摘 要

随着“互联网+大数据”深入发展，面对纪检监督工作的形势与不足，计划以研发模块、关联数据、完善流程为着力点和切入点，通过业务流程镶嵌信息系统提升纪检工作规范化，通过关联数据打破信息壁垒推动纪检工作高效化，通过研发模块监督权力运行推动纪检工作智能化。互联网信息时代的发展带动人们获取信息方式的改变，线下了解附近门店信息的方式已经无法满足用户需求，门店搜索推荐系统可以更加方便用户获取门店信息。本文主要介绍了门店搜索推荐系统的设计与实现。该系统的搜索功能是基于ElasticSearch搜索引擎，推荐功能是基于Spark框架，整个系统是基于Java的SpringBoot框架进行开发，使用MySQL数据库完成数据存储，前端结合了JavaScript、HTML、CSS等技术，应用了Metronic前端主题模板，系统采用InteliJ IDEA作为开发工具。

系统根据用户的地理位置信息等因素，为大众提供个性化的搜索推荐门店功能，主要包括客户端和后台管理系统。系统主要包括用户模块、搜索模块、推荐模块和后台管理模块。用户模块包括用户登录，注销，用户注册，查看首页功能；在搜索模块，用户可以根据门店名称关键词搜索并且可以对搜索结果进行排序，也可以根据搜索关键词的相关性搜索门店信息等；推荐模块应用了推荐算法完成推荐功能；后台管理模块包括商户管理，门店管理，品类管理功能等。

本系统的研发为用户提供了个性化的门店搜索和推荐，更加贴近用户的需求，为用户了解门店信息提供平台。

关键词：ElasticSearch；Spark；搜索；推荐；门店

# Abstract

The development of the Internet information era has led to changes in the way people obtain information. The offline way of learning information about nearby stores has been unable to meet user needs. The store search recommendation system can make it easier for users to obtain store information. This paper mainly introduces the design and implementation of the store search recommendation system. The search function of the system is based on the ElasticSearch search engine, and the recommendation function is based on the Spark framework. The entire system is developed based on the SpringBoot framework of Java. The MySQL database is used to complete the data storage. The front end combines JavaScript, HTML, CSS and other technologies. Front-end theme template, the system uses InteliJ IDEA as a development tool.

The system provides personalized search and recommendation store functions for the public according to factors such as the user's geographic location information, mainly including the client and background management system. The system mainly includes user module, search module, recommendation module and background management module. The user module includes user login, logout, user registration, and home page viewing functions; in the search module, users can search by store name and keyword and sort the search results, or search store information based on the correlation of search keywords; The module uses a recommendation algorithm to complete the recommendation function; the background management module includes merchant management, store management, category management functions, etc.

The development of this system provides users with personalized store search and recommendation, which is closer to the needs of users and provides a platform for users to understand store information.

Keywords: ElasticSearch; Spark; search; recommendation; store

# 目 录

[摘 要 I](#_Toc101728908)

[Abstract II](#_Toc101728909)

[目 录 I](#_Toc101728910)

[第1章 概 述 1](#_Toc101728911)

[1.1系统开发背景和意义 1](#_Toc101728912)

[1.2国内外研究现状 1](#_Toc101728913)

[1.2.1 信息获取类软件国内外研究现状 1](#_Toc101728914)

[1.2.1 推荐系统国内外研究现状 2](#_Toc101728915)

[1.3 本文组织架构 3](#_Toc101728916)

[第2章 开发工具和相关技术 4](#_Toc101728917)

[·2.1 基于Python的Web框架Django 4](#_Toc101728918)

[2.2 DRF框架https://cloud.tencent.com/developer/article/1493264 5](#_Toc101728919)

[2.3 Vue.js 6](#_Toc101728920)

[2.4 Element-UI 7](#_Toc101728921)

[2.5 Apache Echarts 7](#_Toc101728922)

[2.7 本章小结 9](#_Toc101728923)

[第3章 可行性研究及需求分析 10](#_Toc101728924)

[3.1 可行性研究 10](#_Toc101728925)

[3.1.1 社会可行性 10](#_Toc101728926)

[3.1.2 安全可行性 10](#_Toc101728927)

[3.1.3 法律可行性 11](#_Toc101728928)

[3.1.4 文化可行性 11](#_Toc101728929)

[3.1.5 环境可行性 11](#_Toc101728930)

[3.1.6 经济可行性 11](#_Toc101728931)

[3.2 系统的需求分析 12](#_Toc101728932)

[3.2.1 业务需求 12](#_Toc101728933)

[3.2.2 用户需求 14](#_Toc101728934)

[3.2.3 功能性需求 15](#_Toc101728935)

[3.2.4 非功能性需求 15](#_Toc101728936)

[3.3 本章小结 16](#_Toc101728937)

[第4章 系统总体设计 17](#_Toc101728938)

[4.1 系统功能模块设计 17](#_Toc101728939)

[4.2 系统数据库设计 18](#_Toc101728940)

[4.2.1 数据库概念结构设计 18](#_Toc101728941)

[4.2.2 数据库逻辑结构设计 19](#_Toc101728942)

[4.3 本章小结 22](#_Toc101728943)

[第5章 系统的详细设计与实现 23](#_Toc101728944)

[5.1 用户模块的详细设计与实现 23](#_Toc101728945)

[5.1.1通用的异常处理机制 23](#_Toc101728946)

[5.1.2 用户注册功能 24](#_Toc101728947)

[5.1.3 用户登录功能 24](#_Toc101728948)

[5.2 ELK机制同步数据的详细设计与实现 25](#_Toc101728949)

[5.3搜索模块的详细设计与实现 27](#_Toc101728950)

[5.3.1根据关键词、品类搜索功能 27](#_Toc101728951)

[5.3.2 根据排序、标签查看搜索结果功能 30](#_Toc101728952)

[5.3.3 根据同义词扩展查看搜索结果功能 31](#_Toc101728953)

[5.3.4 根据相关性查看搜索结果功能 32](#_Toc101728954)

[5.4 推荐模块的详细设计与实现 33](#_Toc101728955)

[5.4.1 ALS算法训练实现 33](#_Toc101728956)

[5.4.2 LR算法训练实现 35](#_Toc101728957)

[5.4.3 推荐接入的实现 36](#_Toc101728958)

[5.5 后台管理模块的详细设计与实现 37](#_Toc101728959)

[5.5.1 品类管理功能 37](#_Toc101728960)

[5.5.2 商户管理功能 38](#_Toc101728961)

[5.5.3 门店管理功能 38](#_Toc101728962)

[5.6 本章小结 38](#_Toc101728963)

[第6章 系统测试 39](#_Toc101728964)

[6.1 测试方法 39](#_Toc101728965)

[6.2 系统测试用例 39](#_Toc101728966)

[6.3 测试分析总结 42](#_Toc101728967)

[6.3.1 系统的缺陷报告分析 42](#_Toc101728968)

[6.3.2 系统的测试覆盖分析 43](#_Toc101728969)

[6.3.3 系统存在的问题与建议 43](#_Toc101728970)

[6.4 本章小结 43](#_Toc101728971)

[结 论 45](#_Toc101728972)

[参考文献 46](#_Toc101728973)

[致 谢 48](#_Toc101728974)

# 第1章 概 述

## 1.1系统开发背景和意义

传统的纪检工作主要通过真人发送通知，下发文件和统计纪检工作结果等等。通过这种传统的纪检方法，在应对小型组织、企业和公司来说不失为一种极具效率的方法。通过少量的人力即可实现对于所有组织成员或员工的纪检工作管理。

现如今，随着互联网的快速发展和Web技术的飞速进步，传统的纪检工作方法无法起到很好的效果。究其原因首先为现如今的组织和企业的员工人数越来越大，传统的通过人工进行的作业无法保证全面覆盖和100%通知到达。其次现在我们逐渐进入到了一个快节奏的时代，对于各项工作的效率要求和质量要求越来越高，传统的纪检方法根本无法满足此类要求。

综上所述，传统的纪检工作方法终究会被逐渐淘汰。现如今的组织、企业和公司却是非常急需要一套标准和规范化的纪检工作流程和方法。

此项目的开发就是为了帮助组织、企业和公司高效、便捷和规范的进行纪检项目的推动和进行。企业可以通过这个系统充分的完成自己对纪检工作的需求。

旨在为用户提供搜索门店与推荐门店的平台，开发了门店搜索推荐系统，方便用户在家中同样可以获取门店的信息。时间宝贵，更何况是如今忙碌奔波在大街小巷的人们，在工作之余好不容易节省出来的闲暇时间，更应该最大化地合理利用，有了门店搜索推荐系统，人们就可以根据信息提前制定计划，合理安排时间。

## 1.2国内外研究现状

### 1.2.1 信息获取类软件国内外研究现状

科技的发展连带着时代的进步，人们生活面貌发生极大的变化，人们主动获取信息的方式主要是搜索。面对现如今海量数据的时代，简单的搜索方式在大量的数据中对比查找效率太低，因此搜索引擎技术也在不断地更新进步。搜索引擎从最开始的人工时代、文本检索时代到整合分析时代、用户中心时代再到现在的生活生态圈搜索时代，经历了一个飞速发展的过程。在互联网新时代，搜索功能不再是仅仅局限于浏览器中，而是有更广阔的搜索空间，越来越多的APP、网站都在不断优化其内部的搜索功能。现如今人们想要搜索餐厅、电影院，甚至是出租车，都有相应的含有此类搜索功能的APP或者网站为用户提供平台[3]。在国内的非浏览器类可为用户提供个性化搜索服务的APP或网站中，有供用户获取实事新闻或社交的，有供用户参考附近门店信息的。对于第一类在国内最典型的是微博，第二类有美团网、糯米网等。下文以微博和美团网为例进行介绍。

（1）新浪微博：新浪微博是国内较为主流的信息交流平台，新浪微博热搜按分钟更新，其内部的搜索功能可以方便用户搜索各种类型的信息，包括新闻、娱乐消息、明星超话等。微博用户量极大，也就意味着其所包含的信息量庞大，如今用户可以通过微博的搜索功能搜索信息，与搜索引擎对比，微博可能会搜索出相对个性化的结果。对于用户来说，如果更想要搜索出符合心意的视频、图片等，微博可能比起传统的百度等搜索引擎来说是一个更不错的选择[4]。

（2）美团网：美团网是一个团购网站，有着“吃喝玩乐全都有”的宣传口号，美团网有许多服务，美食、酒店、旅游、团购、电影票等，由于人们使用手机要比使用电脑方便的多，美团网的移动端的客户量要比网页端的客户量大。美团网的客户端即移动端可以随时随地的查看身边的门店以及优惠团购信息。可以按距离、人气、价格等排序，更直观的为用户显示信息[5]。

对于国外信息获取类软件来说，第一类与微博比较相似的是美国的Twitter；第二类具有代表性的是美国的Zagat survey。

（1）Twitter：Twitter是一家美国社交网站，是全球互联网访问量最大的十个网站之一。Twitter如今已经正式在Twitter.com网页版推出了通用的搜索功能，通用搜索即是指搜索引擎提供的标配服务，搜索结果中融合呈现网页、图片、视频、新闻、本地等搜索结果。Twitter的搜索在原有的搜索结果中，整合进人和照片，方便用户在搜索信息流的同时，也可以找到相关的人和图片。这意味着，Twitter的搜索正在从站内搜索向主流搜索引擎转变[6]。

（2）Zagat survey：用户可以在Zagat.com网站上发表自己对于饮食、住宿、娱乐、旅游等消费领域的评论以及看法。用户还可用根据这些消费信息，搜索自己感兴趣的方面，以此作为消费决策的参考。该网站在世界上已经享有声誉，基于顾客消费体验的消费信息让其在世界范围内获得了大量的忠实用户。

总体来说，虽然非搜索引擎类信息获取软件一时之间完全取代搜索引擎的可能性不大，但是其自身的搜索功能都在不断发展更新，未来其内部搜索功能可能也会占据市场的一席之地。

### 1.2.1 推荐系统国内外研究现状

据了解，目前国内同行业使用的纪检平台完成了信息公开、发布重要通知、展示工作成果、开展信息交流、网络信访举报等功能，缺少“智慧筛查”、“智慧诊断”等功能。国内少数省市政府部门使用的纪检平台除了信息公开、发布通知等常规功能外，利用大数据平台集查询、比对、分析、研判功能于一体，为监督办案工作提供数据支撑，不仅为相关工作带来了效率的提升，还有工作模式的改变，但系统过于庞大，不适合企业使用。

搜索和推荐是相辅相成的两种信息获取方式，对于推荐系统，有两个主要研究方向，一是对推荐算法的研究，二是推荐系统在工程中和商业上的应用。在第一方面，常用的推荐算法协同过滤推荐算法，包括基于物品的协同过滤，基于用户的协同过滤，基于模型的协同过滤，混合推荐算法等[7]。对于第二方面，其实大部分推荐系统都是作为一个模块包含在一个大的项目或者APP中，国内具有较好的推荐功能应用项目有淘宝内部的推荐、今日头条，抖音短视频推荐等，国外有YouTube的视频推荐。淘宝网的商品推荐一般情况下是使用用户画像，通过记录用户的浏览行为和购买历史，构建物品相同类目或者是相关类目属性模型和用户社会属性模型，并将用户、商品以及店铺建立关联。其推荐的原则包括基于物品信息、关联规则、属性筛选推荐，或增加用户画像，对优质商品划分排序等。

技术上推荐系统可以使用基于内存或邻域推荐，基于模型的推荐，或以矩阵分解的传统机器学习方式推荐。YouTube的视频推荐使用了基于User-Video图游历算法，基于搜索以及协同过滤算法，目前已经转用深度学习神经网络[8]。

如今的推荐算法大致可以分为三类，基于协同过滤的推荐，基于内容的推荐，基于知识的推荐算法等。它们适合的应用场景不同，不同的算法也就会带来不同性能上的体现[9]。

## 1.3 本文组织架构

本文主要分为六章对系统进行介绍，组织结构如下：

第1章概述部分主要对系统的开发背景和意义进行介绍。

第2章主要介绍了系统开发所需要的关键技术和开发工具，介绍了其原理。

第3章主要对系统开发进行多方面的可行性研究和需求分析。

第4章主要对系统进行总体设计，包括对功能模块的设计和数据库的设计。

第5章主要对系统开发的详细设计与实现进行介绍，并展示部分用户界面。

第6章对系统测试进行介绍，包括测试方法测试用例等。

论文最后是本文的结论、参考文献和致谢。

# 第2章 开发工具和相关技术

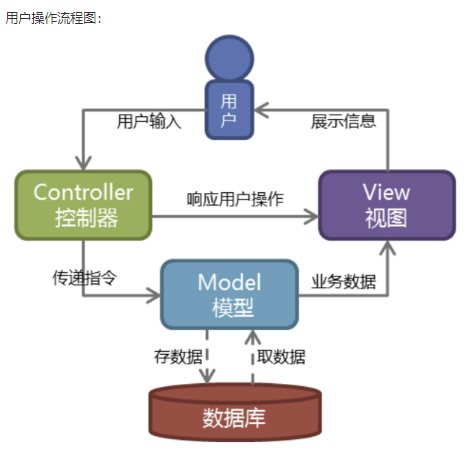
## ·2.1 基于Python的Web框架Django

Django框架是一个由Python语言编写的一个开放源代码的Web应用框架。

Django框架最初被设计用于具有快速开发需求的新闻类站点，目的是要实现简单快捷的网站开发。现如今随着Django框架的不断迭代，Django框架的应用面已经越来越广泛，比如国际知名视频平台YouTube就是由Django框架完成的开发，国内也有例如知乎等大型平台使用Django框架作为其开发基础。

使用Django框架，只要很少的代码，Python的程序开发人员就可以轻松地完成一个正式网站所需要的大部分内容，并进一步开发出全功能的Web服务。

Django框架本身基于MVC模型，即Model（模型）+ View（视图）+ Controller（控制器）设计模式，是软件工程中的一种软件架构模式，Model层主要负责编写程序应有的功能，负责业务对象与数据库的映射，View层主要负责图形界面，负责与用户的交互，Controller层主要负责转发请求，对请求进行处理。如下图2-1所示：



MVC模式使后续对程序的修改和扩展简化，并且使程序某一部分的重复利用成为可能。

MVC模式相比于传统的Web开发具有低耦合、开发便捷、部署方便、可重用性高、维护成本低等巨大优势。

Python加Django是快速开发、设计、部署网站的最佳组合。

## 2.2 DRF框架https://cloud.tencent.com/developer/article/1493264

使用Django框架可以快速的完成对一个Web项目的开发，但是在使用Django框架的途中会出现很多的问题，构建流程非常不便捷，很多冗余代码需要反复编写，极大的降低了开发的效率。这就为DRF框架的出现奠定了坚实的基础。

目前很多项目都在使用DRF（Django REST framework）框架来构建Restful WebApi，由此可见它在Web项目的开发过程中起到了非常重要的作用。

官网中介绍Django REST framework框架是一个强大且灵活的Django工具包，用于在Web后台构建Restful接口。其实在Django下构建Restful接口的工具不止一个，比如Tastypie，但是当下最出名的就是Django REST framework，Django REST framework版本更新速度快，社区活跃度好并且最重要的一点是它的文档十分完善。

Django REST framework主要的优势为：提供了可视化的API调试界面，开发者可以在线测试接口，提供了各种开箱即用的API认证授权工具，如OAuth2，提供了orm数据和非orm数据的序列化和反序列化支持，支持函数视图，它还拥有大量的文档和社区支持。

Django REST framework框架的学习成本非常高，几乎是又学了一个Django框架。所以使用Django REST framework框架的成本就需要针对项目仔细斟酌了。

从项目规模来看，如果是小型项目，使用Django足矣，简单明了，学习成本低。但是一旦进入到了中大型项目甚至是多人参与的项目，则需要使用Django REST framework框架，虽然其学习路线较长，但是从长期来看，它的效率和规范性都要比Django高出许多。

Django框架在编写简单的业务逻辑时会非常容易，但是一旦在视图函数中添加校验逻辑等功能，就会和业务功能耦合在一起，非常容易出现重复代码多、函数过长、不利于维护等问题。除此之外，接口地址的命名和接口返回格式等都需要额外规范来统一。

通过借助Django REST framework，我们可以将接口的开发过程规范化，形成一个统一的代码工作流：

其中Serializers负责数据的序列化和反序列化，结合Validators可以低成本的视线参数校验，Permissions负责权限校验，Routers负责路由自动注册，通过重写框架定义的函数，还可以统一接口返回格式一级异常处理等。总而言之，Django REST framework框架为我们定义好了在什么位置些什么代码，开发都按照这个规范来写，接口才能正常工作。

通过Django框架和Django REST framework框架共同作用，我们还可以很快速的配置出一套Django模型的CRUD接口，将一些开发工作编程配置工作，借助一些Django REST framework周边工具，例如django\_filters，可以快速实现模型数据的过滤类接口，降低了开发成本，并保证了接口的规范以及统一。

其次，使用Django REST framework框架的一个重要的原因是项目分工的精细化，通过使用Django REST framework框架，可以将项目转化为前后端分离的项目，页面渲染的数据都通过Restful接口来提供，前端工程化，后端服务化，代码解耦，开发效率更高。

## 2.3 Vue.js

Vue是一套用于构建用户界面的渐进式框架。与其它大型框架不同的是，Vue 被设计为可以自底向上逐层应用。Vue 的核心库只关注视图层，不仅易于上手，还便于与第三方库或既有项目整合。另一方面，当与现代化的工具链以及各种支持类库结合使用时，Vue 也完全能够为复杂的单页应用提供驱动。

现如今主流的前端JS框架主要有Vue、React和AngularJS，通过对这三种JS框架的对比，最终决定在本项目中使用Vue框架。

Vue框架采用了Virtual DOM，从本质上来说，Virtual DOM知识一个简单的JS对象，并且最少包含tag、props和children三个属性，它们分别代表标签名（tag）、属性（props）和子元素对象（children）。Virtual DOM和DOM对象有一一对应的关系。Virtual DOM最大的特点就是将页面的状态抽象为JS对象的形式，配合不同的渲染工具，是跨平台渲染成为可能。此外，在进行页面更新的时候，借助Virtual DOM，DOM元素的改变可以在内存中进行比较，再结合框架的事物机制，将多次比较的结果合并后一次性更新到前端页面，从而有效地减少页面渲染的次数，提高渲染效率。

Vue框架同时也提供了响应式（Reactive）和组件化（Composable）的视图组件。响应式表示数据和视图同步，即数据绑定，组件化则表示可以将封装好的代码注册成标签，重复使用。

Vue使用了MVVM的设计模式。MVVM其实表示的是Model-View-ViewModel，模型层（Model）主要负责处理业务逻辑和与服务器端进行交互，视图层（View）主要负责将数据模型转化为UI展示出来，比如说HTML页面，视图模型层（ViewModel）主要用于连接Model和View，是Model和View之间的通信桥梁。在MVVM的设计模式下，Model层和View层没有直接的联系，而是通过ViewModel层进行交互。而ViewModel层通过双向数据绑定实现了将View层和Model层连接起来，这样就会使得View层和Model层的同步工作完全自动化。对于开发者来说，只需要关注业务逻辑，就无需手动操作DOM元素了。

Vue还将注意力集中保持在了核心库，而将其它功能如路由和全局状态管理交给相关的库。常见的使用广泛的库比如说VueRouter和VueX。Vue的路由库和状态管理库都是有官方维护支持且与核心库同步更新的，这样就会使得集成起来的Vue框架兼容性极佳。

Vue官方还提供了CLI脚手架，能让我们通过交互式的脚手架引导非常容易地构建项目，我们甚至可以使用VueCLI快速开发组件的原型。

Vue框架的学习路西十分平缓，如果要学习Vue，你只需要有良好的HTML和JavaScript基础即可，有了这些技能，你就可以很快的上手Vue框架。如果只是轻量级的使用，只需在页面中引入Vue.js文件即可使用，开发者只需要不到一天的时间阅读Vue的官方文档即可建立简单的应用程序。

## 2.4 Element-UI

Element-UI框架是一套为开发者、设计师和产品经理准备的基于Vue2.0的桌面组件库。

Element-UI遵循一致性、反馈、效率和可控的设计原则。

其中一致性（Consistency）表现为与现实生活一致：与现实生活的流程、逻辑保持一致，遵循用户习惯的语言和概念；在界面中一致：所有的元素和结构需保持一致，比如：设计样式、图标和文本、元素的位置等。反馈（Feedback）表现为：控制反馈：通过界面样式和交互动效让用户可以清晰的感知自己的操作；页面反馈：操作后，通过页面元素的变化清晰地展现当前状态。效率（Efficiency）表现为：简化流程：设计简洁直观的操作流程；清晰明确：语言表达清晰且表意明确，让用户快速理解进而做出决策；帮助用户识别：界面简单直白，让用户快速识别而并非回忆，减少用户记忆负担。可控（Controllability）表现为：用户决策：根据场景可给予用户操作建议或安全提示，但不能代替用户进行决策；结果可控：用户可以自由的进行操作，包括撤销、回退和终止当前操作等。

Element-UI框架提供了基于Vue的大量组件，使用Element-UI框架封装好的代码可以帮助开发人员进行快速开发。迅速搭建前端架构。

## 2.5 Apache Echarts

Apache Echarts是一个基于 JavaScript 的开源可视化图表库。

ECharts，一个使用 JavaScript 实现的开源可视化库，可以流畅的运行在 PC 和移动设备上，兼容当前绝大部分浏览器（IE8/9/10/11，Chrome，Firefox，Safari等），底层依赖轻量级的矢量图形库 ZRender，提供直观，交互丰富，可高度个性化定制的数据可视化图表。

ECharts 提供了常规的折线图、柱状图、散点图、饼图、K线图，用于统计的盒形图，用于地理数据可视化的地图、热力图、线图，用于关系数据可视化的关系图、旭日图，多维数据可视化的平行坐标，还有用于 BI 的漏斗图，仪表盘，并且支持图与图之间的混搭。

Echarts包含了许多特性：

丰富的可视化类型: 提供了常规的折线图、柱状图、散点图、饼图、K线图，用于统计的盒形图，用于地理数据可视化的地图、热力图、线图，用于关系数据可视化的关系图、treemap、旭日图，多维数据可视化的平行坐标，还有用于 BI 的漏斗图，仪表盘，并且支持图与图之间的混搭。多种数据格式无需转换直接使用: 内置的 dataset 属性（4.0+）支持直接传入包括二维表，key-value 等多种格式的数据源，此外还支持输入 TypedArray 格式的数据。千万数据的前端展现: 通过增量渲染技术（4.0+），配合各种细致的优化，ECharts 能够展现千万级的数据量。移动端优化: 针对移动端交互做了细致的优化，例如移动端小屏上适于用手指在坐标系中进行缩放、平移。 PC 端也可以用鼠标在图中进行缩放（用鼠标滚轮）、平移等。多渲染方案，跨平台使用: 支持以 Canvas、SVG（4.0+）、VML 的形式渲染图表。深度的交互式数据探索: 提供了 图例、视觉映射、数据区域缩放、tooltip、数据刷选等开箱即用的交互组件，可以对数据进行多维度数据筛取、视图缩放、展示细节等交互操作。多维数据的支持以及丰富的视觉编码手段: 对于传统的散点图等，传入的数据也可以是多个维度的。动态数据: 数据的改变驱动图表展现的改变。绚丽的特效: 针对线数据，点数据等地理数据的可视化提供了吸引眼球的特效。通过 GL 实现更多更强大绚丽的三维可视化: 在 VR，大屏场景里实现三维的可视化效果。无障碍访问（4.0+）: 支持自动根据图表配置项智能生成描述，使得盲人可以在朗读设备的帮助下了解图表内容，让图表可以被更多人群访问。

2018年3月全球著名开源社区Apache宣布百度ECharts进入Apache孵化器。**2.6 系统开发及运行环境**

 1．系统开发环境及工具插件

（1）开发平台：Pycharm2022.1，WebStorm2022.1

（2）开发语言：Python，HTML，JS

（3）开发框架：Django，Django REST framework，Vue，Element-UI

（4）数据库：MySQL8.0.27

（5）其它工具：Echarts

2．系统运行环境及硬件设备

（1）操作系统：Windows11专业版

（2）文档编写工具：Typora，Office365

（3）浏览器：Google Chrome

（4）处理器：AMD Ryzen 7 5800H with Radeon Graphics

（5）内存：16.0 GB DDR4

（6）硬盘：512G SSD

2.7 本章小结

本章主要介绍了网站开发工具和相关技术，开发语言为Python，开发工具主要为Pycharm和WebStorm，相关技术有Django，Django REST framework，Vue，Element-UI，Echarts，MySQL数据库等。同时介绍了系统需要配置的软件环境和硬件环境等。

# 第3章 可行性研究及需求分析

## 3.1 可行性研究

可行性研究是在项目的决策阶段的最重要的工作。对系统可行性研究的过程既是多方面比对的过程，也是对相关的资料深入调查研究的过程。通过对系统可行性进行研究，可以在很大程度上对于项目的建设模式有整体性的了解，以便于对后续的工作进行安排。

在对系统进行可行性研究时，为了保证系统后续开发实现的质量，在社会、安全、发绿、文化、环境、经济等方面对系统的功能或开发方式等进行深入的可行性研究，以便规避开发风险，为系统后续的开发计划做必要的准备。可行性研究为后续建设项目投资决策和编制任务书等工作提供了重要的依据。

### 3.1.1 社会可行性

对系统的社会可行性分析可以涉及项目对社会的影响，包括政治体制、方针政策、经济结构、法律道德、宗教民族、妇女儿童及社会稳定性等。智慧纪检平台的研发旨在解决“互联网+大数据”时代，传统纪检监督工作的形势与不足，计划通过研发模块、关联数据、完善流程为着力点和切入点，通过业务流程镶嵌信息系统提升纪检工作的规范化，通过关联数据打破信息壁垒推动纪检工作高效化，通过研发模块监督权力运行推动纪检工作智能化。对社会体制等并无影响，而且可以高效并且规范的完成企业的纪检工作，提升企业工作效率，规范化监督组织成员，对推动企业工作和社会稳定起到了积极向上的作用。

### 3.1.2 安全可行性

一个企业内部的纪检系统的系统安全性是一个非常重要的指标，随着爬虫和信息挖掘技术的不断发展，一个合格的企业项目应该把用户的信息放在首要的考虑位置。

在设计系统的同时一定要全面的考虑到安全和权限等问题，不仅要采用合理的权限认证框架，对部分数据进行加密处理也是十分有必要的。智慧纪检平台在设计之初就将信息安全作为系统考量的一个重要的指标：每一步操作都要对用户的权限进行验证，对于密码等敏感数据采用加密处理，最大程度的保护用户的信息安全与隐私。

整个项目开发过程中使用的开发工具、框架、第三方扩展都使用官网提供的正版，所有用到的技术都采取查阅官方文档的形式进行学习和使用。完全符合安全标准。

### 3.1.3 法律可行性

对于智慧纪检平台的开发工作，完全遵循法律守则。所有环节严格把控，决不允许出现“擦边球”或者违规情况。无论是系统开发使用到的IDE开发工具，还是第三方工具都严格查阅开源许可协议。IDE开发工具采用官网的直链下载，第三方工具会查看开源协议，确保不会出现法律层面的问题，不会出现法律纠纷。

智慧纪检平台的开发务必做到安全、完善，系统做到权限认证覆盖接近100%，不留任何安全陷阱和所谓的“后门”，从根源上杜绝可能产生的法律问题，外部人员无法通过任何手段获取到用户信息，打造一个安全、完善、有保障的安全系统环境。智慧纪检平台会确保满足发绿可行性的要求。

### 3.1.4 文化可行性

在文化方面，智慧纪检平台遵循传统文化的特点，将界面设计的朴素并且简洁，确保用户直接可以上手使用。UI设计和界面文字完全符合社会主义核心价值观中对于文化的要求。界面设计采用白色为主调，搭配部分深色的设计理念，重要的按钮或者提示，都采用明亮的颜色标示出来，确保用户可以直接注意到。

智慧纪检平台确保系统中没有任何垃圾信息，例如广告和推销信息等等。界面设计和功能模块符合传统文化的特点，尽量贴合大众审美，每一处的设计都经过深思熟虑，措辞经过反复地推敲，确保整个系统符合传统文化的特点。

### 3.1.5 环境可行性

在现如今的大环境下，对于环境的保护话题越来越受到人们的重视。保护环境作为一个亘古不变的话题，我们要坚持可持续发展观，国家发展和环境保护同步进行。任何现代化的企业都必须要花费心思确保将对于环境的污染降低至最低。智慧纪检平台的系统开发不会对环境产生任何直接性的影响，完全满足对于环境可行性的评估。

### 3.1.6 经济可行性

对于一个系统的经济可行性的评估和深入分析，有利于最大程度的减少开发成本，合理规划资金的分配。

对于智慧纪检平台的开发，预计采用基于开源框架Django-Vue-Admin框架进行二次开发。通过开源框架可以极大的简化开发流程和节约成本。IDE采用教育版免费正版软件，密码加密和权限认证等功能采用集成第三方开源框架的方法来简化流程。主要支出为项目的数据库采用阿里云的云数据库，由于此项目为前后端分离的项目，原本的电脑已不足以再支持开发，所以更换了更好的笔记本电脑。

总体来说，对系统各方面的可行性分析发现，系统开发可以符合社会、安全、发绿、文化、环境、经济等可行性，可以进行下一步工作，合理开发。

## 3.2 系统的需求分析

在系统工程及软件工程中，需求分析指的是在创建一个新的或改变一个现存的系统或产品时，确定新系统的目的、范围、定义和功能时所要做的所有工作，其中包括考虑来自不同利益相关者的需求，确认是否冲突，在冲突的需求之间进行取舍，并针对软件需求及系统需求进行分析、记录、确认以及管理[TODO]（Kotonya, Gerald; Sommerville, Ian. Requirements Engineering: Processes and Techniques需要免费注册. Chichester, UK: John Wiley and Sons. 1998. ISBN 9780471972082.）。

需求分析是软件项目或系统项目中的关键过程，关系项目的成败[TODO]（Alain Abran; James W. Moore; Pierre Bourque; Robert Dupuis (编). Chapter 2: Software Requirements. Guide to the software engineering body of knowledge 2004. Los Alamitos, CA: IEEE Computer Society Press. March 2005 [2007-02-08]. ISBN 0-7695-2330-7. It is widely acknowledged within the software industry that software engineering projects are critically vulnerable when these activities are performed poorly.）。理想的需求要整理成文件、可以执行、可以量测、可以测试、可以追踪、和识别到的商业的需求或机会有关，而且要有系统设计的相关设计细节。

### 3.2.1 业务需求

业务需求对于用户和市场来说是软件的更高级层的目标和要求。通过对于现有业务的进行评估，包括但不限于业务流程、业务组织建模、改进业务流程和领域建模等等方面，并且考虑到在未来的一段时间内可能需要对业务进行维护，甚至于扩展，需要对于业务需求有一个很好的分析与总结。总的来说就是站在业务逻辑的角度对项目成功的预期效果进行完整、详细的分析与总结。在确定业务需求之前，还应在与各类与相关人员在业务需求分析山达成一致的共识，业务需求相对于整个系统来说无异于基石的作用，其他的需求都必须与之相对应。

对于系统的业务背景，随着“互联网+大数据”深入发展，面对纪检监督工作的形势与不足，计划以研发模块、关联数据、完善流程为着力点和切入点，通过业务流程镶嵌信息系统提升纪检工作规范化，通过关联数据打破信息壁垒推动纪检工作高效化，通过研发模块监督权力运行推动纪检工作智能化。

传统的纪检工作，主要由人工参与进行监督和管理。这样的纪检方式，对于一些小型的企业来说确实是足够了。但是对于一些较为大型或者人员众多的企业或组织来说，传统的纪检方式就相对显得捉襟见肘了。无论是通知的下发还是后续的回执都有着诸多不便。首先对于企业，纪检管理人员就是一笔不小的负担，人数越多，需要的纪检工作人员就会相对的增加，这样会对企业的纪检成本造成不小的负担，而且通知的效率也将无法被保证，传统的纪检工作无论是以邮件通知的方式、钉钉群的方式或者微信区的方式都无法保证所有人及时的收到消息，难免会存在有人忘记看消息，或者说消息众多，通知被“顶掉”的现象。最后所有员工的回执也是非常难处理的一点，没有一个统一的提交方式，最终很容易造成文件丢失甚至于整个纪检工作的进度被打乱。

智慧纪检平台为企业的纪检工作提供了一个一站式的解决方案。整个智慧纪检平台的设计遵循解放思维、聚智聚力，加大人工智能手段及数据治理能力在纪检监督工作领域的创新运用，主要赋予了智慧纪检平台三个突出的功能：“智慧筛查”：利用纪检系统数据库信息进行碰撞和比对分析，察觉可疑情境和潜在问题，落实重点领域、重要岗位的严排查，提供问题异常的预防性风险提示。“智慧诊断”：利用基础数据构建相应的算法模型，推导出纪检监督数据背后的内在规律、相互关联和发展趋势，智慧诊断出问题线索的“病灶”，并做出必要的智能预判。“智慧决策”：通过比对历史数据与模型分析，根据“智慧诊断”的信息，辅助人工决策做出科学严谨的“智慧决策”，为党风廉政建设和反腐败工作提供高效合理的判断依据。

根据和实际人员的沟通发现了公司已经具备了一些优势条件：一是公司在智能油田的建设中积累了大量有益经验。公司的数字化、智能化建设实现了油田生产经营管理的扁平化、精益化，并且形成了可借鉴、可复制、可推广的成熟经验，为构建智慧纪检工作平台提供了智能思维及技术支撑。二是公司纪检监督工作基础比较扎实。一直以来，公司纪委坚持以稳中求进提升纪检监督工作质量，建立了《落实上级情况清单》清单、《监察对象信息数据库》《与油田发生业务往来重要企业人员数据库》和《领导人员廉政档案》等资料，为构建智慧纪检工作平台提供了抓手。

业务需求预计主要研发五个模块：1.党风廉政建设工作模块。以《领导干部履行党风廉政建设责任工作手册》为基础，将领导干部履行党风廉政责任规定点项和行动计划以电子化的方式在模块中呈现，利用信息技术手段实现进度提醒、记录查询、检查考核等功能，保证领导干部履行党风廉政建设责任工作不落项、内容有记载、责任可追溯。2.党委巡察工作模块。梳理《油田党委巡察工作任务清单》，建立巡察方式及对应资料数据库，利用信息技术展示各数据库内资料是否齐全或哪项资料缺失，为开展自查、迎接巡察等提供保证。同时完善问题整改、“回头看”等工作流程，实现问题整改的闭环管理。3.监督执纪工作模块。根据中国石油天然气集团公司贯彻落实《中国共产党纪律检查机关监督执纪工作规则（试行）》实施细则，将信息技术手段贯穿审查调查、案件审理等各项业务流程，加强过程管控，让信息技术成为规范执纪执法的重要手段。4.政治生态综合分析模块。从政治生活、队伍建设、选人用人、信访举报、违纪违法案件等多个维度收集数据、案情，利用信息技术手段纵向对比当前数据与前一年度或前一阶段数据变化，各单位之间进行横向比较，通过图表进行展示。借助人工智能手段，对公司整体及各基层单位的政治生态情况进行综合研判，作出正面评价、问题分析、意见建议。5.日常监督智能模块。初步分为“四风”问题监督、廉洁从业监督及监督再监督等3个小项。总体思路为：根据实际情况确定告警阈值，通过数据关联进行及时告警，根据告警内容进行结果处置。例如，“四风”监督中的办公用房使用，以各级工作人员办公室使用面积为数据基础，以各办公室面积为告警阈值，通过智能模块对当前使用面积超标办公室进行告警，利用数据分析对办公室进行合理分配；廉洁从业监督中的化公为私治理，以《管理人员廉洁从业信息表》为数据基础，以《与油田发生业务往来重要企业人员数据库》为告警阈值，通过智能模块进行复核筛查，对存在问题进行及时告警，按照清退要求进行清退；监督再监督中的合规监督，以各职能部门业务流程关键节点为数据基础，以各类相关制度规定实施细则为告警阈值，利用智能模块对违反制度规定行为进行告警，通过告警处置倒逼职能部门压实监管责任，促进合规管理水平。此智能模块存在涉及业务部门多、业务内容覆盖面广等特点，还需与各职能部门尤其是与经营财务办公进行深度结合，进一步深入调研、深入探索。

总而言之，智慧纪检平台为企业提供了一个一站式的纪检工作解决方案，企业可以直接通过智慧纪检平台完成几乎所有的日常纪检工作的需求，后续还会根据实际情况确定维护项目和扩展功能等。

### 3.2.2 用户需求

用户需求是站在使用系统的企业用户角度描述本软件产品主要完成的功能模块。用户需求一般来说是在用例模型的文档当中有具体的描述，从用户需求的概念中还可以引申出软件的质量属性。

对于用户需求分析的工作重点就在于如何搜集用户的需求，通俗的说就是如何确定软件系统为用户提供的具体功能，还有软件系统和外部环境因素的互相影响等等。获取用户的需求是一件简单的事情，问题在于如何完整，充分的获取用户的需求，因为许多的需求是隐性的，并非直接明确的，有的需求还会含糊不清，很难通过简单的沟通直接获得较好的成效，部分需求还会经常性的产生变更。

对于智慧纪检平台来说，通过和专业的纪检人员沟通，站在用户的角度，对用户的需求进行分析，完成用户想要的功能，具体分为以下几个方面：纪检管理人员为纪检对象分配账号，纪检人员可以完成纪检干部管理、监督对象管理、重点人员管理、监督检查管理、党风监督管理、规章制度管理、系统管理等功能。纪检对象登录之后可以查看纪检任务，参与纪检考试等。

### 3.2.3 功能性需求

功能性需求作为规定开发人员必须在产品中实现的软件功能，用户必须利用这些功能来完成任务，满足业务的需求。功能性需求也常被称作为行为需求，因为习惯上总是用“应该”来对其进行一定的描述。功能性需求一般来说就是描述的开发人员需要完成实现什么具体的功能。

智慧纪检平台系统的功能性需求主要为：

TODO（展开介绍）

对纪检干部的管理，包括对纪检干部的增加、删除、修改和查询，同时也可以通过筛选条件进行筛选；对支部纪检委员的管理，包括查询、导入、单条增加、修改、删除和导出等功能；在线答题，纪检对象在登陆之后可以查看到自己的所有考试信息，可以查看自己每套试卷的最高成绩，如果考试时间未截止，则可以继续考试，记录自己的最高成绩；答题结果统计，以不同的颜色区分已完成和正在进行，可以查看已答人员及分数和未答人员；监督对象信息管理，包括查询、导入、单条增加、修改、删除、导出等功能；党政党纪处分管理，包括查询，导入组织措施、党纪处分和政务处分格式验证、单挑增加、修改和删除等功能；子女升学筛查，可以通过年龄范围和身份证号筛选、人工核准、确认提醒、承诺书、廉洁提醒、承诺书签字上传等等功能；婚丧嫁娶筛查，功能要求和子女升学筛查一致；廉洁从业提醒，包括业务往来企业管理信息管理，管理人员廉洁从业信息管理、重点人员筛查；工作运行计划管理，包括导入、查询、导出、工作提醒和完成情况统计等等；落实上级通知，通知登记表管理、完成时间提醒、通知落实管理和导出等功能；工作计划完成统计情况，包括查询和导出；履行责任监督，包括资料完整性检查和导出；支部责任落实督促，包括支部工作完成情况报告上传；按文件名查询规章制度和按内容查询规章制度；支部信息管理；职称、分级分类、职级、熟悉专业、从事纪检工作年限、单位类别、履行责任项、人员类别、违法行为和司法处理情况的数据字典管理；用户管理；题库维护；试卷管理；规章制度管理等功能。

1. 3.2.4 非功能性需求

非功能性需求是软件为了满足用户的业务需求之外必须具有的特性。非功能性需求影响软件产品是否能够持续稳定并且提供高效的服务。在实际项目中比较常见的非功能性需求主要有：性能需求、可靠性需求、安全性需求、易用性需求、可维护性需求和可扩展性需求。

对于性能需求，要求页面相应及时，系统资源利用率不能接近满负荷，否则会造成系统奔溃的隐患。可靠性需求要求系统保证完善异常处理，定时备份数据，防止发生意外情况。安全性需求要求系统采用权限框架，并对部分数据实现加密处理。易用性需求要求界面简洁大气，用户可以简易的上手。可维护性需求和可扩展性需求应该在开发时按照标准开发，写好文档，预留扩展接口。

3.3 本章小结

本章主要对系统进行可行性和需求分析，对社会、安全、发绿、文化、环境、经济等方面进行可行性研究，从业务需求、用户需求、功能性以及非功能性需求方面对系统进行需求分析。

# 第4章 系统总体设计

## 4.1 系统功能模块设计

系统功能模块就是按照系统的功能的从属关系绘制而成的图表，图中的每一个方框都可以称之为一个功能模块。功能模块同时可以根据系统的具体情况进行相应的分解，分解得到的最小的功能模块可以是一个程序中的每个处理的相应过程。而较大的功能模块则可以是完成某一个任务的相应的一组程序。功能模块图是对硬件、软件和解决方案等进行剖析的结果，详细的描述了功能列表的结构和构成。

智慧纪检平台主要分为七个模块，包括纪检干部管理模块、监督对象管理模块、重点人员管理模块、监督检查管理模块、党风监督管理模块、规章制度管理模块和系统管理模块，系统功能模块图如图4-1所示。

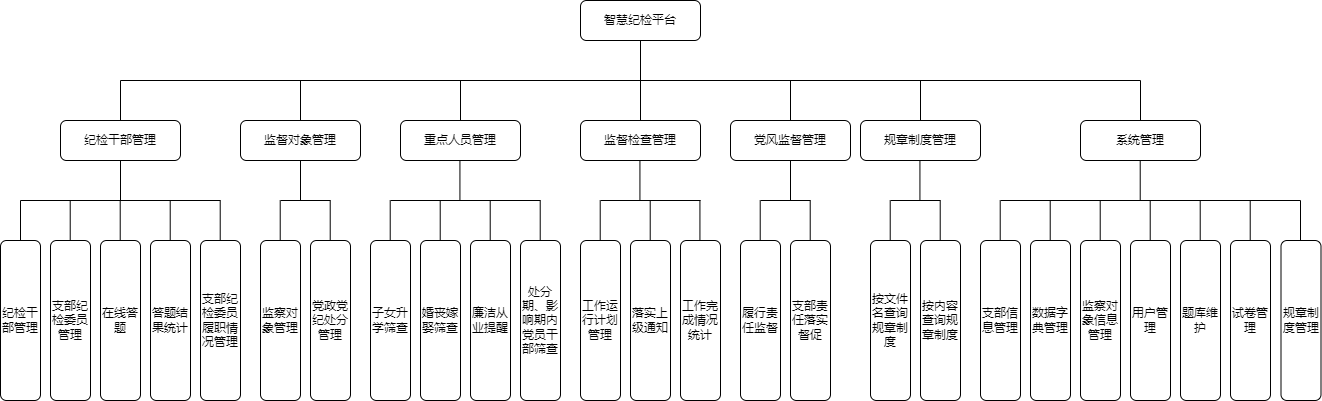


图4-1 系统功能模块图

1. 纪检干部管理模块

纪检干部管理模块的主要功能有：纪检干部管理、支部纪检委员管理、在线答题（支部纪检干部）、答题结果统计（管理人员可查看）和支部纪检委员履职情况管理。

纪检干部管理主要包括纪检干部管理，可以对纪检干部进行查询，查询又分为按照专职和兼职进行条件查询，对纪检干部进行导入操作，同时也可以进行单条增加（部分信息可以由字典选择），可以对纪检干部的信息进行修改和删除，对纪检干部的信息进行导出，导出时可以选择导出字段并且可以调整字段顺序。支部纪检委员管理，可以对支部纪检委员进行查询操作，支持批量导入支部纪检委员管理，同时也支持单条增加（部分信息可以由字典选择），可以对支部纪检委员的信息进行修改和删除操作，同时也支持对支部纪检委员的导出功能，可选择导出字段或者调整字段顺序。在线答题模块，此模块仅对支部纪检干部开放，纪检干部在进入此功能模块时，会以表格的形式展现此用户的所有考试任务，所有的考试任务主要分为已结束和未结束，已结束的考试可以查看答题记录，未结束的考试则可以重复进入答题，取最高分作为最终的成绩。答题结果统计模块，此模块仅有管理员可以进行查看，当管理员进入此页面时，试卷列表会以表格的形式进行展示，管理员可以查询已作答的人员和他们的分数，或者查看未回答的人员。支部纪检委员履职情况管理，在此功能模块中，纪检干部可以在此功能模块中上传履职报告，上传附表并且进行履职情况的综合分析。

纪检干部管理模块主要就是为了对整个企业中的纪检干部的信息做基本的维护，通过简单的操作就可以基本实现对于所有纪检干部的操作，可以使支部纪检干部完成在线考试的功能，查看答题结果，对支部纪检干部的履职情况进行报告和分析。

1. 监察对象信息管理

监督对象信息管理模块主要分为检查对象信息管理模块和党政党纪处分管理模块。

监察对象信息管理模块的主要功能为对监察对象信息的查询，支持通过文件批量导入监察对象，同时也可以选择单条增加监察对象（部分信息可以由字典选择），对监察对象信息的修改和删除操作，最后可以选择导出监察对象。党政党纪处分管理模块支持对党政党纪处分信息的查询操作，可以通过各种条件进行查询，支持导入党政党纪处分信息，在导入时要确定一些条件，比如说，组织措施和党纪处分、政务处分不可同时存在，党纪处分和政务处分可以同时存在，组织措施、党纪处分、政务处分不可以同时为空，组织措施非空或者党纪处分、政务处分两项非空（党员），除了批量导入以外，同时也支持单条增加的功能（部分信息可以由字典进行选择），支持对党政党纪处分信息的修改和删除，导出功能支持选择导出字段、调整字段顺序、按照时间加单位的形式进行导出操作。

监察对象信息管理模块主要为了管理监察对象的信息和对党政党纪处分信息的高效化管理和操作。

1. 重点人员管理模块

重点人员管理模块的主要功能有子女升学筛查，婚丧嫁娶筛查（包括本人和亲属），廉洁从业提醒（包括本人和亲属），处分期、影响期内党员干部筛查。

子女升学筛查模块主要的功能通过设置年龄范围和身份证号进行条件查询，人工核准、确认提醒、承诺书和廉洁提醒导出，用户可以上传廉洁提醒书和签字上传承诺书，如果用户没有签承诺的话，系统将会在六月份对用户进行自动提醒。婚丧嫁娶筛查模块的主要功能和子女升学筛查模块大同小异，同样也包括通过设置年龄范围和身份证号进行条件查询，人工复核、确认提醒、承诺书和廉洁提醒导出，最后还支持廉洁提醒功能和承诺书的签字上传。廉洁从业提醒模块的主要功能包括业务往来企业信息的管理，包括对业务往来企业信息的查询、导入和导出功能，管理人员廉洁从业信息管理模块主要包括对从业信息的查询和导出功能，最后还有一个重点人员筛查功能。

重点人员管理模块主要是为了高效便捷的管理重点人员的重要事项，包括他们亲戚的一些事情，对人员的廉洁从业监督起到了重要的作用。

1. 监督检查管理模块

监督检查管理模块主要包括以下3个功能模块：工作运行计划管理、落实上级通知和工作完成情况统计。

工作运行计划管理模块主要包括可以通过文件批量导入工作运行计划。通过各种条件对工作运行计划进行查询，支持导出工作运行计划，当月未完成的任务可以在月末时对用户进行工作提醒，同时，工作运行计划管理模块也支持对完成情况的操作，包括上传工作运行计划报告和对工作运行计划完成的确认功能。落实上级通知功能模块主要包括了通知登记表管理，对上级通知落实的完成时间提醒。对落实上级通知的管理，包括本部的上传多种落实证明材料和支部的上传多种落实证明材料，同时也支持将落实上级通知的情况进行导出操作。工作完成情况统计包括工作计划和上级的临时通知。此模块主要包括对工作完成情况的查询和导出操作。

监督检查管理模块主要的功能就是对企业的工作计划和上级的通知，还有工作完成情况这3个大的方面进行简单并且高效的管理。对于企业的纪检工作进度起到了非常重要的作用

1. 党风监督管理模块

党风监督管理模块主要包括2个大的部分，一个是履行责任监督模块和另一个是支部责任落实督促模块。

履行责任监督模块主要包括按照年份进行统计资料完整性的检查。同时也支持按照年份进行导出的操作，支部责任落实督促模块则主要负责支部工作完成情况的报告的上传。

党风监督管理模块主要的功能是对支部工作的一些考核和对履行责任监督模块的资料进行按照年份的统计等操作。

1. 规章制度管理模块

规章制度管理模块主要只有2个功能。一个是按照文件名查询规章制度，另一个则是按照内容查询规章制度。

对于规章制度的查询主要分为两种，一种是按照文件名进行查询规章制度，这种查询方式比较普通，采用模糊查询的方式来匹配文件名中是否有符合查询条件的字符。而按照内容查询规章制度的话。TODO

1. 系统管理模块

系统管理模块主要分为六大模块。分别为支部信息管理、数据字典管理、用户管理、题库维护、试卷管理和规章制度管理。

其中，支部信息管理模块主要是对支部的信息进行比较基础的操作，包括但不限于，增、删、改、查、导入和导出功能。数据字典模块则主要是为了智慧纪检平台服务，在此功能模块中定义了常用的10种数据字典，包括职称、分级分类、职级、熟悉专业、从事纪检工作年限、单位类别、履行责任项、人员类别、违法行为和司法处理情况。这10种数据字典对于整个系统来说都是非常重要的，因为这些数据字典几乎在每个功能模块中都被涉猎到了，完成一个优秀的数据字典对于整个系统来说都算是磨刀不误砍柴工。用户管理模块则主要是为了纪检管理人员能够录入用户数据，并且对用户角色进行管理与授权操作，确保了每一个用户的权限都在可接受的范围之内，对于整个系统的安全性和稳定性都有了极大的保障。题库维护模块主要是为了给纪检干部进行出题等操作。它和后面的试卷管理模块相互配合，完成了一套相对功能较为完善的在线考试系统。题库维护模块不仅支持批量导入，而且支持手动创建单选题、多选题和判断题，并对前端的页面进行了特殊的优化，使得用户在录入题目的时候获得最佳的体验。在后端中也对题目的类型和对应的答案进行了诸多的判断，保证了题目的稳定性。因为题目一旦生成之后，如果再要进行修改或删除操作的话，会导致试卷信息完整性不能保障，进而导致整个试卷的崩溃。所以题库维护模块只仅仅支持新增题目，如若修改和删除题目的话，就会对试卷产生影响，故不予以修改或删除权限。试卷管理模块主要是为了创建试卷，在创建试卷的同时可以进行一些设置，包括起止日期、截止日期和各种题目类型的题目数量和分数的选择。可以对每一类题目设置不同的个数和分数。整个创建试卷的过程非常灵活，几乎所有的需求都可以自定义来操作。规章制度管理模块则支持多种格式的规章制度的文件的上传，同时包括了对规章制度文件的增删改查等基本操作。

整个系统管理模块可以算得上是智慧纪检平台这个系统中最为重要的一个模块了。在这个模块中不仅完成了数据字典的管理，而且还有支部信息和用户管理的功能，最后还有一个在线考试的系统，整个功能模块的工作量和难度都是非常有挑战性的。完成之后，系统管理模块的功能将会非常强大。

## 4.2 系统数据库设计

数据库技术是信息资源管理最有效的手段。数据库设计是建立数据库及其应用系统的核心和基础，它要求对于指定的应用环境，构造出较优的数据库模式，建立起数据库应用系统，并使系统能有效地存储数据，满足用户的各种应用需求。本节是对智慧纪检平台的数据库设计的相关介绍。

### 4.2.1 数据库概念结构设计

智慧纪检平台主要涉及到的实体有72个之多。将会在数据表中进行详细的介绍。

用户实体是存放用户具体信息的实体，其属性主要有：用户ID、用户密码、用户手机号、用户昵称、用户性别、创建时间、修改时间等。

品类实体是用来存放门店所属品类信息的实体，其属性主要有：品类ID、品类名称、品类标志图片、品类显示优先级、创建时间、修改时间等。

商户实体是存放商户具体信息的实体，其属性主要有：商户ID、商户名称、商户评分、是否禁用标识、创建时间、修改时间等。

门店实体是存放门店具体信息的实体，其属性主要有：门店ID、门店名称、门店评分、人均价格、门店地理位置纬度信息、门店地理位置经度信息、所属品类ID、标签、营业时间、关闭时间、门店地址、门店所属商户ID、门店标志图片、创建时间、修改时间等。图4-2为系统的E-R图。

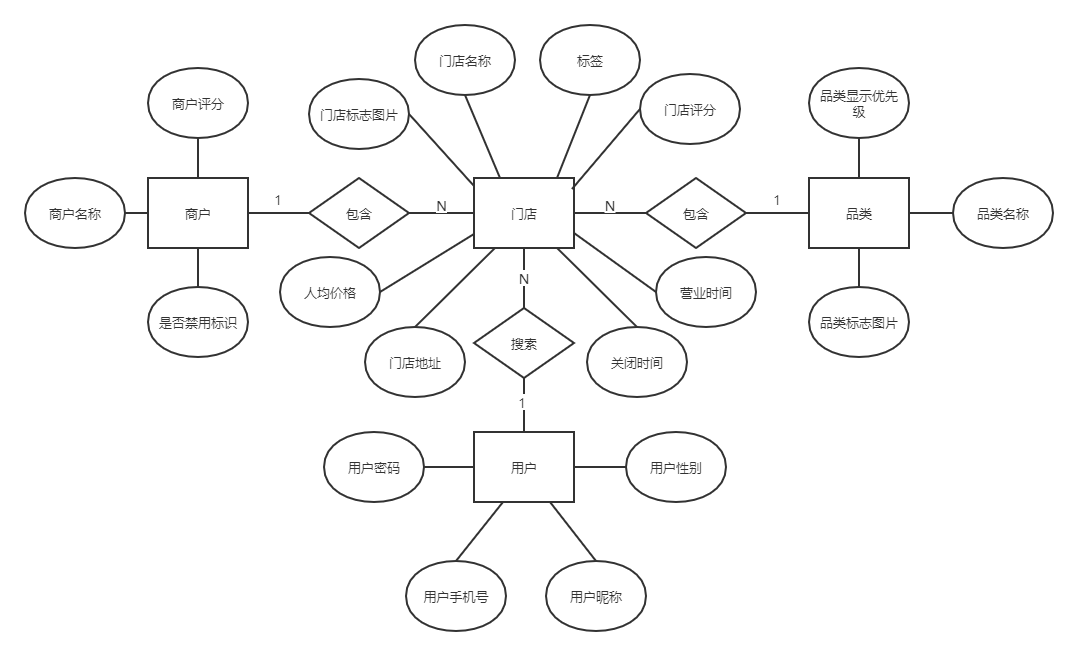


图4-2 系统E-R图

### 4.2.2 数据库逻辑结构设计

本系统主要使用MySQL数据库存储数据，系统主要涉及到的表有用户信息表、管理员信息表、品类信息表、商户信息表、门店信息表、推荐信息表。数据库模型如图4-3所示。

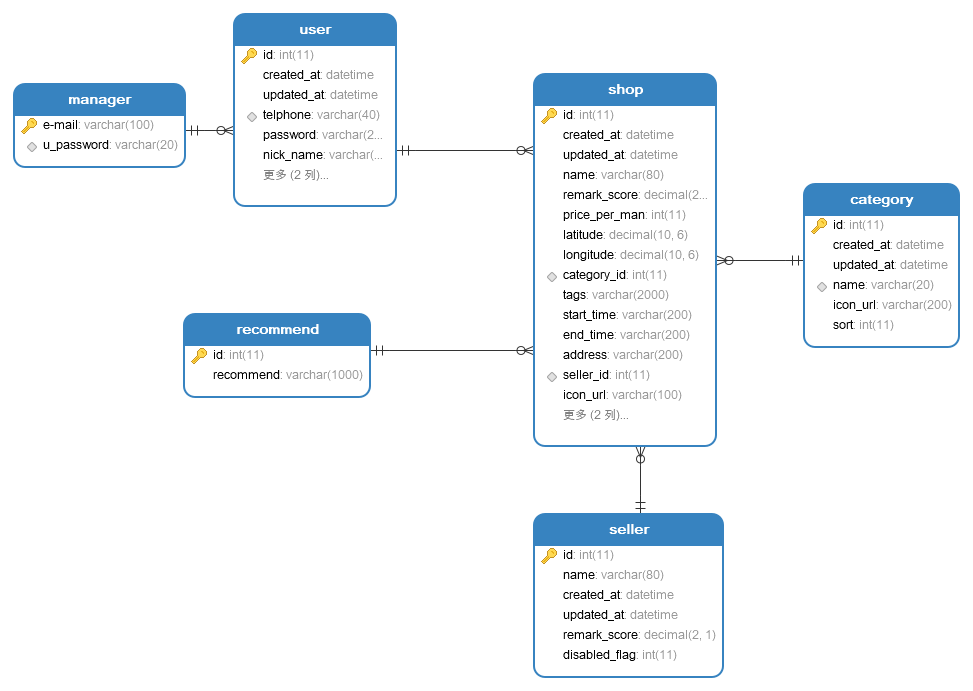


图4-3 数据库模型图

1、用户信息表

用户信息表主要包括ID、用户名、用户密码、用户性别、用户手机号等字段的信息。该表主要是记录用户的基本信息，详见表4-1。

表 4-1用户信息表字段介绍

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 允许空 | 主键 | 说明 |
| Id | INT | 11 | 否 | 是 | 用户ID |
| nick\_name | VARCHAR | 40 | 否 | 否 | 用户名 |
| Password | VARCHAR | 200 | 否 | 否 | 用户密码 |
| Gender | INT | 11 | 否 | 否 | 用户性别 |
| Telephone | VARCHAR | 40 | 否 | 否 | 用户手机号 |
| created\_at | DATETIME | — | 否 | 否 | 创建时间 |
| updated\_at | DATETIME | — | 否 | 否 | 修改时间 |

2、管理员信息表

管理员信息表主要包括管理员邮箱、密码字段的信息，该表主要用来记录管理员信息，详见表4-2。

表4-2管理员信息表字段介绍

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 允许空 | 主键 | 说明 |
| e-mail | VARCHAR | 100 | 否 | 是 | 管理员邮箱 |
| u\_password | VARCHAR | 20 | 否 | 否 | 管理员登录密码 |

3、品类信息表

品类信息表主要包括品类ID、品类名称、品类标志图片等字段的信息。此表主要记录门店所属品类的具体信息，详见表4-3。

表4-3品类信息表字段介绍

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 允许空 | 主键 | 说明 |
| id | INT | 11 | 否 | 是 | 品类ID |
| name | VARCHAR | 20 | 否 | 否 | 品类名称 |
| icon\_url | VARCHAR | 200 | 否 | 否 | 品类标志图片路径 |
| sort | INT | 11 | 否 | 否 | 品类优先级排序 |
| created\_at | DATETIME | — | 否 | 否 | 创建时间 |
| updated\_at | DATETIME | — | 否 | 否 | 修改时间 |

4、商户信息表

商户信息表主要包括商户ID、商户名称、商户评分等主要的字段信息。该表主要用来存放商户的具体信息，详见表4-4。

表4-4 商户信息表字段介绍

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 允许空 | 主键 | 说明 |
| id | INT | 11 | 否 | 是 | 商户ID |
| name | VARCHAR | 80 | 否 | 否 | 商户名称 |
| remark\_score | DECIMAL | 2 | 否 | 否 | 商户评分 |
| disabled\_flag | INT | 11 | 否 | 否 | 是否禁用 |
| created\_at | DATETIME | — | 否 | 否 | 创建时间 |
| updated\_at | DATETIME | — | 否 | 否 | 修改时间 |

5、门店信息表

门店信息表主要包括门店ID、门店名称、门店评分、门店标签、门店地址等字段信息。门店信息表主要用来存放门店的具体信息，以及一些和商户、品类等的关联字段，详见表4-5。

表4-5 门店信息表字段介绍

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 允许空 | 主键 | 说明 |
| id | INT | 11 | 否 | 是 | 门店ID |
| name | VARCHAR | 80 | 否 | 否 | 门店名称 |
| remark\_score | DECIMAL | 2 | 否 | 否 | 门店评分 |
| price\_per\_man | INT | 11 | 否 | 否 | 人均价格 |
| latitude | DECIMAL | 10 | 否 | 否 | 地理位置纬度 |
| longitude | DECIMAL | 10 | 否 | 否 | 地理位置经度 |
| category\_id | INT | 11 | 否 | 否 | 品类ID |
| tags | VARCHAR | 2000 | 否 | 否 | 门店标签 |
| start\_time | VARCHAR | 200 | 否 | 否 | 营业时间 |
| end\_time | VARCHAR | 200 | 否 | 否 | 关闭时间 |
| address | VARCHAR | 200 | 否 | 否 | 门店地址 |
| seller\_id | INT | 11 | 否 | 否 | 商户ID |
| icon\_url | VARCHAR | 100 | 否 | 否 | 门店标志图片路径 |
| created\_at | DATETIME | — | 否 | 否 | 创建时间 |
| updated\_at | DATETIME | — | 否 | 否 | 修改时间 |

6、推荐信息表

推荐信息表主要包括用户ID、推荐门店ID列表字段信息。此表主要记录为登录用户推荐门店的相关信息，详见表4-6。

表4-6 推荐信息表字段介绍

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 允许空 | 主键 | 说明 |
| id | INT | 11 | 否 | 是 | 登录用户ID |
| recommend | VARCHAR | 1000 | 否 | 否 | 推荐门店列表 |

4.3 本章小结

 本章主要通过系统功能模块图、E-R图等介绍了门店搜索推荐系统的功能模块设计，以及各个模块的主要功能，本系统主要包括用户模块、搜索模块、推荐模块、后台管理模块四个模块。对系统进行数据库的设计，列出了主要的数据信息表，以便清晰的展示系统面向的主要实体。

# 第5章 系统的详细设计与实现

## 5.1 用户模块的详细设计与实现

用户模块主要包括用户登录和注销、用户注册、查看系统首页等功能。在完成用户模块的搭建之前先封装通用的异常处理机制，以便后续统一调用。

### 5.1.1通用的异常处理机制

在完成系统的各种功能时，经常会遇到用户输入不符合规范或者各种各样的情况，此时需要一些对用户的提示信息，或是需要捕获异常，这就需要系统有完善的异常处理机制，为了方便调用，本系统定义了通用的异常处理机制。

定义status和data属性，若status为success，则返回对应的json数据，若status为fail，则返回定义的通用的错误码对应的格式、提示信息等。在CommonRes中定义一个通用的创建返回对象的方法，即在Controller层中最后返回结果集调用此方法。遇到异常情况即显示提示信息，需要定义通用的错误处理CommonError，在其中定义错误码和错误描述属性，定义一个枚举类型EmBusinessError，在其中存放自定义的一些错误码。定义通用的捕获异常的类BusinessException继承Exception，这样节约了Controller中的代码冗余，只需要throw抛出异常即可。定义GlobalExceptionHandler类，添加@ControllerAdvice令所有Controller都被其包含，定义doError方法，当Controller抛出的异常被应用程序捕获后被最上层的异常捕获，会被这个方法拦截，并返回CommonRes。异常处理流程图如5-1所示。

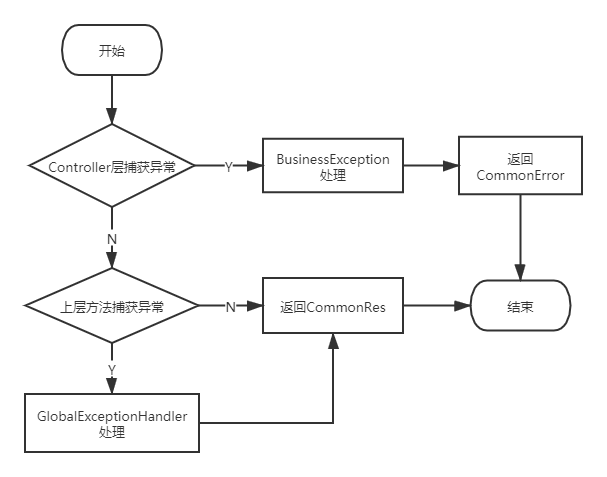


图5-1 异常处理流程图

### 5.1.2 用户注册功能

对于用户注册，在Service层的UserServiceImpl中的register方法添加@Transactonal标签，表示当前所有操作在一个事务中进行，保证事务的完整性和用户模型的闭环性。由于将telphone字段设置了一个unique\_index唯一索引，需要抛出异常，当数据库发生主键或唯一键冲突时捕获异常，并提示用户。出于为用户的安全性考虑，为用户密码设置MD5加密的形式，在encodeByMd5方法中，get到MD5算法的因子，声明BASE64Encoder对象，直接通过encode方法进行加密，并且需要在注册前用密文密码替换掉明文的密码，同样在验证时也需要将明文密码转换为加密格式与数据库中对比。定义RegisterReq在其中对电话、密码、性别等字段进行非空校验，并在Controller层完成与前端界面的交互。用户注册界面如图5-2所示，加密的关键代码如下：

MessageDigest messageDigest = MessageDigest.getInstance("MD5");

BASE64Encoder base64Encoder = new BASE64Encoder();

return base64Encoder.encode(messageDigest.digest(str.getBytes("utf-8")));

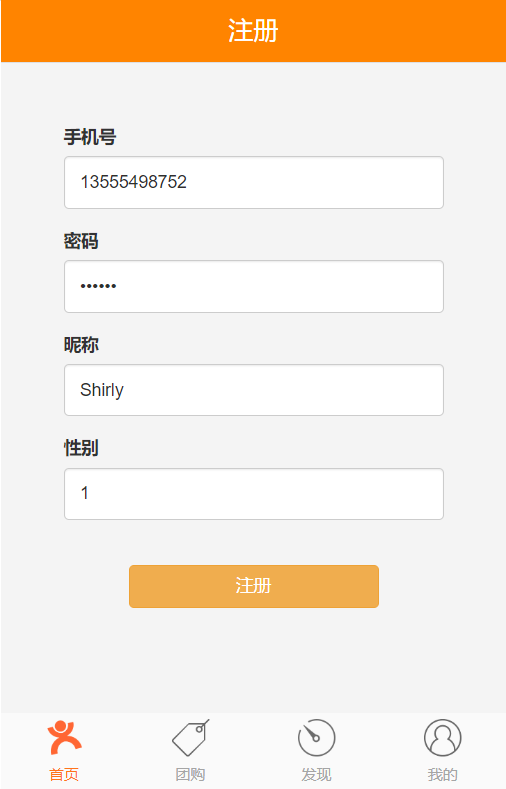


图5-2 用户注册界面图

### 5.1.3 用户登录功能

对于用户登录实际上和用户注册相似，在UserServiceImpl中通过telphone电话和password密码查找用户，将密码的MD5加密格式转义，并抛出异常，当用户输入错误会显示提示信息。Controller层与前端页面进行交互，与注册不同的是登录需要使用HttpSession，当用户登录成功就放入Session内，定义常量CURRENT\_USER\_SESSION，当用户请求成功将UserModel放入Session内，返回通用的CommonRes，要求UserModel继承Serializable表示可以序列化，才可以真正意义上将Session序列化实例化。

在前端的ajax请求中设置当用户登录成功后，首页的登录按钮即变为注销按钮，注销操作即将Session对象invalidate，返回null即可。登录界面如图5-3所示。

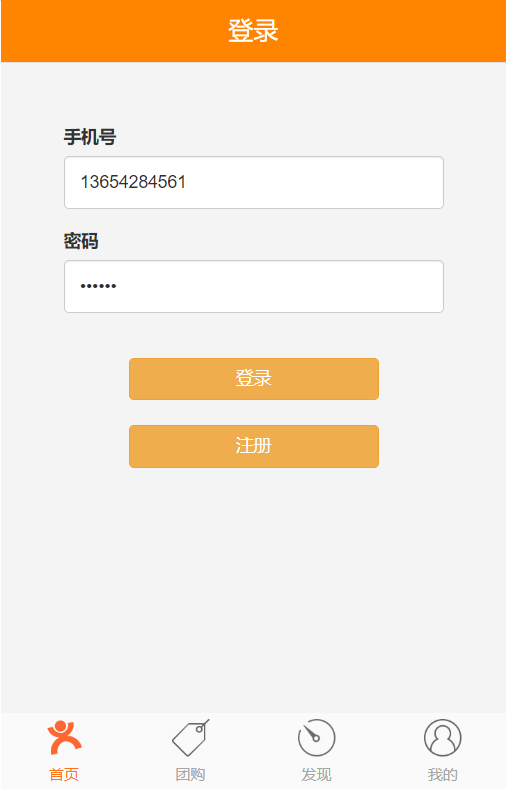


图5-3 用户登录界面图

## 5.2 ELK机制同步数据的详细设计与实现

由于搜索模块是使用ElasticSearch全文搜索引擎工具支撑，需要将MySQL数据库中的数据同步到ElasticSearch中，才可供用户搜索使用。本系统在此处同步数据的操作是采用ELK的机制，即ElasticSearch存储数据，Logstash-input-jdbc同步数据，Kibana可视化工具监控数据。Logstash相当于一个管道，将MySQL数据库中的数据同步到ElasticSearch中，而这个过程，通过控制相关参数使得Logstash只同步支持搜索功能需要的数据，而过滤掉冗余的不需要的数据。

ELK机制同步数据的流程如图5-4所示。

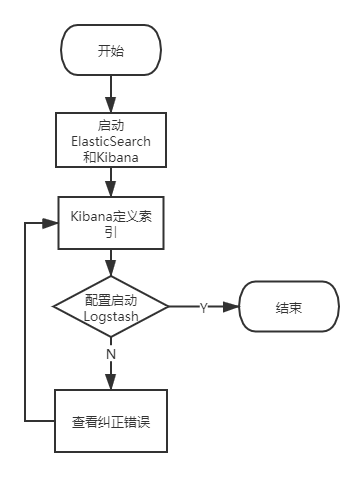


图5-4 ELK机制同步数据流程图

在官网下载相关的工具后，对ElasticSearch进行部署配置，配置ElasticSearch的集群名称，集群中所有的机器都需要共享此名称；结点名称，每一个结点在对应的集群中都应当有一个唯一的名字；启动IP，即配置集群的结点启动的IP地址，并使用http.port区分不同的结点，用来响应Restful请求；配置transport端口用来做集群间的指定通信，即结点通过对应的transport端口做集群的协商以及指定的传输；并配置http.cors允许前端跨域访问；配置discovery.seed\_hosts用来发现对应集群的其它结点；由于集群结点初始无数据，需要初始化master\_nodes，即配置在master\_nodes列表中的结点都有资格竞选为主结点，竞选规则是要结合内部的算法。

出于对系统性能的考量，考虑使用中文分词器，下载中文分词器后部署到ElasticSearch的插件文件中。使用Kibana定义门店索引，需要定义字段逻辑、字段类型、字段的分词方式等。定义的索引的属性最好与数据库表中的字段保持一致，保证和应用之间进行平滑的数据转义。定义门店id属性，类型为integer；门店名称name属性类型为text，指定分词方式为ik\_max\_word，即最大化分词，搜索关键字的分词方式定义为ik\_smart方式；标签tags属性类型为text，由于MySQL数据库中当一个门店有多个标签时，采取的是空格的方式分隔多个标签，所以此处分词方式设为whitespace空格分词器，此处体现了ElasticSearch使用分词器的一大优点，规避了SQL查询语句的局限性，fielddata设为true，表示可以支持标签字段做聚合操作；地理位置信息location属性编码方式为geo\_point，即地理坐标类型；评分remark\_score属性，类型为double；人均价格price\_per\_man属性，类型为integer；品类category\_id属性类型为integer；品类名称category\_name属性，类型为keyword，keyword属性与text属性的区别在于，text类型可以进行分词操作，而keyword类型仅仅将属性作为普通的字符串不进行分词，防止分词产生语义歧义；商户seller\_id属性，类型为integer；商户评分seller\_remark\_score属性，类型为double；商户是否禁用seller\_disabled\_flag属性，类型为integer。定义索引结构的操作在Kibana中完成。

对于索引的构建，可以采用全量和增量的方式构建索引。使用logstash-input-jdbc工具作为管道收集日志，将MySQL数据库中的数据抽取，input到ElasticSearch中。下载logstash并部署，引入MySQL的jdbc的jar包作为驱动，新建jdbc.conf文件和jdbc.sql文件。在jdbc.conf中配置input，即从数据库中抽取的配置，配置jdbc数据库的连接，指定数据库名称、用户名、密码等，加载jdbc驱动，配置驱动类名，执行的SQL文件的路径和名称。设置监听的时间间隔schedule，格式为分、时、天、月、年，每分钟更新一次。output的配置即向ElasticSearch输出的配置，配置ElasticSearch的IP地址及端口、索引名称，关联数据库中的id字段，最终以JSON的格式输出。编辑jdbc.sql文件，在其中编写SQL语句，根据规则过滤不需要的数据，只保留需要的数据作为元数据同步。其中使用concat拼接经纬度作为location地理位置信息字段，以category\_id对category表做inner join的连接，获取category\_name，以seller\_id对seller表做连接，得到seller\_remark\_score和seller\_disabled\_flag。对于增量索引的构建，由于logstash默认会选择0时区的时间，为了更新时间的准确需要在jdbc.conf中设置timezone设为北京时间；设置last\_run\_metadata\_path新建last\_value文件用来记录上一次更新时间。增量索引构建的方式需要在jdbc.sql文件中赋时间轴的初值。使用logstash -f mysql/jdbc.conf的命令启动logstash，完成数据的同步。

## 5.3搜索模块的详细设计与实现

搜索模块包括根据关键词搜索、根据品类、标签、价格排序等进行搜索，根据同义词、相关性等搜索功能。友好的界面为用户提供了方便的交互体验。

### 5.3.1根据关键词、品类搜索功能

搜索系统最基础的功能一般是根据关键词搜索的功能。用户输入想要搜索门店的关键词，系统根据索引匹配并返回搜索结果展示给用户。点击酒店品类，即显示含有关键词的所有酒店，搜索结果界面如图5-5所示。



图5-5 关键词品类搜索界面图

用户在首页的搜索框内输入关键字，点击键盘的Enter键跳转到搜索结果页面。具体后端的逻辑实现为，首先引入ES的服务，在Java层面接入ElasticSearch服务有三种方式，可以使用Node API方式直接接入，可以基于Transport API接入，借助9300端口的Transport API接入ElasticSearch的集群，可以使用轻量级的HTTP Rest Client方式接入ElasticSearch，对于第一种接入方式，将Java应用程序作为一个结点接入集群，在其内部可以进行分片，可以访问、通讯所有ElasticSearch的信息，但是这种方式的缺点在于，Java应用程序在此种情况下有多套业务环境，其仅仅提供的是业务属性层的服务，并不需要感知分片，也不需要知道具体分片的存储位置，更不需要存储分片以及备份的数据信息，因此Java应用程序不适合使用此种方式接入。第二种以Transport通信端口9300进行交互，Transport Client连接所有ElasticSearch中的9300端口，但其本身并不是集群的一部分，而是通过9300端口并且符合其对应的transport协议为集群中的结点发送指令，获取数据，在ElasticSearch5之前，许多应用程序都是通过Transport方式交互，效率相对高，但ElasticSearch5之后，逐渐转移到rest client方式。对于rest client方式，即基于HTTP通信，其优点在于，仅仅需要连接ElasticSearch中任何一个结点，就可以通过发送HTTP请求交互，并不需要感知集群中网络拓扑的连接状态。是最为适合应用程序开发的方式，因其不受语言、地理环境的限制，仅仅使用最普通的HTTP协议即可满足对应的要求。本系统使用的是ElasticSearch7.3.0版本，ElasticSearch7之后已经废弃Transport API协议的支持，本系统使用第三种rest client方式。

在pom.xml文件中引入ElasticSearch的相关jar包，在application.yml配置文件中声明ElasticSearch的服务地址，并且在config配置包下的ElasticSearchRestCliet类中存储对ElasticSearch的配置，保证能够成功对接ElasticSearch，该类应该加@Configuration注解表示其是一个配置的Bean，将配置文件中的IP和Port设给HTTPHost，并且指定HTTP协议，使用new的方式构造了HTTP HighLevelRest API即可在其它的Bean通过name的注入可以获取对应的数据。在Service层的shopServiceImpl中引入RestHighLevelClient的Bean，并且编写searchES方法，通过Request方式对接ElasticSearch的Restful API，此处可以认为Request是基于ElasticSearch的rest client的Request，使用GET方法，索引名称为shop，对索引进行search操作。为了增强系统的可扩展性和可维护性，选择使用fastjson的JSONObject的方式构建请求，经历初次构建之后，后期的操作直接在其上增添即可，避免了从Kibana中重复粘贴控制语句。使用put的方式构建字段，每次构建下一个属性之前需要使用getJSONObject或getJSONArray的方式获取上一步构建的参数，反复直至语句完成。需要传入前端页面的参数时，直接将对象参数替换成前端需要的动态参数即可。需要构建的是source部分，自定义的距离字段，即使用haversin函数根据获取的位置信息和经纬度信息计算出距离，代码如下：

jsonRequestObj.getJSONObject("script\_fields").getJSONObject("distance").

getJSONObject("script").put("source","haversin(lat,lon,doc['location'].lat,doc['location'].lon)");

query部分，即按照关键词查询功能，先构建外围的function\_score，在其内部构建控制查询的语句，构建match匹配关键字等查询信息，在must其中有许多term属性，此处直接定义一个queryIndex，加载了一个term之后将queryIndex++即可继续加载下一个term。第一个term是关键词的term，对于关键词将boost打分设为0.1；第二个term是启用标志disabled\_flag的term，若为0表示商户启用可以被展示给用户；第三个term是品类名称，若收到了前端传入的categoryID，即需要根据品类分类搜索，加载第三个term中的内容。

将最终封装的结果对象放入Request的setJSONEntity中，调用RestHighLevelClient的getLowLevelClient方法返回给HTTP的Response。将Response的实体拿给EntityUtils，转化为字符串，此处的字符串即为发出请求后查询出的一系列结果的字符串格式。将该字符串转化为JSONObject的形式，取到JSONArray形式的hits结点下的一系列结果，遍历JSONArray，获取其中的查询结果，由于获取方式复杂，可以选择只获取ID和距离，根据后台封装好的SQL查询的方式，直接根据ID查询出门店的所有信息，并将距离信息set到shopModel中。获取距离字段信息的代码如下：

BigDecimal distance=new BigDecimal

(jsonObj.getJSONObject("fields").getJSONArray("distance").get(0).toString());

将ElasticSearch中返回的千米的函数转换为米后向上取整，获得最终精细化后的距离。最后在Controller层实现与前端页面的交互即可实现整体的需求。

### 5.3.2 根据排序、标签查看搜索结果功能

系统的排序功能有默认排序和低价排序，根据上节介绍，要实现默认排序功能首先要构建functions部分，在其中定义排序规则，对于默认的排序规则，系统规定根据距离和评分等因素进行排序，本系统为了更精确排序模型，对距离使用ElasticSearch的高斯衰减函数打分，需要origin参数，设为用户所在经纬度；scale参数，表示到设定值为衰减的拐点，设置为100km；offset参数，起始点0km；decay参数，设为0.5。ElasticSearch支持的衰减函数有三类，高斯guass、指数exp和线性linear，线性函数是条直线，一旦直线与横轴相交，所有其他值的评分就均为0；指数函数是先剧烈衰减再变缓，高斯函数是钟形的，它的衰减速率先是缓慢，然后变快，最后又放缓。高斯函数支持四种参数，origin中心点，或者是字段有可能的最佳值，落在中心点上的文档评分为满分1分，支持数值、时间以及经纬度地理坐标点字段；以origin为中心，为其设一个便偏移量offset覆盖一个范围，在此范围内的所有评分同样都是和在中心点上一样的满分；scale是衰减率，是一个文档从origin中心点下落时，评分改变的速度；decay是从origin衰减到scale所得的评分，默认为0.5。对于本系统来说，与用户所在距离为0的门店评分为1分，在此范围之外，评分会发生衰减，衰减值由scale值和decay值决定。默认排序除了需要距离的打分之外，还需要门店评分和商户评分，为了评估准确将五分制归一化，将门店的weight属性设为0.2，商户的设为0.1。最终系统的搜索结果是将系统默认的TF-IDF打分，与定义的functions中的打分融合，需要设定score\_mode为sum即functions内部打分的分值相加的形式，boost\_mode为sum是functions内部和TF-IDF结合的形式同样为相加。为了加大距离因子对打分结果的影响，将其的weight设为9，最后系统以打分结果降序的形式显示给用户。

继续上节介绍的JSONObject操作，构建functions的部分，定义functionIndex，每写一条属性将functionIndex++，并在其中增加包含的控制语句，构建sort排序字段，即评分降序排序，便可完成默认排序功能。同理对于低价排序，当orderby不为空时，前端接收到低价排序的信号，只需要在其中扩展低价排序规则，影响排序函数因素field\_value\_factor的是price\_per\_man人均价格，由于只有一个影响因素，weight属性直接设为1，不同的是，boost\_mode设为replace，即查询的评分只影响召回，不影响排序，排序字段为asc价格升序排序。

对于根据标签显示查询结果，首先在query条件里添加一个标签查询的term，前端接收到标签查询信号的时候，启用此功能。为查询添加一个聚合字段aggs，名称为group\_by\_tags，由于对标签使用的是空格分词，每次查询聚合后，对标签关键字做doc\_count的统计就可看出有几家门店拥有此标签。最终需要将搜索到的标签信息展示，获取到JSONArray形式的aggregations其中的buckets结果，获取其中JSONObject形式的key字段即标签名称，doc\_count字段即含有此标签的门店个数，返回前端页面。按照价格排序后标签筛选的界面如图5-6所示。



图5-6 排序标签筛选查询结果界面

### 5.3.3 根据同义词扩展查看搜索结果功能

以搜索“凯悦”关键词为例，搜索该关键词时由于分词器内部的机制，会出现其它的含有“悦”字的门店，用户交互体验不好，为了避免此类情况，将“凯悦”拓展为分词器词库中的词，即可解决这个问题。在analysis-ik分词器插件的词典中新增词典new\_word.dic，在其中添加凯悦作为新词，并且在配置文件IKAnalyzer.cfg.xml中的entry结点的ext\_dict中加载扩展词典，在Kibana中使用\_update\_by\_query的方式更新索引即可。

用户在使用的过程中，可能会遇到这种情况，搜索“苹果”时，系统自动也为其显示iPhone的内容，实际上这是系统开发时期为“苹果”设置同义词的体现，本系统以为“凯悦”设定同义词为例。同样在分词器插件的字典列表中新增一个字典synomyns.txt，在其中为“凯悦”设定同义词“红桃”。在Kibana中重构shop索引，在settings中定义analysis字符分析规则，使用同义词过滤的方式解析同义词词典；自定义分析器，第一个分词原则是ik\_smart，第二个分词原则是ik\_max\_word，过滤器是定义的analysis分析原则；修改name字段的分词方式为自定义的两个分词原则，即可完成对同义词搜索方式的扩展。

### 5.3.4 根据相关性查看搜索结果功能

对于相关性的重塑，是希望搜索引擎可以理解语义的过程，最后会对召回和排序的结果有一定的影响，通过调整影响打分的参数来对结果产生影响。

首先需要在系统中构建分词函数识别器，构建init方法，添加@PostConstruct标签，表示方法在Bean初始化容器之后就会被执行，在其中存放关键词，例如将“吃饭”、“下午茶”等关键词与美食品类的ID关联，将“休息”、“住宿”等于酒店品类关联等。在getCategoryIdByToken方法中获取token并进行比较即根据获得的关键词与categoryList中的类目做比较，识别成功即返回对应类目的ID。在analyzeCategoryKeyword方法中进行Restful请求，与ElasticSearch交互。最后在构建query操作的过程中对相关性的功能进行扩展，获取上个方法返回的Map，为该功能设定一个bool查询条件默认为true，判断当此Map被匹配，且bool条件被应用时，执行相关性的功能，与前几节类似，设定filterQuery参数，每多嵌套一层就会被加一，在bool查询中嵌套should属性，match匹配到输入的关键字中含有“住宿”等这类设定的关键词，遍历相关性的词库，在Map中有多少条category\_id被命中，filterQuery就会被加多少层，相当于多增加了term的属性，即若设定“唱歌”关键词既关联KTV品类，也关联休闲娱乐品类，则在should属性中相应的打分就会增加。

相关性的搜索不止可以对召回结果有影响，也可以对排序结果也有一些影响，对于排序结果的影响，即在系统中显示的先后顺序，但是在实际开发过程中一般不会同时设定影响召回和排序结果，一般情况下召回策略不容易调整，但是排序模型非常容易调整，在orderby排序处添加对排序的影响字段，使用filter影响排序，为了体现关键词对于排序的影响，权重值weight通过调整设为5。具体的搜索结果界面如图5-7所示，以“凯悦休息”关键字为例，根据分词器的分词结果，将“凯悦”和“休息”分为两个词，可以发现，虽然“凯悦咖啡”也匹配了“凯悦”的关键字，但是由于将“休息”相关性的关键词设置的权重增强，“凯悦咖啡”不属于酒店品类，所以搜索显示的结果会相对靠后一些。



图5-7 相关性搜索结果界面

## 5.4 推荐模块的详细设计与实现

推荐模块主要完成猜你喜欢的推荐功能，通过ALS算法进行召回粗排，将粗排的结果存入数据库中，再使用LR算法进行精排。

### 5.4.1 ALS算法训练实现

Spark MLlib包中主要包含的算法有分类、回归、聚类、评估、特征、推荐等，推荐算法有矩阵分解算法和交替最小二乘算法，即ALS算法，本系统使用的是ALS算法。从协同过滤的角度来说，ALS算法属于混合协同过滤，ALS算法同时考虑了用户和产品两方面，通过观察所有用户为产品的打分情况，来推测每个用户的喜好并且向用户推荐合适的产品。将用户和产品的关系抽象为一个三元组<user,item,rating>，而其中的rating代表的是用户相对于产品的评分，表示的是用户对产品的喜好程度，在一批用户数据中，有m个用户和n个产品，根据用户的特征矩阵乘以物品的特征矩阵得到了评分矩阵，矩阵中的元素表示的是第u个用户对第i个产品的评分，因为一个用户不可能为所有的产品评分，所以该评分矩阵是一个稀疏矩阵。ALS算法的作用是为了预测稀疏矩阵中的空点，若该空点的打分值比非空点高，就将空点对应的产品推荐给用户。ALS算法分别将用户矩阵和产品矩阵的特征分解，要求都分解出同样数量的特征，并且将用户矩阵和产品矩阵的转置矩阵进行相乘，即V=UPT。经过不断的拟合运算得到最终的推荐结果。

Spark Core的Map Reduce可以将复杂的操作封装成RDD的操作，可以使我们轻松进行数据转换。Spark会将很多条数据拆分成若干条，map后将所有数据合并，进行计算，最后将结果发给目标机器。对于ALS算法的具体实现，引入Spark的MLlib依赖，在AlsRecallTrain类中先加载模型进行训练计算，首先必要的是初始化Spark的运行环境，在Java中是将Java操作转化成JavaRDD的操作，加载behavior.csv文件，用来构造稀疏矩阵，其中主要存放三个参数，用户ID，门店ID和rating评分，提取其中的元素封装定义为Rating数据结构。将所有的训练数据进行二八分，即百分之80用来训练，百分之20做测试。设置最大的迭代次数为10次，将矩阵的特征分解为5个，设置RegParam正则化系数为0.01，防止过拟合，即模型训练出的内容过于趋近真实数据情况，一旦真实数据有误差或者出现问题，导致模型的预测结果不尽如人意；欠拟合则相反，模型的预测结果与真实数据偏差太大。当出现过拟合的情况，可以采取增大数据规模，减少特征维度，或增大正则化系数来调整；出现欠拟合采取相反的措施即可。使用fit方法训练模型，将训练好的模型文件通过save方式保存到合适的路径中。为了直观的评测模型，使用rmse参数对模型评测，rmse是均方根误差，预测值与真实值的偏差的平方除以观测的次数，再开根号，rmse参数的值越小，表示训练数据越准确，可以通过调整迭代次数、特征数量或正则化系数来调试rmse的分数。ALS算法进行训练的流程如图5-8所示。

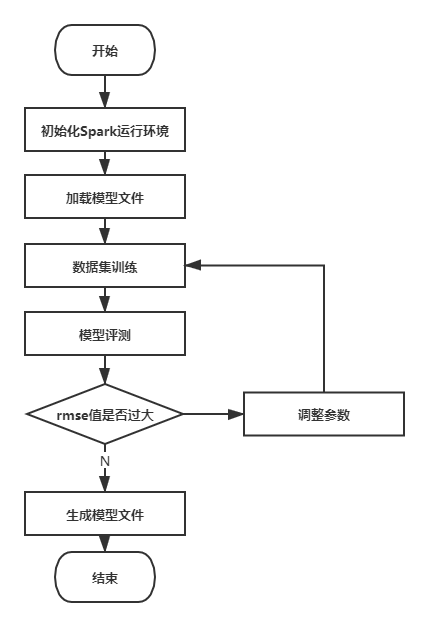


图5-8 ALS算法训练流程图示例

在AlsRecallPredict中进行预测，为用户做预测，并根据内部算法将预测概率最高的前20个门店取出，存入MySQL数据库中。核心代码如下：

Dataset<Row> userRecs = alsModel.recommendForUserSubset(users , 20);

PreparedStatement preparedStatement = connection.prepareStatement("insert into recommend(id,recommend) values (?,?)");

### 5.4.2 LR算法训练实现

LR算法是通过用户历史数据，对未来可能发生的结果的概率进行预测。例如将用户可能被点击的概率看做因变量Y，将用户的ID、年龄、性别以及商品的一些属性，本系统是门店ID、评分、人均价格、是否被点击等特征作为自变量X，LR逻辑回归算法根据这一系列特征，每个特征会占有一个权重，当这些权重进行了逻辑乘和线性加之后，即Y=AX1+BX2+CX3+DX4……，会获得一个概率，这个概率越接近1，就越有可能被用户喜爱，越接近0，表示被用户喜爱的可能性越小。

由于LR算法的计算过程只接受int、double等类型，所以需要对特征进行预处理，将用户年龄、人均价格采用bucket方式编码，分为四个阶段，根据用户的年龄在某一个阶段值就为1，其余为0，价格同理。处理后的结果是一个12维的.csv文件，如下图所示，A-D表示年龄的分类，EF是性别的分类，G是评分，H-K表示人均价格，L表示是否被点击。在LRTrain中对LR的模型文件进行训练，加载特征文件，对其处理，与ALS算法类似，转化为Row的格式，不同的是ALS算法中直接定义Rating的数据结构，转化直接使用Rating.class，LR算法中使用RowFactory的create，将特征放入Vetors向量中。 使用StructType的方式定义列，含有两个field，label和feature，是以上自定义的VectorUDT。对于训练集和测试集同样二八分，LR算法使用的类是LogisticRegression，最大迭代次数10，正则化系数0.3，弹性网络参数ElasticNetParam设为0.8，Family设置为multinomial即多分类问题，可以防止概率值过拟合。训练模型生成model，使用MulticlassClassificationEvaluator方法对模型评价，观察模型的准确率，即预估正确样本的数目和预估错误的样本数目占总样本的百分比。feature.csv文件的数据构成如图5-9所示。

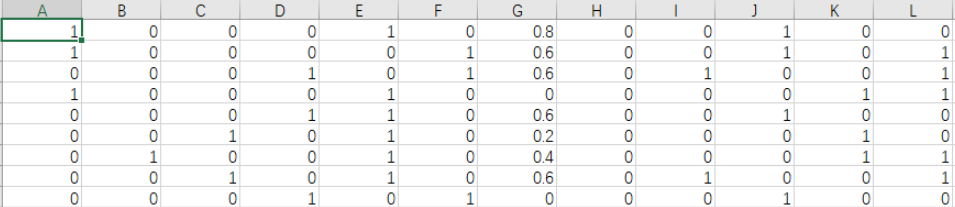


图5-9 feature.csv部分数据示例

### 5.4.3 推荐接入的实现

当系统初次启动未有用户登录时，在首页的前端ajax请求内获取地理位置经纬度，根据SQL的球体表面距离计算公式计算出距离，根据距离和评分因素进行推荐，距离越近评分越高越排序越靠前。门店纬度和所在地纬度的差距，除上纬度的平面，经度的差距除360，做了指数加sin,cos的运算，根据95%因素受地理位置的影响，5%的因素受评分的影响，地理位置以10为对数进行取值，将数值范围控制在10以内，距离越大，参数就会缩小，即参数归一化的操作。门店也进行归一化操作，即除以5。关键代码如下：

select<include refid="Base\_Column\_List"/>,ceil(1+1000\*(2\*6378.137\* ASIN

(SQRT(POW(SIN(PI()\*(#{latitude}-latitude)/360),2)+COS(PI()\*#{latitude}/180)\* COS(latitude\* PI() / 180) \* POW(SIN(PI() \* (#{longitude} - longitude)/360),2))))) AS distance from shop where name like CONCAT('%',#{keyword},'%')

当用户登录后，系统获取到UserModel，获取UserModel中的id，根据id查询推荐表完成访问粗排结果，在RecommendSortService中加载LR模型，使用LR内部的predictProbability方法预测，得到二维数组形式的结果，第一位是为正样本的概率，将正样本的概率值作为打分，并存入新建的ShopSortModel，将预测结果进行二次排序比较，将分值结果倒序排列，即正样本概率高表示会被用户喜爱的概率高，排在前面。推荐结果页面如图5-10所示，核心代码如下：

Vector result=lrModel.predictProbability(v);



图5-10 猜你喜欢页面

## 5.5 后台管理模块的详细设计与实现

后台管理模块包括品类管理、商户管理、门店管理等功能。系统管理员可以通过指定的邮箱和密码登录系统，在登录首页可以看到品类、商户、门店、注册用户的数量，可以对门店等进行管理操作。管理运营平台方便系统管理员查看管理信息，为门店创建等操作提供平台，友好的界面体验可以避免管理员直接更新底层数据库而感到不便。

### 5.5.1 品类管理功能

品类管理功能是管理员对品类进行管理的功能，管理员可以创建品类，可以查看品类列表等。管理员在创建品类时，可以为品类设定权重，系统会根据设定的权重按照由高到低的顺序排序展示。

在数据库的品类表中，品类的名称在系统中认为是唯一的，为品类创建索引，令索引类型为UNIQUE即唯一，索引方法为BTREE。在获取品类列表时，调用categoryModelMapper的selectPrimaryKey即根据品类ID查询方法即可。对于创建品类，在CategoryServiceImpl接口实现类的create方法中将CreatedAt和UpdatedAt设置为当前时间，并调用categoryModelMapper的insertSelective方法将获取categoryModel，返回根据品类ID查询出的结果列表。需要注意的是，由于对品类名称做了唯一索引，此处需要捕获异常，若品类名已存在则无法重复添加品类。最终在Controller层直接调用Service服务层的方法并传送给前端页面展示即可。品类管理界面如图5-11所示。

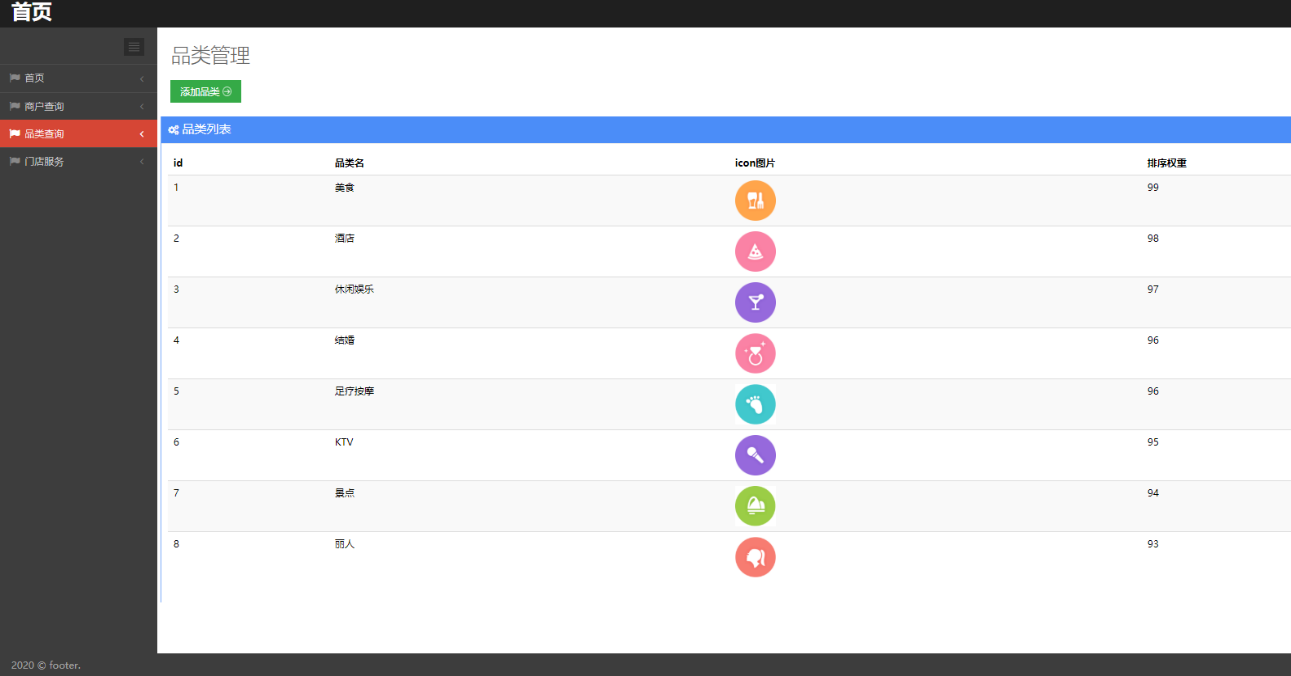


图5-11 品类管理界面

### 5.5.2 商户管理功能

商户管理主要是管理员对商户的创建以及查询，管理员可以启用或禁用商户，被管理员禁用的商户旗下的门店则不可以被展示出来。对于商户的创建以及商户列表的查询，业务逻辑与品类管理相似，不作多余赘述，对于商户的启用禁用，在Service层的changeStatus方法中，将修改后的disabledFlag传入sellerModel。同样在Controller层调用需要的方法更改可用状态即可。

### 5.5.3 门店管理功能

门店管理是对门店的查询、创建等。由于门店列表繁多，此处采用分页展示。具体的查询门店列表以及添加门店的逻辑与以上相似，此处的分页方式使用了PageHelper插件，默认每页显示20条数据。对于经纬度的解析，在ajax请求中，将URL参数设置为百度地图网址，拼接申请的AK参数，启动百度地图服务解析地理位置经纬度。门店管理界面如图5-12所示。

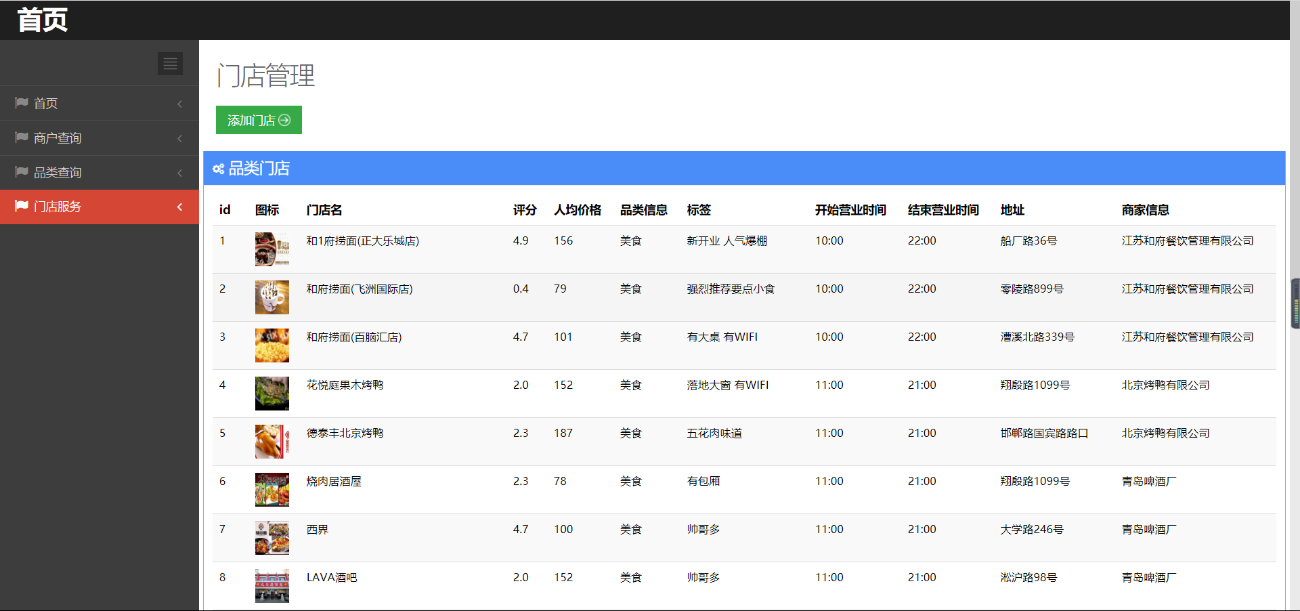


图5-12 门店管理界面

5.6 本章小结

 本章主要描述了系统的详细设计，介绍了用户模块和后台管理模块的主要功能，详细介绍了搜索模块和推荐模块ElasticSearch接入方式的分析、对参数的调整、推荐算法的选择等。

# 第6章 系统测试

软件测试是为了发现软件存在的故障或者缺陷，为了对软件的质量进行衡量，动态的对程序代码中的各种错误和问题进行查找的过程，要最大可能的找出系统的错误。软件测试是在软件含有缺陷和故障的假设下进行的，科学合理的设计最能够暴露问题的测试用例是实现这个目标的关键。对于门店搜索推荐系统来说，验证搜索和推荐的合理性是系统开发的主要目标，而测试要做的就是找出潜在问题，以及不合理的地方，为系统升级改进做出参考[24]。

## 6.1 测试方法

其实在软件开发时期，也有对程序进行调试的过程，对于模块化的系统来说，可以是编好了一个模块后，先在计算机上进行运行测试，发现错误及时纠错，修改后再重复进行测试。在正式的测试阶段，软件测试的方法有很多，常用的有黑盒测试和白盒测试。

黑盒测试是当不了解系统内部逻辑结构时，将被测试的程序看做一个黑盒，对系统的功能进行测试，检查程序的功能是否符合要求。所以对于黑盒测试又可以称为功能测试或数据驱动测试。主要测试的内容是发现输入能否正确接收，能否正确的输出结果，是否存在不正确或者遗漏的功能，是都存在初始化或者终止性的错误，性能上是否能够满足要求[25]。

白盒测试顾名思义，与黑盒测试正相反，白盒测试将程序看做是一个透明的白盒，对程序的内部结构及处理过程可见，对程序内部的逻辑结构进行测试，检查程序是否正常工作。白盒测试又称为结构测试或者逻辑驱动测试。

本章对门店搜索推荐系统进行测试，主要使用黑盒测试的方式，为系统安排测试方案，设计测试用例，对系统进行功能测试，并根据测试的结果找出系统需要改善的地方，提出解决方案并对系统进行优化。

## 6.2 系统测试用例

对门店搜索推荐系统通过黑盒测试的方式进行测试，并设计测试用例。此处主要对搜索模块和后台管理模块进行测试。对搜索模块的根据关键词、分类、排序、标签、同义词、相关性搜索设计测试用例进行测试，对后台管理模块的商户管理、品类管理、门店管理等进行测试。测试用例如表6-1和6-2所示。

表6-1 搜索模块测试用例表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 项目  名称 | 门店搜索推荐系统的设计与实现 | | | 程序版本 | | 1.0.1 |
| 测试  环境 | 硬件环境：Core i7-6500U @ 2.50GHz双核，内存12G DDR3L ，硬盘512G； | | | | | |
| 软件环境：1、操作系统：Windows10；  2、开发平台与数据库：IntelJ IDEA2018，MySQL5.5，ElasticSearch7.3.0； | | | | | |
| 功能  模块名 | 搜索模块 | | | | | |
| 测试  目的 | 验证搜索结果，保证能够显示相对正确符合用户要求的搜索结果 | | | | | |
| 预置  条件 | ElasticSearch服务已经成功启动 | | | | | |
| 用例编号 | | 测试步骤 | 预期结果 | | 测试结果 | | |
| 001 | | 搜索“凯悦”，点击键盘的Enter键 | 显示含有“凯悦”关键字的门店 | | 显示含有“凯悦”关键字的门店 | | |
| 002 | | 搜索“凯悦1”，点击键盘的Enter键 | 显示含有“凯悦”和“1”关键字的门店 | | 显示含有“凯悦”和“1”关键字的门店 | | |
| 003 | | 搜索“和府”，点击键盘的Enter键 | 显示含有“和府”关键字的门店 | | 显示含有“和府”关键字的门店 | | |
| 004 | | 搜索“凯悦”，点击键盘的Enter键，点击搜索结果下拉列表，点击“美食”品类 | 显示含有“凯悦”关键字的美食类门店 | | 显示含有“凯悦”关键字的美食类门店 | | |
| 005 | | 搜索“凯悦”，点击键盘的Enter键，点击搜索结果下拉列表，点击“酒店”品类 | 显示含有“凯悦”关键字的酒店类门店 | | 显示含有“凯悦”关键字的酒店类门店 | | |
| 006 | | 搜索“和府”，点击键盘的Enter键，点击搜索结果下拉列表，点击“低价排序”规则 | 显示含有“和府”关键字的门店，且按照价格由低到高排序 | | 显示含有“和府”关键字的门店，且按照价格由低到高排序 | | |
| 007 | | 搜索“凯悦”，点击键盘的Enter键，点击搜索结 | 标签筛选列表中显示含有“凯悦”关键字所有门店的所有标 | | 标签筛选列表中显示含有“凯悦”关键字所有门店的所有标 | | |

续表6-1

|  |  |  |  |
| --- | --- | --- | --- |
|  | 果下拉列表，点击标签筛选其中一个标签 | 签，并在标签后显示含有该标签门店的个数，点击其中一个标签，显示含有该标签的门店 | 签，并在标签后显示含有该标签门店的个数，点击其中一个标签，显示含有该标签的门店 |
| 008 | 搜索“凯悦”关键字的同义词“红桃”，点击键盘的Enter键 | 显示含有“凯悦”关键字的门店信息 | 显示含有“凯悦”关键字的门店信息 |
| 009 | 搜索“休息”关键字，点击键盘的Enter键 | 显示所有与“休息”关键字有相关性的门店即所有酒店等信息 | 显示所有与“休息”关键字有相关性的门店即所有酒店等信息 |
| 010 | 搜索“凯悦休息”关键字，点击键盘的Enter键 | 显示含有“凯悦”关键字和与“休息”关键字有相关性的门店信息 | 显示含有“凯悦”关键字和与“休息”关键字有相关性的门店信息 |
| …… | …… | …… | …… |

表6-2 后台管理模块用例表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 项目  名称 | 门店搜索推荐系统的设计与实现 | | 程序版本 | | 1.0.1 |
| 测试  环境 | 硬件环境：Core i7-6500U @ 2.50GHz双核，内存12G DDR3L ，硬盘512G； | | | | |
| 软件环境：1、操作系统：Windows10；  2、开发平台与数据库：IntelJ IDEA2018，MySQL5.5，ElasticSearch7.3.0； | | | | |
| 功能  模块名 | 后台管理模块 | | | | |
| 测试  目的 | 测试对系统的商户、品类、门店等的管理性能 | | | | |
| 预置  条件 | 管理员账号登录成功 | | | | |
| 用例  编号 | 测试步骤 | 预测结果 | | 测试结果 | |
| 001 | 点击“商户查询”按钮 | 查询出所有商户的信息 | | 查询出所有商户的信息 | |
| 002 | 在第一条商户信息的操作中点击“禁用”按钮 | 显示提示信息“确定要禁用商户吗？” | | 显示提示信息“确定要禁用商户吗？” | |
| 003 | 在第一条商户信息的操作中点击“禁用”按钮，并确定禁用商户 | 显示提示信息“禁用商户成功”，“禁用”按钮变为“启用”按钮 | | 显示提示信息“禁用商户成功”，“禁用”按钮变为“启用”按钮 | |
| 004 | 点击“启用”按钮 | 显示提示信息“确认要启用商户吗？” | | 显示提示信息“确认要启用商户吗？” | |
| 005 | 点击“添加商家 ”按钮， | 跳转回商家列表并将添加的商 | | 跳转回商家列表并将添加的商 | |

续表6-2

|  |  |  |  |
| --- | --- | --- | --- |
|  | 输入商户名点击创建 | 户信息显示在列表最后 | 户信息显示在列表最后 |
| 006 | 点击“品类查询”按钮 | 显示品类列表 | 显示品类列表 |
| 007 | 点击“添加品类”，输入品类名、图片URL路径、权重，点击“创建” | 跳转回品类列表并将添加的品类信息根据权重显示在列表相应位置 | 跳转回品类列表并将添加的品类信息根据权重显示在列表相应位置 |
| 008 | 点击“门店服务”按钮 | 分页显示所有门店信息 | 分页显示所有门店信息 |
| 009 | 点击“添加门店”，输入地址，点击“解析经纬度” | 将解析出的经纬度信息显示在下方文本框中 | 将解析出的经纬度信息显示在下方文本框中 |
| 010 | 点击“添加门店”，输入所有信息，解析经纬度，点击“创建” | 跳转回门店列表并将创建的门店信息显示在列表最后 | 跳转回门店列表并将创建的门店信息显示在列表最后 |
| …… | …… | …… | …… |

## 6.3 测试分析总结

### 6.3.1 系统的缺陷报告分析

通过对系统的各项功能设计测试用例进行黑盒测试，对系统的各方面缺陷进行总结整理，系统缺陷报告表如表6-3所示。

表6-3 系统缺陷报告表

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 项目编号 | | SleepyR01 | | | 产品名称 | | 门店搜索推荐系统 | |
| 测试人员 | | 孙宇锐 | | | 测试类型 | | 功能性测试 | |
| 测试时间 | | 2020-5-15 | | | 测试地点 | | 东北石油大学 | |
| 序号 | 测试项 | | 子测试项 | 问题描述 | | 缺陷类型 | | 回归测试结果 |
| 1 | 搜索模块 | | 相关性搜索 | 搜索“吃饭”显示信息较少 | | 中等 | | 通过 |
| 2 | 搜索模块 | | 关键词搜索 | 未输入关键词不显示提示信息 | | 微小 | | 通过 |
| 3 | 后台管理模块 | | 门店管理 | 分页功能的下一页无响应 | | 微小 | | 通过 |
| 4 | 后台管理模块 | | 门店管理 | 添加门店时输入格式错误无提示信息 | | 微小 | | 通过 |
| 问题类型说明： （1）严重：导致死机或误删信息或程序无法控制；主要功能未实现或流程缺陷。  （2）中等：一般功能点未实现或流程缺陷； 数据结果错误或界面信息错误。  （3）微小：不影响功能实现的其他缺陷。  （4）建议：建议的改进。 | | | | | | | | |

对缺陷报告表进行分析发现，测试过程中一共发现4个问题，搜索模块出现一个中等缺陷问题，对主要功能没有较大影响，回归测试后发现缺陷进行了改进。系统性能响应方面较好。

### 6.3.2 系统的测试覆盖分析

对于本系统来说，测试用例的设计应该覆盖测试需求，一个需求应该有一个或多个测试用例覆盖，覆盖详情见表6-4所示。测试用例完全执行，并且通过，覆盖率达到100%。

表6-4 系统测试覆盖表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 需求/功能 | 用例个数 | 执行总数 | 未执行 | 未/漏测分析和原因 |
| 功能测试 | 20 | 20 | 0 | 通过 |
| 性能测试 | 1 | 1 | 0 | 通过 |
| 用户界面测试 | 1 | 1 | 0 | 通过 |
| 兼容性测试 | 1 | 1 | 0 | 部分浏览器界面显示不同 |
| 文档审核测试 | 1 | 1 | 0 | 通过 |
| 安全性测试 | 1 | 1 | 0 | 通过 |

测试用例测试执行通过了25个，1个未通过，通过率96%。最终功能性测试用例的效率说明了测试用例具有发现缺陷的能力。对于性能、安全性测试等采用检查的方式测试，并未计入测试用例的效率范围之内。

### 6.3.3 系统存在的问题与建议

门店搜索推荐系统实现了需求分析规定的功能性需求等。界面风格简洁明快，方便用户寻找功能。对于系统存在的问题，在开发的过程中对于细节的处理或者参数的调整可能不是特别完美，虽然不会影响正常的功能和用户的使用，但是可能显示出的结果不够绝对精确，需要继续努力改进优化，以期提高用户对系统的使用体验。

6.4 本章小结

本章主要对系统的测试进行阐述，介绍了黑盒白盒测试方法，并采用功能测试的方式为系统设计测试用例进行测试，并以搜索模块和后台管理模块为例对测试结果进行了分析得出结论。

# 结 论

门店搜索推荐系统的整体架构是采用Java的SpringBoot框架，使用全文搜索引擎ElasticSearch支持搜索功能，Spark框架支持计算，使用IntelJ IDEA作为开发工具，Maven管理jar包，使用MySQL数据库进行数据存储，最终完成了需求分析的要求，本系统将功能划分为四个模块，用户模块、搜索模块、推荐模块和后台管理模块，具体完成以下功能：

用户模块完成了用户的登录、注销、注册、查看首页等功能。

搜索模块完成用户对门店名称关键词的搜索、对品类的分类搜索、搜索结果排序、根据标签搜索、同义词搜索和相关性搜索等功能。

推荐模块是根据推荐算法进行召回排序和逻辑回归，并最终将推荐结果展示给用户的过程。

后台管理模块有管理员登录、商户管理、品类管理、门店管理等功能。

系统的开发过程中还使用Maven项目管理jar包的机制，可以方便的查看jar包冲突等带来的问题，运营平台的前端页面使用了Metronic模板，并且使用thymleaf模板引擎加载模板，使用了jQuery、ajax、BootStrap等前端技术，为了模仿滑动效果，引入了swipe.js。

本系统为用户提供了方便获取门店信息的平台，搜索引擎为用户更精确的搜索信息提供了技术支持，一些二次分类功能可以让用户拥有更加亲切的搜索体验，推荐功能为用户提供了参考信息，以便用户在不清楚门店名称的情况下也可以获取门店的信息。

# 参考文献

1. 杜贺. 大数据分析在大众点评的运用[J]. 西部皮革,2020,42(02):59.
2. 黄际洲,孙雅铭,王海峰,刘挺.面向搜索引擎的实体推荐综述[J].计算机学报,2019,42(07):1467-1494.
3. 王佳安,王可心,李直旭.美食精准搜索与智能推荐平台的设计与实现[J].福建电脑,2019,35(08):27-30.
4. 刘珂嘉.大数据背景下微博在高校管理中的作用分析[J].才智,2019(36):244-245.
5. 吴嘉睿.团购类APP现状与发展探究——以美团网为例[J].新闻传播,2016(03):63-64.
6. 齐欣.基于推特的美国智库专家涉华信息源分析——以新美国安全中心为例[J/OL].情报理论与实践:1-10.
7. Yutian Hu,Fei Xiong,Dongyuan Lu,Ximeng Wang,Xi Xiong,Hongshu Chen. Movie collaborative filtering with multiplex implicit feedbacks[J]. Neurocomputing,2020,398.
8. 吴涛. 推荐系统中推荐算法研究及其应用[D].北京交通大学,2019.
9. Jianshan Sun,Rongrong Ying,Yuanchun Jiang,Jianmin He,Zhengping Ding. Leveraging friend and group information to improve social recommender system[J]. Electronic Commerce Research,2020,20(3).
10. 王博. 基于ElasticSearch的科技型企业数据筛选与分析管理系统[D]. 山东大学,2019.
11. 许文龙,胡金霞. 基于Elasticsearch的高校评教大数据分析平台的设计与实现[D]. 科技资讯,2019,17(35):10-12.
12. Praveen M Dhulavvagol,Vijayakumar H Bhajantri,S G Totad. Performance Analysis of Distributed Processing System using Shard Selection Techniques on Elasticsearch [J]. Procedia Computer Science,2020,167.
13. 孟峰,荆诚.全文搜索引擎在煤矿运维系统中的应用[J].工矿自动化,2019,45(03):103-108.
14. 史宇. 基于深度学习的中文分词方法研究[D]. 南京邮电大学,2019.
15. 潘春华. 基于ELK的高校数据中心运维多层面监控平台研究与实践[J].中国教育信息化,2020(07):93-96.
16. Mrs. I. Nivetha,Mrs. J. Preethi.Big Log Search and Analysis Using Elk Environment [J]. Middle-East Journal of Scientific Research 24 (Techniques and Algorithms in Emerging Technologies): 420-424, 2016
17. 张雪飞,张秉海.基于Spark MLlib的罪犯数据聚类分析[J].科技经济导刊,2019,27(13):18.
18. 徐林.基于Spark MLlib协同过滤算法的美食推荐系统研究[J].吉林大学学报(信息科学版),2019,37(02):181-185.
19. 董银.基于ALS协同过滤算法的个性化推荐研究与应用[J].无线互联科技,2016(06):125-126+142.
20. 李珍,吴青洋.基于Spark平台ALS模型推荐算法的研究与优化[J].电脑知识与技术,2019,15(13):19-22.
21. 刘黎志,彭贝.Spark集群中还贷问题的逻辑回归模型研究[J].武汉工程大学学报,2020,42(01):113-118.
22. 张振球. 基于SpringBoot的校园文章发布系统设计与实现[J]. 无线互联科技,2019,16(23):41-43.
23. Jhon Masso,Francisco J. Pino,César Pardo,Félix García,Mario Piattini. Risk management in the software life cycle: A systematic literature review[J]. Computer Standards &amp; Interfaces,2020,71.
24. 苏春莉,曹新淼.如何设计好的测试用例[J].价值工程,2020,39(03):219-220.
25. Engineering - Software Engineering; Reports Outline Software Engineering Study Findings from University of Texas Dallas (How Does Combinatorial Testing Perform In the Real World: an Empirical Study)[J]. Journal of Technology &amp; Science,2020.

# 致 谢

大学四年时光说长也长，说短亦是时光飞逝。这四年，有幸福，有快乐，有遗憾亦有不舍。冬末春至的这一次突如其来的疫情更是让我们不知所措，无法回到朝思暮想的大学校园，无法见到亲切的老师，也无法见到可爱的同学们。但是疫情只会阻挡我们的脚步，阻挡不了我们互相惦念彼此的心。回首四年，我看到了大一青涩幼稚的自己逐渐成长为大四成熟稳重的学姐，我得到了蜕变，亦心怀感恩。

无论是过去还是现在，我都要感谢我所经历的一切，感谢我身边给予我鼓励的每一个人，感谢我的指导教师薛继伟老师，老师不厌其烦的督促教导我们，为我们指出错误和不足之处，提出建议和修改方案，使得我们收获了许多无论是知识层面还是心理层面。在这里，非常想要对继伟老师说一句，老师您辛苦了。还有我身边的朋友们，李丽同学、姜岩松同学、丁凯悦同学在我感到焦虑时安慰和鼓励我，在我开发过程中遇到困难时，徐晨同学和黄山奇同学帮我远程调试了程序，技术上不懂的问题白松凡同学帮我一一理清，还有许多帮助过我的人，因为有了他们的帮助，在我遇到困难时能够尽快的克服。

在准备毕业设计的这段日子里，我收获了许多知识，学习到了以前没有接触过的技术，更加的体会到了，不断地学习有多么的重要。在这个技术飞速发展的时代，只有紧跟前沿技术，才能在未来的路上有所突破。大学四年的时光美好而又短暂，还来不及告别就要匆匆结束了，也意味着我们的人生都将踏上新的征程。我真切的祝愿所有的同学都可以在未来的道路上砥砺前行，前途光明。