

浙江农林大学

本科生毕业设计（论文）

开题报告（含文献综述）

（2019 届）

信息工程学院

题 目：在线商品用户评论分析系统设计与实现

学生姓名 陈怡 .

学 号 201505010323 .

专业班级 计算机科学与技术 153 班 .

学院名称 信息工程学院 .

指导教师 黄雷君 .

2018 年 12 月 28 日

目录

1	论文研究的背景.....	3
2	论文研究的意义.....	3
3	系统需求及可行性分析.....	3
3.1	网络购物现状.....	3
3.2	总体性能需求目标.....	4
3.3	功能需求.....	4
3.4	可行性研究.....	4
3.4.1	技术方面可行性.....	4
3.4.2	操作方面可行性.....	4
4	系统的详细设计.....	5
4.1	管理模块.....	6
4.2	图表显示模块.....	6
4.3	爬虫模块.....	6
4.4	分析模块.....	6
4.5	交互模块.....	6
5	预期成果.....	7
6	研究进度安排.....	7
7	参考文献.....	7

1 论文研究的背景

这是一个互联网的时代，随着网络的普及和电脑技术的发展，人们的生活越来越离不开互联网，其中很为明显的就是网络购物。随着电子商务的急剧增长，网络购物越来越受到欢迎，占据了零售的很大一部分。而类似“双 11”之类的活动的推出更是将网络购物推向了巅峰。现在的生活节奏越来越快，以往的慢节奏已经不大适应现在的人们了，显然网络购物这种快节奏的活动方式更适合现在的人们。

随着我国互联网普及率和人民生活水平的提高，网上购物已成为很多人愿意选择的购物方式。据调查显示，近几年来网购所占购物数额的比例逐年大幅度提升。而与此同时，电子商务平台上商品的评论数量也呈几何式上升，使得想要购买商品的顾客很难在纷繁的评论中提取到自己所需要的信息。即使能得到信息也过于片面，不能系统合理评测商品。本人站在用户的角度上开发本系统，目的是希望能降低顾客网络购物风险。本系统主要采用爬虫技术，文本情感分析中的一些算法和一些当前比较先进的开源框架和 AI 平台来实现，用户可参考系统给出的评测结果从而选择是否购买该产品。

2 论文研究的意义

目前是一个网络时代，数据是生活中必不可少的一个元素，在大数据横行的时代下，我们可以利用数据做些什么。网络购物已经深入我们生活中的方方面面，在我们想要购买一个不是立即需要的商品时，我们往往想到的是网络购物，例如淘宝、天猫、京东或者考拉，因为在网络上的商品常常会比实体商铺里的更加便宜也更加方便，我们可以足不出户就购买到物美价廉的商品。但是，这其中也存在许多问题，现在的网络监管还没有达到严格的程度，网络购物偶尔也会掉入陷阱，价廉的商品可能并不物美。而网购平台给出了一个解决方案，购买者可以在所购商品下添加评论，当商品质量上成，商家服务周到时，购买者会留下美好的夸赞之语，而对所购商品不满意时就会恰恰相反。这时水军进入了商家的视线，通过花费一定的成本，可以使自己的商品布满好评，掩盖差评的存在。在评论数量过多时，肉眼已经无法分辨商品是否存在不好之处，这时候我们可以利用爬虫技术和自然语言情感分析的技术来让机器为我们分析，既可以节省宝贵的人力时间，而且机器并不会被欺骗。

3 系统需求及可行性分析

3.1 网络购物现状

成年人工作繁忙，学生学业繁重，人们普遍没有多余的时间去实体店购物。而互联网的发展和网上购物平台的出现完美的解决了这个问题。人人都有手机，何时何地都可以在网上“逛街”，免去了时间和空间上的限制。同时网上购物平台的出现，商家也可免去了高额的店面租借费，那

么商品的价格自然也就可以低价销售出。这对于商家和用户都是有利的。然而网上购物平台上的商品质量良莠不齐，作为购买者也无法直观的观察商品的细节，不能第一时间发现商品的好坏，所以商品的评论是购买者是否购买商品的一大主要因素，购买者通过查看商品的评论来判断商品是否良好，服务是否友善。但是另一方面，由于评论数量过多，购买者无法通过短暂的浏览从中分析出正确的结论。

3.2 总体性能需求目标

本次设计目标开发基于 python3 的在线商品用户评论分析系统，采用 b/s 模式。以下是本设计要开发完成的几个需求目标：

1. 实用性：设计界面美观实用简洁，软件与用户交互性较强。
2. 稳定性：软件目标是实现用户通过网页界面可实现在线商品的评论分析。用 python 成熟框架编程，稳定性较高。
3. 友好性原则：操作简单是任何软件的开发的目標，操作简单并不意味着功能简单，将复杂的功能简单化会给用户留下很深的印象。软件在各个阶段给用户以良好的提示。能保证用户的使用体验。
4. 安全性原则：每个使用者无需经过任何的身份验证。各项连接出错要有良好的出错提示。关闭服务端的信息泄露，保证服务器的安全。
5. 可扩展性原则：采用开放的接口和标准，便于系统向更大的规模和功能扩展。对业务代码服务化，交互和逻辑分离，为了以后更好拓展，更好维护。

3.3 功能需求

软件要求不断完善，功能也是逐步添加，以下为该在线商品用户评论分析系统的基本功能需求。

1. 使用者在搜索栏输入需要进行评论分析的商品链接，系统将商品链接返回后台处理。
2. 读取商品链接，如该商品已经进行过评论分析，则读取该商品的分析总结发送给使用者，否则通知爬虫引擎进行爬取。
3. 爬虫引擎读入商品链接，爬取商品评论，存入数据库，等待分析引擎处理。
4. 分析引擎读取商品评论，分析评论信息，生成商品分析总结，存入数据库并发送给使用者。

3.4 可行性研究

3.4.1 技术方面可行性

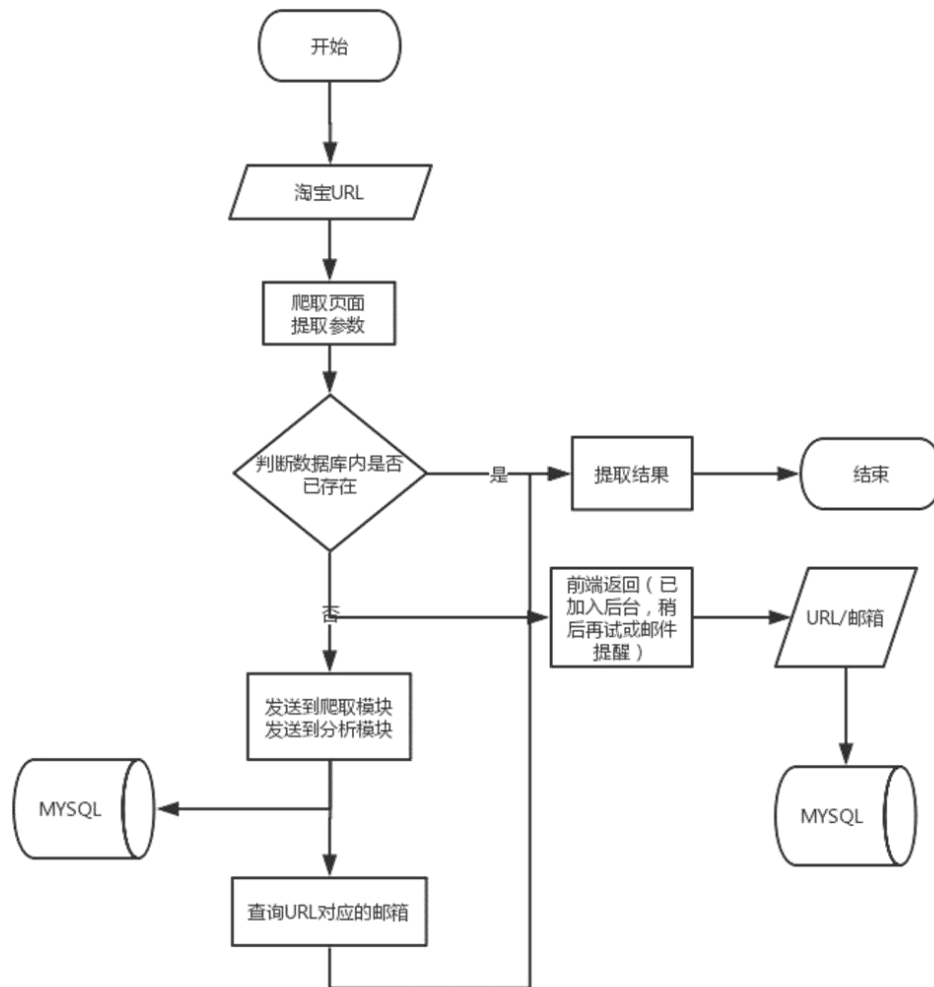
系统使用 django 作为开发基准框架。python 用途广泛，可以用于数据挖掘、数据分析、web 交互等。技术比较成熟，可行性高。

3.4.2 操作方面可行性

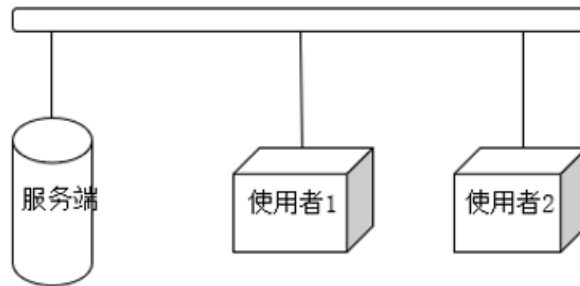
本软件采用网页的形式，pc 端可以通过网址访问。硬件环境采用云服务器，软件环境为 unix 操作系统，编程环境为 python，程序员均较为熟悉，操作可行。

4 系统的详细设计

程序流程图如下所示：



该软件为在线商品用户评论分析系统，总的系统规划基于传统的 b/s 模式。可以实现多用户同时在线，如下图所示：



系统主要模块包括管理模块、图表显示模块、爬虫模块、分析模块、交互模块。

4.1 管理模块

管理员可以登录管理员界面查看系统运行情况，进行系统常规设置等。

4.2 图表显示模块

读取分析结果，生成简单易懂的图表，为用户直观的呈现分析结果。

4.3 爬虫模块

将商品链接发送到爬虫引擎，爬取商品链接下的评论，存入数据库等待分析模块运行。适当使用 IP 代理池和反爬技术，设置合理的爬取速度防止 IP、账户被封。

4.4 分析模块

将爬取到的商品评论发送到分析引擎，进行评论处理和分析，去除无用信息，合并相似信息，分析评论信息，最后生成分析结论。

4.5 交互模块

用户通过网页来查看分析的结果，也可以通过邮件接收分析的结果。由用户输入商品链接，传输到服务器，服务器进行评论爬取、分析，并生成分析结果发送给用户。商品链接可一次性输入多条，使用分号（“;”）或回车分割。用户可以填写邮箱，将分析结果通过邮件发送，减少不必要的等待时间。

5 预期成果

实现在线商品评论分析系统的基本功能（用户操作界面、爬虫功能、分析功能），整个系统流畅完整的运行，不出现常规错误。具有用户操作友好的界面和良好的视觉效果。最后完成论文的书写。

6 研究进度安排

序号	毕业设计（论文）各阶段内容	时间安排	备注
1	文献查阅；外文翻译。	2018.12	
2	系统设计；开题报告。	2019.01-02	
3	系统实现、调试和测试。	2019.03-04	
4	论文撰写；答辩。	2019.05	

7 参考文献

- [1]于娟, 刘强. 主题网络爬虫研究综述[J]. 计算机工程与科学, 2015, 37 (02): 231-237.
- [2]魏韡, 向阳, 陈千. 中文文本情感分析综述[J]. 计算机应用, 2011, 31 (12): 3321-3323.
- [3]王洪伟, 郑丽娟, 尹裴, 史伟. 在线评论的情感极性分类研究综述[J]. 情报科学, 2012, 30 (08): 1263-1271+1276.
- [4]尹裴, 王洪伟. 面向产品特征的中文在线评论情感分类: 以本体建模为方法[J]. 系统管理学报, 2016, 25 (01): 103-114.
- [5]陈旻, 朱凡微, 吴明晖, 应晶. 观点挖掘综述[J]. 浙江大学学报(工学版), 2014, 48 (08): 1461-1472.
- [6]Kumar Ravi, Vadlamani Ravi. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications[J]. Knowledge-Based Systems, 2015, 89.

指导教师意见：

对课题的背景及现状进行了一定的阐述，研究的目的与意义表达清晰，研究内容明确，研究方案基本可行，进度安排较合理，同意开题。

指导教师签名：

年 月 日

开题小组评审意见：

开题小组组长签名：

年 月 日

学科意见：

开 题 报 告 答 辩 结 果： ☐ 通 过 ☐ 不 通 过

学科负责人签名：

年 月 日