效果不好。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 长度 | 是否可为空 | 简要说明 |
| id | int | 12 | 否 | id |
| username | varchar | 50 | 否 | 用户名 |
| role | int | 1 | 否 | 0表示管理员，1表示用户 |
| agentName | varchar | 100 |  | 代理处名称 |
| name | varchar | 30 |  | 姓名 |
| profession | varchar | 50 |  | 职称 |
| phone | varchar | 20 |  | 电话号码 |
| addressId | int | 12 |  | 用户地址 |
| register\_time | datetime | 30 | 否 | 注册时间 |
| last\_time | datetime | 30 |  | 用户最近访问时间 |
| idCard | varchar | 30 |  | 身份证号码 |
| produceId | Int | 12 |  | 产品信息id |

6.4地理位置信息表（tb\_address）

记录用户、管理员、代理的地理位置。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 长度 | 是否可为空 | 简要说明 |
| addressId | int | 12 | 否 | Id |
| country | varchar | 50 | 否 | 国家 |
| province | varchar | 50 |  | 所在省 |
| city | varchar | 50 |  | 所在市 |
| deati | varchar | 200 |  | 详细地址 |

6.5产品信息表（tb\_produce）

记录旅游路线信息。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 长度 | 是否可为空 | 简要说明 |
| produceId | int | 12 | 否 | 旅游路线id |
| title | varchar | 100 | 否 | 旅游路线主题名称 |
| scenics | varchar | 500 | 否 | 路线所包含所有景点，用“青秀山，大明山”这种格式 |
| waySId | varchar | 60 | 否 | 出行方式，中间用逗号隔开 |
| totalAmount | double | 255 | 否 | 路线原价，单位元 |
| price | double | 255 | 否 | 路线现价 |
| Images | varchar | 255 |  | 路线的特色图片，多张图片用逗号隔开 |
| description | varchar | 1000 |  | 对路线的文字描述 |
| start\_time | datetime | 30 |  | 路线运营的初始时间 |
| End\_time | datetime | 30 |  | 路线运营的结束时间 |

6.6出行方式表（tb\_ways）

记录从起始地点到目的地的出现方式。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 长度 | 是否可为空 | 简要说明 |
| id | int | 12 | 否 | 出行方式id |
| startPlace | varchar | 100 | 否 | 起始地 |
| endPlace | varchar | 100 | 否 | 目的地 |
| amount | double | 255 | 否 | 花费金额 |
| way | varchar | 30 | 否 | 出行方式，如飞机，0飞机，1火车，2汽车，3自驾，4其它 |

6.7用户访问以及购买记录表（tb\_accessRecord）

记录每条路线的访问的记录。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 字段类型 | 长度 | 是否可为空 | 简要说明 |
| id | int | 12 | 否 | 记录id |
| userId | int | 12 |  | 为空，表示游览用户非平台注册用户 |
| produceId | int | 12 |  | 产品id，空表示对平台的访问，非空表示对路线的访问 |
| createTime | datetime | 30 | 否 | 记录生成时间 |
| buy\_tool | int | 2 |  | 0表示访问记录，1表示通过携程购买，2表示省油灯购买，3表示在代理处购买等 |
| agentId | int | 12 |  | 空，表示非代理处购买，否则在代理处购买并可以找到此次购买的代理人 |
| buy\_way | int | 2 |  | 0表示现金支付，1微信支付，2支付宝支付 |
| buy\_time | datetime | 12 |  | 购买时间，为空表示访问时间 |
| buy\_price | double | 30 |  | 所花金额，单位元 |
| count | int | 3 |  | 买票的张数 |
| produce\_id | int | 12 |  | 产品id |

# 功能设计

7.1系统功能图

图7-1 系统功能图

7.2系统功能列表

表7- 1查询系统功能列表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 产 品 名 称 | | 广西文化旅游 | 版本号 | V1.0 |
| **软件功能项目** | | | **功能说明** | |
| 查询各产品访问量走势 | 游  客  操  作 | 日访问量走势 | 提供用户该产品每日访问量走势，并给出每日所有产品总访问量的排名 | |
| 月访问量走势 | 提供用户该产品12个月内访问量走势，并给出每月中所有产品总访问量的排名 | |
| 年访问量走势 | 提供用户该产品每年内访问量走势，并给出每年内所有产品总访问量的排名 | |
| 当前平台总访问量 | 该产品访问量 | 提供给用户本平台内该产品的历史总访问量 | |
| 客源地势力分析 | 该产品访问游客来源 | 提供用户本产品的主要游客来源 | |
| 全国主要城市游客来源 | 提供用户全国范围内主要的游客来源 | |
| 游客信息统计 | 男女比例分析 | 显示访问本平台游客男女比例分析 | |
| 年龄段统计分析 | 显示访问本平台游客年龄比例分析 | |
| 消费水平 | 显示访问本平台的游客不同消费水平所占比例分析 | |
| 出行方式分析 | 显示访问本平台的游客选择的出行方式比例分析 | |
| 各市级代理销售情况 | 各市级代理销售情况分析 | 统计各个市级代理在各平台上的销售情况 | |

表7- 2专家系统功能列表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 产 品 名 称 | | 广西文化旅游 | 版本号 | V1.0 |
| **软件功能项目** | | | **功能说明** | |
| 产  品  销  售  分  析 | 用  户  操  作 | 查询产品销量 | 提供各产品销售量查询功能 | |
| 销售途径统计 | 提供各产品官网直销统计和OTA分销统计功能 | |
| 关键词分析 | 提供关键词分析并向目标人物推荐产品的功能。 | |
| 游  客  预  测 | 情感分析 | 提供对游客进行情感分析的功能 | |
| 年龄预测 | 提供对游客进行年龄分析的功能 | |
| 职业预测 | 提供对游客进行职业分析的功能 | |
| 性别预测 | 提供对游客进行性别预测的功能 | |
| 地  区  销  售  分  析 | 查看各地区销售 | 提供查看各地区代理销售情况的功能 | |
| 销售情况分析 | 提供销售情况分析的功能 | |

7.7模块功能详细设计

|  |  |
| --- | --- |
| 模块编号：001 | 模块名称：查询各产品访问量走势 |
| 模块描述：游客在选择查询本产品的访问量走势后，会以折线图的形式显示出该产品本月以来每日的日访问量，当年内所有月份的月访问量以及近几年来的年访问量，并且还提供给用户每日的产品访问量排名、每月的产品访问量排名以及每年的产品访问量排名。 | |
| 界面描述：  在产品界面选择访问量走势功能，以下图7-1、图7-2、图7-7的折线图的形式显示出日访问量、月访问量以及年访问量的走势。  图7-1  图7-2  图7-7 | |

|  |  |
| --- | --- |
| 模块编号：002 | 模块名称：当前平台总访问量 |
| 模块描述：提供给用户本平台内该产品的历史总访问量，每当游客浏览或者点击该产品，则访问量加一。 | |
| 界面描述：在每款商品的下方显示该产品的总访问量。 | |

|  |  |
| --- | --- |
| 模块编号：007 | 模块名称：客源地势力分析 |
| 模块描述：提供用户本产品的主要游客来源以及全国范围内游客的主要来源地。 | |
| 界面描述：  以柱状图的方式呈现出各产品的客流来源，如图7-4 。以热力图的方式呈现全国游客主要来源，如图7-5所示。  图7-4  E:\qq\631035080\FileRecv\Y8AJN8QC[~HOH]2O1$L8X%X.png  图7-5 | |
| 模块编号：004 | 模块名称：游客年龄段统计 |
| 模块描述：显示访问本平台游客男女比例、年龄比例、不同消费水平所占比例以及游客选择的出行方式比例，以供用户进行分析。 | |
| 界面描述：  从产品展示界面选择男女比例功能后，会以饼状图的方式展示游客男女比例分析，如图7-6所示：  C:\Users\Administrator\AppData\Roaming\Tencent\Users\631035080\QQ\WinTemp\RichOle\}Q53U3H8P[1RT9Z%M$U)3[M.png  图7-6  从产品展示界面选择年龄比例功能后，会以饼状图的方式展示游客年龄比例分析，如图7-7所示：  C:\Users\Administrator\AppData\Roaming\Tencent\Users\631035080\QQ\WinTemp\RichOle\AB}(@PRQEMJOAOPUA$R6JRX.png  图7-7  从产品展示界面选择消费水平比例功能后，会以柱状图的方式展示游客消费水平比例分析，如图7-8所示：  图7-8  从产品展示界面选择游客出行方式功能后，会以雷达图的方式展示游客出行方式比例分析，如图7-9所示：  图7-9 | |

|  |  |
| --- | --- |
| 模块编号：005 | 模块名称：各市级代理销售情况 |
| 模块描述：提供给用户各市级和各平台本产品的销售情况。 | |
| 界面描述：从产品展示界面选择各市级代理销售情况功能后，统计各个市级代理在美团、携程、去哪儿等各平台的销售情况，如图7-10所示。  图7-10 | |

|  |  |
| --- | --- |
| 模块编号：006 | 模块名称：产品销售分析 |
| 模块描述： 专家型用户可以点击产品销售分析可以查看各种产品的销量情况，点击产品销售统计可以查看各产品官网直销统计和OTA分销统计，点击关键词分析可以根据游客查询的关键词进行分析并推荐相关包含该关键字的产品。 | |
| 界面描述：  在产品销售分析模块点击产品销售分析查看各种产品的销量情况，如图7-11所示。  WVIK0))4SH(KWARK((B88`C  图7-11  点击产品销售统计可以查看各产品的官网直销统计和OTA分销统计，如图7-12、图7-17所示。  3TZVQXHLDDUD7M8[MWZ1L`L  图7-12  @DB)X_IO78Y`89R(0H(P@}Q  图7-17  点击关键词分析可以根据游客查询的关键词进行分析并推荐相关包含该关键字的产品，如图7-14所示。  (_B%`~5)I~3A2AP9MA0C~ZJ  图7-14 | |

|  |  |
| --- | --- |
| 模块编号：007 | 模块名称：游客预测 |
| 模块描述： 专家型用户可以点击情感分析查看游客的浏览购买评价等情况，分别点击年龄分析、职业分析、性别预测可以查看游客的年龄分布、职业分布、性别分布情况。 | |
| 界面描述：  用户在游客预测模块点击情感分析后，显示情感分析雷达图如图7-15所示  F)}0Y0J502_89{Q(ZIL%1}1    图7-15  点击年龄预测，显示年龄预测分布图，如图7-16所示。  3C60DN027J0HLZ@`E~E{WZ6  图7-16  点击年龄预测，显示职业预测分布图，如图7-17所示。  S}8D9ZY0XCS@MKL~B]R13A2  图7-17  点击性别预测，显示性别预测分布图，如图7-18所示。  V{%H7505)7F]V]C7XV[8}6X  图7-18 | |

|  |  |
| --- | --- |
| 模块编号：008 | 模块名称：地区销售分析 |
| 模块描述：提供给商家用户各地区和各平台内该产品销售情况，便于商户进行销售分析。 | |
| 界面描述：从产品展示界面选择各市级代理销售情况功能后，统计各个市级代理在美团、携程、去哪儿等各平台的销售情况，便于商户进行销售分析，如图7-19所示。  图7-19 | |