



Viiision

APP评论数据分析系统

方案概要



ComVision 团队

1. 需求分析与解决思路

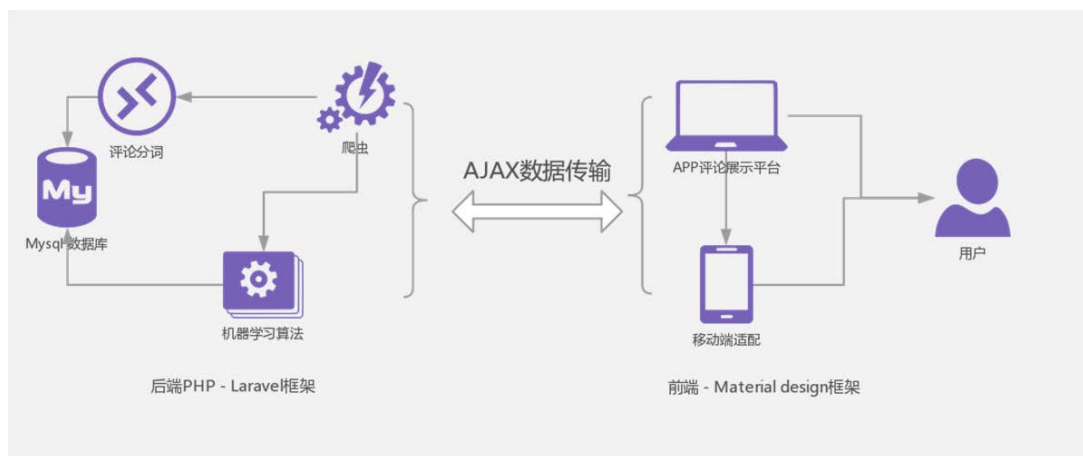
1.1. 需求分析

发包方虹软公司期望能有这样一个系统，在获得海量用户的反馈基础上，能够较为快速地分析出不同地区不同人员对于 APP 的使用反馈，让公司决策者和开发者能够多维度直观的了解 APP 的使用情况。

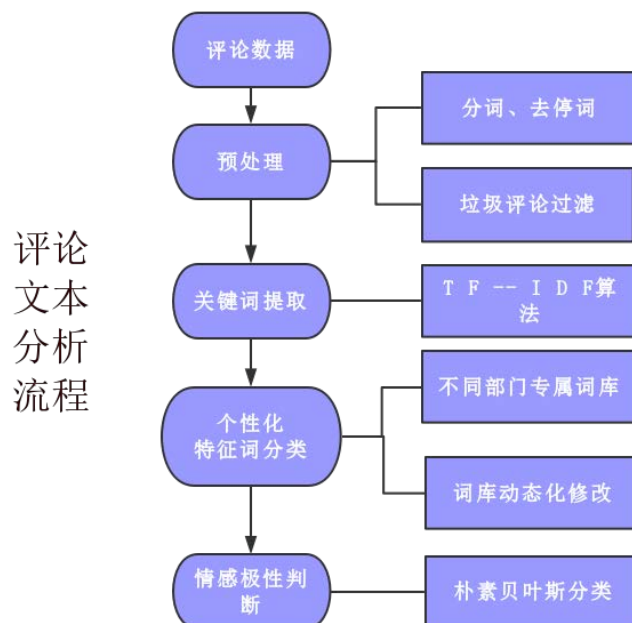
1.2. 解决思路

ComVision 团队致力于数据分析方面的软件研发，针对本次课题利用爬虫技术、数据可视化技术、分词与关键词提取技术、模式匹配、机器学习等方法协作完成一款 Viiision—APP 评论数据分析系统解决发包方需求。

本系统技术总框架如下：



评论数据分析流程图：



2. 功能简介

该系统通过对各大应用平台数据的实时抓取，并对评论进行过滤、分词、情感分析等处理，将统计分析后的数据以多样的图表形式呈现给用户，同时给予用户准确及时的反馈。

功能模块如下：



3. 特色综述

3.1. 全方位的评论数据获取和预处理

本系统利用爬虫技术，管理员只需输入 APP 名称并配置其发布的平台，即可实现评论数据的自动获取和保存。同时利用有监督学习的朴素贝叶斯二分类器机器学习算法，对实时抓取的评论进行数据预处理，将垃圾评论自动过滤以提高后续统计分析的准确性。

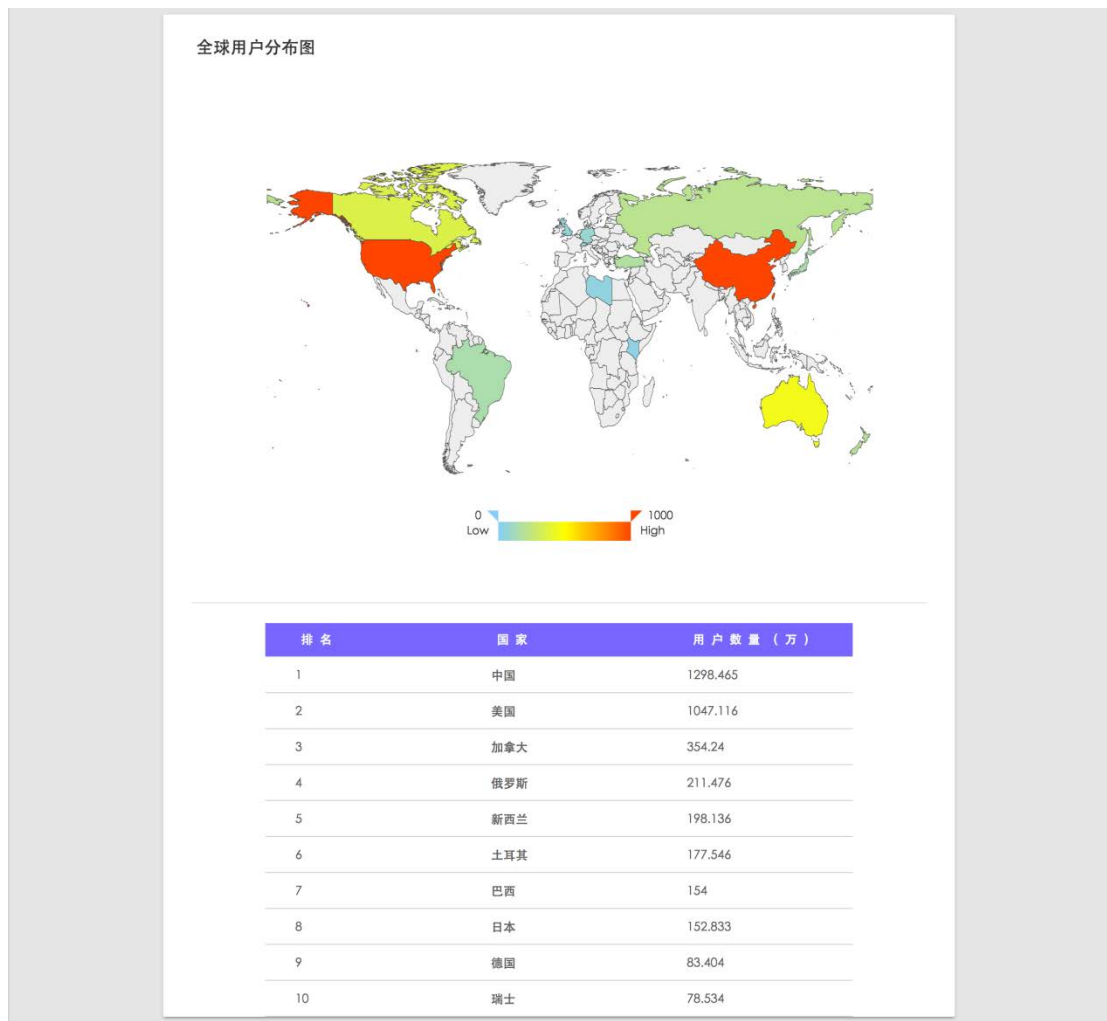
3.2. 直观的多维图表呈现

针对信息量大难以直观呈现的问题，我们通过 Echarts，将统计分析后的重要信息以多类型的图表形式呈现，且在同一图表中反馈多元化的内容，使用户能够直观了解更多的信息。

词云图如下：



区域分布图如下:



3.3. 动态化+个性化的分词管理

因不同职能部门需求不同，我们将分词分类，不同部门可根据所需查看不同类别的分词，例如运营部可以查看活动类的分词；同时用户可根据所需对时下热词进行选择添加，并在后续查询时优先显示。



3.4. 精确的产品特征情感分析

本系统利用关键词提取技术对评论中的产品特征进行提取，进而利用模式匹配和机器学习相结合的算法分析出各产品特征的用户态度，用户可查看各产品特征的好评差评详情及其百分比。

热词详情

画面 内存 版本 像素 广告 手机 垃圾

好评

- 1.太好了！就是有时候有东西要下，**内存**占的很大。
- 2.可以 如果**内存**占小些就OK了
- 3.**内存**有点大，以后更新的时候，**内存**小一点 不过还是力推的
- 4.好玩，就是**内存**大了一点（少给你一颗怕你骄傲啦）
- 5.这**内存**。。好那门
- 6.刚开始挺喜欢的后来图片就不存到手机上了。不是**内存**的问题。连续卸载下载了好几次。又好了。不太稳定。。不过还是很喜欢。用了半年了。
- 7.**内存**比较大，手机比较卡，不过这个软件挺好的，所以给个好评！
- 8.**内存**大
- 9.还好啦，就是有些占地儿，然后突然有一天图纸都不见了，还是占着**内存**所以就删了从下个图就不给五星了？
- 10.特别喜欢，之前下过，但是占**内存**，所以给了四星，不怎么卡
- 11.不太好，**内存**大
- 12.用了一段时间了，软件很实用，但也有很多缺点，比如闪退、卡，一些特效需要下载这倒不算什么，但特别占**内存**，然后今天我是因为这个软件更新到一半怎么也不更新了，怎么试都不行，还不能卸载了重下，就是有个正在下载的图标在那里又不能删，特别恶心
- 13.就是有点占**内存**，耗电快
- 14.好好好好好好，就是**内存**占的大
- 15.你们每次玩这个游戏的原因是游戏的**内存**太大了 点个赞！
- 16.一直觉得挺好玩，**内存**也不是很大，平时照相自拍也一直用，闺蜜上次问我这是什么东西真好玩，我就推荐给她，她也很喜欢。最重要的是faceu只要下载就不用联网了，随时随地图拍美美哒，还是我最喜欢的地方

差评

- 1.不好，有些都删不了，占**内存**
- 2.以前使用时觉得是拍照神器，可自从更新新版本后，手机就出现闪退问题，卸载后重新下载还是出现闪退，期间手机**内存**都是够用的，希望工程师上人们能够帮我解决闪退问题。
- 3.根本就拍不了，只会占**内存**。
- 4.占**内存**太大，有点卡，不过功能很好很齐全。??
- 5.**内存内存**。。。
- 6.根本找不到 我保存下来的视频，不知道咋回事~ 难道是**内存**不够？
- 7.越来越占**内存** 不知道开发者怎么想的
- 8.太占**内存**了！！！！
- 9.卡，但是用别的**内存**比这还大就是一点不卡
- 10.卡""了不能用啊，卸载了占**内存**
- 11.真的好卡，之前用不会啊，手机**内存**还剩2G，就是卡
- 12.一卡一卡的，我手机**内存**又不是没有，好气
- 13.一般般功能太少**内存**太大
- 14.挺好的，就是占**内存**
- 15.大家别看好，其实一点也不好，浪费流量，站空间，本来我也喜欢这个软件的，可是站**内存**啊！我删了，可是吧，我星期日中午要去鞍山湖公园，我就又下载了，结果在人家看到站了1.79GB吓到我赶紧删了，并且还下载到一半的表情呢！但是还是想在星期日留几张美美的照片，就下载了，这是我个人的感觉，觉得还行，就是站空间，所以我只是把我的感觉说出来，如果你觉得就这一点**内存**的话就下载，如果觉得不好的，我即使告诉你一声站空间！谢谢！
- 16.好是好玩，只不过有点卡，有点占**内存**，还有一点耗电，希望有改进??
- 17.真的好卡的占**内存**很大，而且有些壁纸下载之后下载不了，还是可以再重新下载，但是真的好烦。
- 18.感觉不怎么样超超占**内存**，而且一用起来手机热的好像要爆炸了卡的不要不要的
- 19.还不错，表情太少，**内存**太大

3.5. 便捷的反馈信息获取

为能使企业及时纠正产品问题，把握市场商机的需求，本系统可根据用户所需，定时分析整理用户关注的内容直接推送给用户。

