

**课程 实 验 报 告**

**课程名称： Python程序设计**

**专业班级： 自实1901**

**学 号： U201916457**

**姓 名： 张皓然**

**指导教师： 卢仁智**

**报告日期： 2020年10月17日**

**人工智能与自动化学院**

# 实验1

## 1题目

利用爬虫程序获得大学排名并进行展示。

## 2设计思路与代码实现

### 2.1设计思路

1.利用requests库和beautifulsoup4库获取网站源代码中的信息。

2.按标签索取内容。

3.调整格式，进行输出。

4.设计新的功能。

### 2.2代码实现中遇到的问题以及解决方案

#### 1.从’td’标签中无法获得大学名称

分析网站的源代码，我发现出现这个问题主要是因为大学名称的td标签下面还有a标签和p标签，但省份、类型等td标签下面并没有a，p标签。直接读取大学名称的td标签，实际应该是得到了其下面p标签的内容，导致u.string为None。

对于这个问题，我的解决方案是，在遍历读取td标签后，再特地读取每个tr标签下第一个td标签，td.find(‘a’)，得到大学名称的信息，并加入allUniv中。

#### 2.输出排版错位

在输出排版时，我发现虽然按照书上的格式写了代码，但却无法得到书上的效果。在仔细分析后我发现，这可能是因为u.string得到的字符串存在空格或换行符。

为了解决这个问题，我们使用u.string.strip()除去其中的无关字符就可以了。

## 3实验截图

程序开始后可以输入数字，用来控制输出的大学排名的范围。

|  |
| --- |
|  |

随后我们可以输入大学的名字，如“华中科技大学”，就可以查询对应大学的排名啦。

|  |
| --- |
|  |

## 实验结果

输出结果符合题目要求，并添加了额外的功能，程序设计成功。