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 | 王屹之 | 专业 | 人工智能 | 班级 | 091181 | 学号 | 09118135 |
| 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**  结合本专业知识，运用软件实践课程所学，构建基于Web的高考志愿推荐系统。 | | | | | | | |
| **二、小组任务和个人任务**  小组任务：构建省份信息知识图谱，并在此基础上实现图谱可视化，以及相关信息显示功能。  个人任务：  1. 作为组长，统筹与协调组内工作，分析并安排任务，并与老师和其他业务相关小组进行沟通。  2. 管理本组项目代码，处理组员的request，进行测试，管理项目进度。  3. 分析前端需求，实现特定部分后端接口函数代码。 | | | | | | | |
| **三、个人任务需求分析**  1. 分析小组任务，进行需求分析与系统设计，确定项目技术路线。对大家都不会的地方以及项目的难点，自己要先学，搞清楚到底应该怎么做以后，再清楚明白地表述给大家，使别人对自己需要做什么有一个清晰的认识。之后对任务做合理的划分和派遣。  2. 定期组织小组会议，展示目前项目进度，确定下一步的工作目标和分工。  3. 分别与数据处理组和UI设计组沟通，获取项目所需的数据文件和HTML模板。  4. 后端接口：  1). 已知省份名称，获取省份ID  2). 已知省份名称，获取省内大学列表  3). 已知省份名称，获取分数节点属性字典  4). 已知省份名称，获取分数线及一分一档属性字典  5). 已知省份ID，获取该省分数线信息  6). 获取省份名称-ID对应字典  注：其中 3)、4) 项所提取的字典，目的是为了作为中间变量，供编写节点获取函数和关系获取函数的同学使用 | | | | | | | |
| **四、实验过程（需附上关键代码及相关说明）**  1). 已知省份名称，获取省份ID  def get\_provinceID(province):  *"""  获取省份ID* ***:param*** *province: 省份, string; 例如: '河南'* ***:return****: 省份ID, int  """*  2). 已知省份名称，获取省内大学列表  def get\_collegesInProvince(province):  *"""  获取特定省份内的大学列表* ***:param*** *province: 省份, string; 例如: '河南'* ***:return****: 一个 大学信息字典 的列表  """*  3). 已知省份名称，获取分数节点属性字典  def get\_scoreinfo(province):  *"""  获取指定省份的分数节点属性字典* ***:param*** *province: 省份, string; 例如: '河南'* ***:return****: 分数节点属性字典  """*  4). 已知省份名称，获取分数线及一分一档属性字典  def get\_scoreinfolink(province):  *"""  获取指定省份的一分一档属性字典* ***:param*** *province: 省份, string; 例如: '河南'* ***:return****: 一分一档属性字典  """*  5). 已知省份ID，获取该省分数线信息  def get\_scoreline(provinceID):  *"""  获取指定省份的分数线信息* ***:param*** *province: 省份, string; 例如: '河南'* ***:return****: 分数信息列表，包含该省在数据库内的所有一本线和二本线信息  """*  6). 获取省份名称-ID对应字典  def get\_provincedict():  *"""  获得 省份ID-省份 的对应字典* ***:return****: 省份ID-省份 字典, dict; 例如:{"1": "北京", "2": "天津", "3":"江苏"}  """* | | | | | | | |
| **五、实验结果与分析**  项目后端部分采用Neo4j图数据库进行存储，全数据库共分三类节点，分别为省份、学校和分数信息。下游业务通过事先设计并编写好的接口获取所需数据。因此知识图谱的可视化不必受到数据库原生结构的限制，实现了应用与逻辑结构间的数据独立性。  项目前端部分共分两个页面，其一为省份选择页面（如图1），其二为省份知识图谱可视化及信息展示页面（如图2、3）。  C:\Users\pc089\AppData\Roaming\Tencent\Users\1624735319\QQ\WinTemp\RichOle\Y2X)$7@SJZP4VXKHK)NG5P0.png  图1  在省份选择页面中，界面主题为一中国地图，用户可根据需求通过点击相应省份色块跳转至省份知识图谱页面。界面显示清晰美观，便于用户操作。  C:\Users\pc089\AppData\Roaming\Tencent\Users\1624735319\QQ\WinTemp\RichOle\JEOYS97@O1YW`KFYM%YA(UF.png  图2  C:\Users\pc089\AppData\Roaming\Tencent\Users\1624735319\QQ\WinTemp\RichOle\_[DIJPPC2KES$XSGT}V@}~9.png  图3  在省份知识图谱页面中，上部为可视化后的知识图谱（如图2），下部为近年来的一、二本分数线折线图（如图3）。知识图谱的展示抛弃了默认的矩阵布局，经过调整，采用了可读性更高的同心圆布局。折线图展示了近三年的分数线波动，以便用户对该省的录取情况进行全面准确的了解。 | | | | | | | |
| **六、实验总结与心得体会**  为期四周的软件实践弹指而过，这一路上我们跌跌撞撞，体会不可谓不深。  首先作为组长，在推进工作的过程中我发现，对任务进行合理的分割和派遣能够极大地提升项目开发的效率和品质。这就要求我们必须清楚地知道“我们要做什么”。先通过讨论确定目标结果，再对工作做合理划分，最后再明确各人的任务输出。这样，每个人就可以把精力真正投入到开发当中，而不是在开始工作后陷入迷惑。  第二，课程项目会不可避免地用到大量我们未曾学过的内容，而这就需要一个人先去学习了解。当负责人学完后，再由他将任务分解并分配给大家。或者也可根据个人专长，每人单独负责一个模块，互不干扰，相对独立。这样即可很好地保证开发的效率和大家的参与度。但在这次的实际操作中，常常是负责人学完后，大部分的任务就已经被完成了。这可能是我们的项目体量还不大，代码量相对较小的缘故。  第三，组内成员之间保持良好且及时的沟通是极为必要的。因为项目开发的过程中几乎总会出现这样那样的问题，而当突发情况发生时，有效的沟通可以使问题在一开始就被纳入处置清单。从容而高效，省去了不知多少麻烦。我们对这一点的感受可能尤为深刻。因为我们开会讨论的频率较高，大家坐在一起，商讨项目当前存在的问题和下一步的计划，很多问题都被化解于无形。虽然很多方面可能都不成熟，但这一点却使我们受益颇深。这对我而言，是要记住的宝贵经验。  回顾整个软件实践课程，我收获了许多经验和教益。无论当初做得是否出色，最后都成为了宝贵的历练，或许这就是这门课程的意义。最后，感谢一直带领我们的孔老师，您耐心的指导和无数付出成就了我们在这门课上的收获。回首这四周软件实践，我们受益无穷。 | | | | | | | |

2020年9月制