暑期学校实验项目：高考志愿填报助手

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 小组名称 | 知识图谱构建B组 | | | | | | |
| 姓 名 | 武逸仙 | 专业 | 人工智能 | 班级 | 091181 | 学号 | 09118130 |
| 实验时间 | 2020.8.31-2020.9.23 | | 指导教师 | 孔祥龙 | | 成绩 |  |
| **一、实验背景和目的**  在中国，高考是一个人一生中的重要转折点之一。然而，大部分的中国高考生都被要求在高考后的三四天内即完成志愿填报。俗话说，“考得好不如填得好”。本项目致力于通过“211工程”高校数据爬虫、知识图谱可视化、机器学习等技术实现一个基于Python的高考志愿填报助手。该助手预通过用户的高考分数、城市偏好、兴趣等特点提供个性化的学校推荐，协助用户高效地完成志愿填报。 | | | | | | | |
| **二、小组任务和个人任务**  知识图谱构建B组以省份这一实体展开知识图谱的构建。  小组任务分为基于Neo4j的知识图谱的构建及可视化和相关的前端设计两部分。  本人参与前期数据库设计及本组主页面和一分一段表格的前端设计。 | | | | | | | |
| **三、个人任务需求分析**  1. 数据库设计  为实现与省份有关的搜索查询需求，为方便后期知识图谱的构建，数据库设计需涵盖省份  信息、分数信息、排名信息等多个实体及关系，力求考虑全面。     1. 知识图谱B组主页面kgB\_index.html前端设计   知识图谱B组以省份为核心实体展开分数、排名、学校等细节信息的精准查询。因而，主页面的作用是提供给用户进入各个省份的入口。由于省份信息较多，同时考虑用户定位兴趣省份的效率和主页面的美观性，我们计划通过地图显示的方式可视化省份入口。    3. 一分一段表格的前端设计  要求表格和UI组的设计风格相符。 | | | | | | | |
| **四、实验过程（需附上关键代码及相关说明）**  1. 数据库设计  基于与小组成员一起总结得到的与省份相关的年份、文理科、一本线、二本线、专科线、一分一段、学校等信息，详细分析其中的各实体和关系，我们将其分为三个关系（表/类）：省份-分数表、一分一段表、省份-学校表。同时定义了每张表具体的属性和主键。如图：    2. 知识图谱构建B组主页面设计  1）主页面的最上栏和最下栏使用了UI组的模版，保持风格的统一。  2）地图部分以https://demodashi.com上基于Echarts的中国地图模版为基础做了以下修改：   * 赋予每个省份固定的色彩值。定义包含各省份的dataList，一个value值对应一个色彩。特别地，我们将南海诸岛注释掉。   var dataList=[  *// {name:"南海诸岛",value:0},* {name: '北京', value: 100},  {name: '天津', value: 100},  {name: '上海', value: 100},  {name: '重庆', value: 100},  {name: '河北', value: 600},  {name: '河南', value: 900},  {name: '云南', value: 200},…]   * 增加各个省份知识图谱页面和主页面的链接。   myChart.on('click', function (params) {  window.open("http://127.0.0.1:8000/kgB/kg\_province/?csrfmiddlewar" +  "etoken=CQ5HrGjyVvSwrUZeiYS4MqB1epb2crZEK1R3m0gddRZ8PDIc0CA405tj" +  "FZttOOaB&select\_province="+params.name);  });   1. 一分一段表格前端设计 2. 主页面的最上栏和最下栏使用了UI组的模版，保持风格的统一。 3. 创建表格结构：<thead>中插入一行两列表头（分数和累计人数）；<tbody>中插入对应内容。   <div class="container">  <table class="table">  <thead>  <tr>  <th>分数</th>  <th>累计人数</th>  </tr>  </thead>  <tbody>  <tr>  <td> 695 </td>  <td> 2 </td>  </tr>  …  </tbody>  <tbody>中的内容使用python读取csv获得。  3）渲染表格   * 斑马线效果：<table class="table table-striped table-boardered">      * 鼠标悬停效果：<table class="table table-striped table-bordered table-hover"> | | | | | | | |
| **五、实验结果与分析**  知识图谱主页面最终效果如下。    通过鼠标选择兴趣的省份，该省份将显示为黄色。点击即进入特定省份的知识图谱页面。    分析：预期达成的前端目标基本实现。但在该页面增加背景后，地图的白色背景显得有些突兀，希望后续能做改进。 | | | | | | | | |
| **六、实验总结与心得体会**  本次暑期实践是我首次参与的多人协作web开发项目。我们小组从最初的没有头绪，经过不断的讨论、学习和试错，在第三周的开始终于明晰我们的任务和分工，并在最终上交了一分具有前后端设计的成品。  在整个过程中，我首次接触了前端设计，基础了解了html、css和JavaScript；首次使用neo4j作为知识图谱的创建和可视化工具，学习neo4j的安装以及如何在终端使用neo4j远程连接数据库，基础了解了Cypher语言。  由于第一次参与web开发再加上人数较多，在整个实践周期中，也显现出了许多问题。  首先，由于web开发经验不足，我们小组耗时两周才真正确定下来小组任务。期间不断试错，导致部分成员耗费时间但却做了一些无用功。从开发效率上来说是不及格的，但是作为初学者，试错经验也是一笔财富。  其次，组内和组间缺乏有效交流。例如地图主页的设计虽然是我和另一位组员的任务，但是最终呈现成品时却发现该页面被增加了背景，但我们并不知情。  最后，项目人数过多导致实际分配到每个人的任务量较少。  总的来说，虽然存在一些小问题，但我们小组成员之间合作非常愉快。我也意识到，无论作为组长还是组员都应该积极主动去做尝试，而不应该被动等待任务分配，否则开发效率会非常低下。唯一不同的是，作为组员，在与组长的合作分工上存在上下级关系，允许组员创新，但也同时需要和组长沟通，只有得到组长的同意的条件下才可以落地组员的想法。另外，这次web开发也初步带我走进了前端的世界，希望自己以后能精进前端技术。 | | | | | | | | |

2020年9月制