基于Spark技术的气温预测分析系统

设计文档

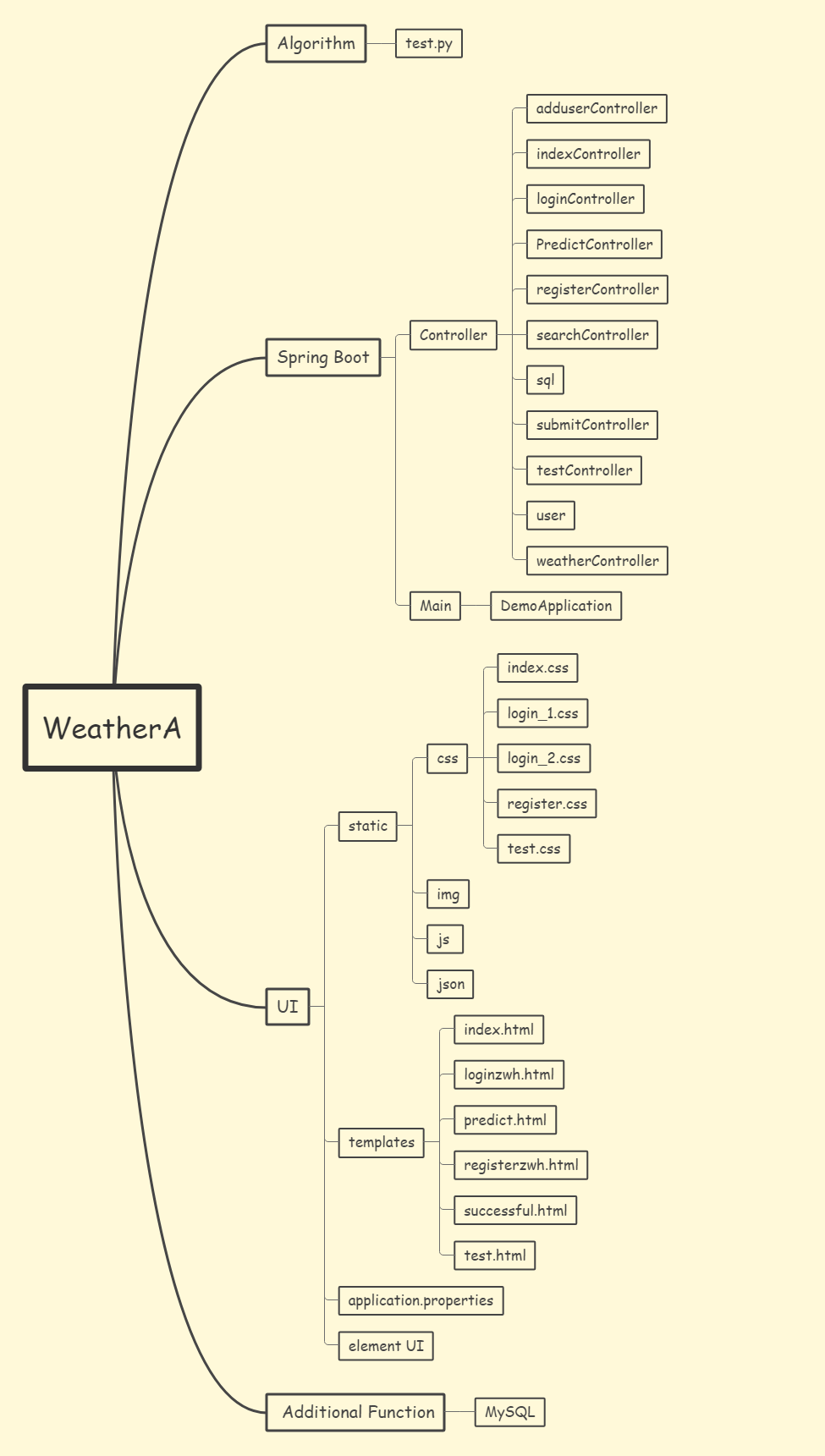
组长：王昱

组员：赵康辉 林小倩

武熠彬 张正阳

1. 系统架构设计

系统架构图



1. 系统概要设计

天气预测分析系统主要分为三大模块，即前端页面设计，框架构建和后端算法设计。

通过前端页面接收输入数据，将数据传递到服务端，服务端调用后端算法进行预测并生成数据传递到前端，在页面上显示出来。

1. 前端页面设计主要包括登录注册页面，天气预测页面，用户管理页面。
2. 涉及到的技术有html，css以及Element UI；
3. 框架构建主要包括服务器的搭建，前端数据接收，向后端传递数据，调用后端函数生成预测数据，将预测数据传递到前端，连接数据库进行用户管理操作。

涉及到的技术有spring boot，MySQL以及axios；

1. 后端算法设计主要包括大数据的下载与清洗，ARIMA模型的构建，生成预测数据，转化成json格式。

涉及到的技术有ARIMA模型，python以及spark。

1. 系统详细设计
2. 前端页面设计
3. Html

属性: class: HTML全局属性，class 属性定义了元素的类名

id: HTML全局属性，id 属性规定 HTML 元素的唯一的 id

style: HTML style 属性,style 属性规定元素的行内样式（inline style）

dir: dir属性规定元素内容的文本方向

1. Css

css样式表的定义

css：（Cascading Style Sheets）层叠样式表；

分类及位置：内部样式-head区域style标签里面

外部样式-link调用

内联样式-标签元素里面

css选择器：

class类选择器可以重复利用

id选择器唯一

标签选择器

选择器：css选择器就是要改变样式的对象

选择器{属性:值;属性:值;}

1. Element UI

使用Vue+element框架制作用户管理界面并实现增删改查

Element框架网址：[https://element.eleme.io/#/](https://element.eleme.io/" \l "/)

1. Vue

使用Vue-axios前后端分离跨域访问的实现

使用Vue+element框架制作用户管理界面并实现增删改查

1. 服务器框架构建
2. Spring boot

利用spring boot构建整个项目的框架。在spring boot中编写控制页面的controller进行数据接收，然后利用函数Runtime.getRuntime( )执行python脚本文件，调用文件中的函数预测未来一周的数据，并生成json文件，利用Bufferedreader接收输入信息流，将得到的数据在java中转换为字符串类型并保存，并使用fastjson包将其转换为json格式的数据，以待传递到前端。

1. MySQL

利用MySQL进行用户管理操作。注册登录用户时，接收前端传递的变量（用户名，密码），将其与数据库中变量进行比对，并提示用户“用户名已存在”“密码错误”等信息；如果信息均无重复则将用户信息保存在数据库中（注册），或是进入天气预测界面（登录）。

1. Axios

利用Axios进行前后端之间的数据传输。Axios 是一个基于 promise 的 HTTP 库，可以用在浏览器和 node.js 中。注册登陆用户与在天气预测界面选择城市与月日时，通过Jquery获得相应控件内的文本值，接着通过Axios将封装好的json文件通过get、post等方式传递，后端在获得相应的路由后即可得到前端传来的数据。

1. 后端算法设计
2. 数据清洗

首先从NOAA网站下载各个省会的天气数据，在使用spark在windows下单节点进行数据清洗，使用python的statemodels、pandas等包筛选所需部分，摒弃无用部分，最后生成csv文件。

1. ARIMA模型
2. 初步设计

首先获得清洗数据，其次根据acf、pacf图和单位根检验确定pdq，在进行残差检验和qq图检验，确定无误后进行预测并输出预测数据

1. 自动化

由于初步设计需要人手工判定，故考虑编写自动化函数实现预测。首先是使用单位根检验原数据和一阶差分二阶差分的数据，确定d；其次手动编写函数，跑pq值从2到7的模型，使用try跳过不可用模型，最后根据bic值确定最佳组合；最后使用最佳的pdq组合进行预测。

1. 最终函数接口

A.编写函数根据城市的不同，读取相应的数据文件；

B.编写一个函数把城市每一天的数据集的pq值都跑一遍并保存，提高后期调用的效率；

C．编写函数从文件中读取相应城市和日期的pq值

D．编写函数生成json数据和预测图

E．编写最终接口函数获得城市和日期参数，并调用上面的函数最终获得预测结果并写入json文件。