**《Python程序设计》**

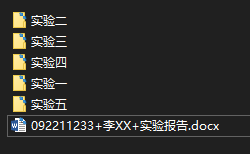
**课程实验**

**一、实验要求：**

1. 操作系统可为下列Windows（V7,V10），Linux（Ubuntu，Centos），MacOS之一；
2. 安装Anaconda，python版本不低于3.5；
3. 开发环境可选择Visual studio .Net，Visual studio Code，jupyter，Spyder,VIM等；
4. 包含库有numpy，matplotlab，pandas，scikit-learn, Pytorch,Beatifulsoup,urllib, jieba, wordcloud等。

**二、实验形式：**

1. 每人单独完成；
2. 实验结束后，撰写实验报告，将所有题目合成一份电子稿报告上交。
   1. 每个人上交的电子稿材料包括：所有题目的程序（按题目分成不同文件夹），一份电子实验报告（**只接受word版本，如下图所示**）



* 1. 将上述材料压缩至一个压缩包，压缩包命名为学号+姓名，例如：092211233+李XX.rar或者是zip包



* 1. 由老师指定一位同学负责收齐材料后再转交给老师。

**三、实验报告形式：**

1. 封面，模板如附件1.；
2. 开发环境描述；
3. 原题目；
4. 题目分析及解题思路；
5. 实验过程（编码过程，以及代码的注释）；
6. 实验结果及实验结果分析。
7. 总结

**四、实验题目：**

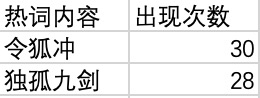
1. 计算1000以内的所有水仙花数，主程序要求使用迭代器和生成器等机制打印输出结果。(5”)
2. m(m>16并且为偶数)个海盗在一条船上，中途遇上风浪，食物短缺，需要扔 m/2名海盗下海。于是海盗们排成一队，排队的位置即为他们的编号。报数，从 1 开始，数到 n(n分别为5或7)的海盗下海。如此循环，直到船上仅剩 m/2名海盗为止。请针对上述问题中的不同的n值分别编制函数，然后在主程序中利用回调机制求解m=40,n=7的海盗下船问题，打印输出留在船上的海盗最初排序的序号。（20”）
3. 1. 制作一个库文件，里面包含至少三种排序模块，比如堆排序、冒泡排序、快速排序，然后通过调用库文件对列表[1,3,5,2,8,7,4]里的元素进行由小到大的排序。（7”）

2. 编辑一个基类，类方法里包含一种排序方法；然后编辑一个类，包含新的一种排序方法，对基类进行继承；再编辑一个类，又包含新的一种排序方法，对前两个类的方法进行继承。调用最后一个类，使用不同方法对列表[1,3,5,2,8,7,4]里的元素进行由小到大的排序。（8”）

1. 1. 上金庸网<https://www.jinyongwang.net/tian/>选择任一部金庸小说，选择某一章，拷贝粘贴到记事本，以“书名第X章.txt”为名保存；（3”）

2. 用流程控制方法统计本章所有段落个数，并将“书名第X章.txt”按照段落总数参照C10code中的testmodule.py方法大致均匀分成5个txt文件，并以文件中第一段中的前4个汉字为名保存；（7”）

3. 结合第12章的知识，利用jieba、wordcloud对5个txt文件中的热词进行提取，并自行选择背景图片绘制5幅词云图，将每个词云图中排名前五的热词内容和出现次数（如下图所示）存储在不同Excel表格中；（15”）

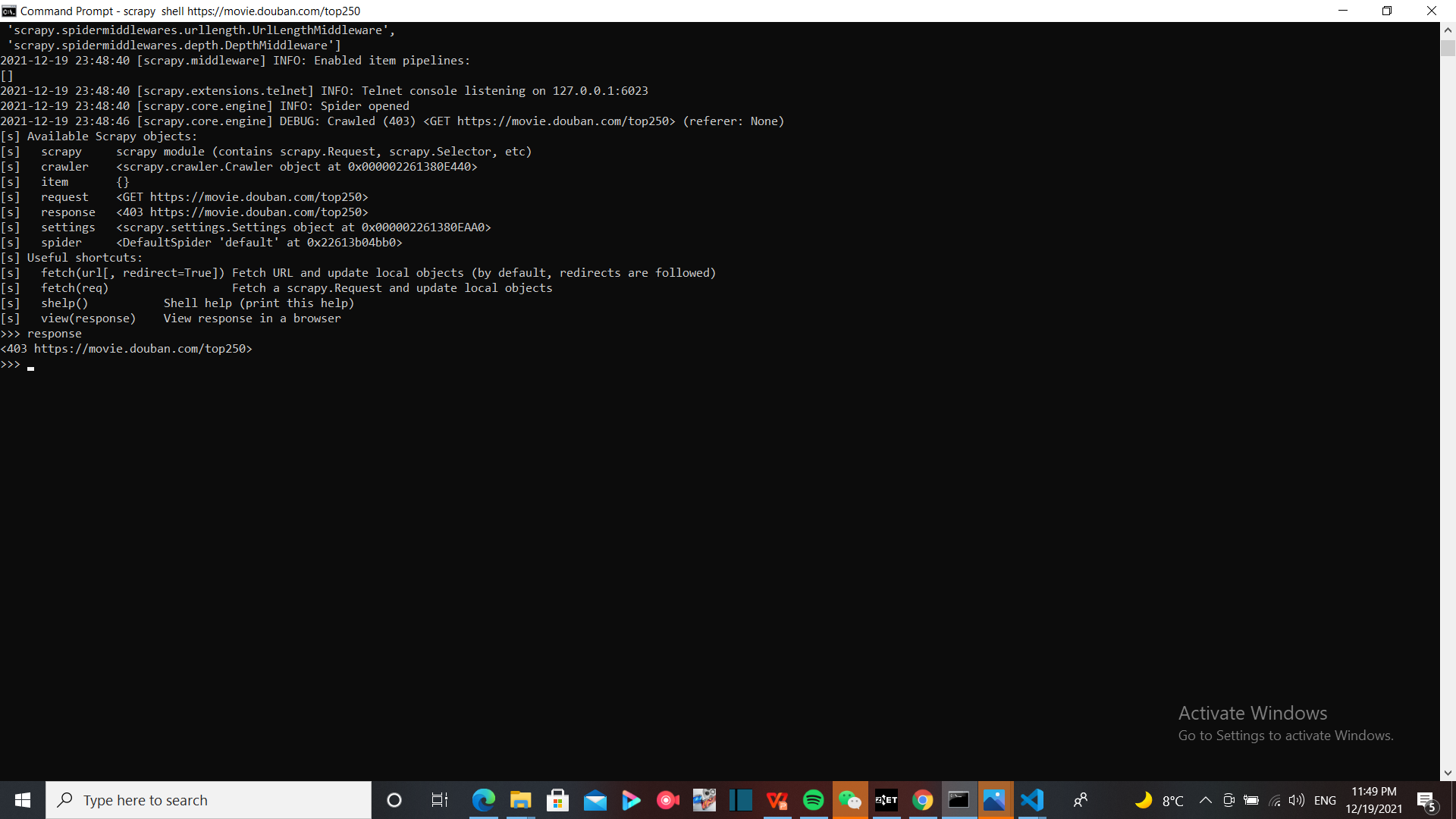


4. 综合5幅词云图，手动选取4个不同的热词，用正则表达式统计4个热词在“书名第X章.txt”中总出现次数。（5”）

实验5. 网络爬虫（30”）

* + 1. 爬取所有豆瓣电影评分Top250的电影的信息(10”)
       1. 正文链接
       2. 英文名（如有），中文名
       3. 等等
    2. 获取每部影片的简介和影评(5”)
    3. 加分项(15”)：
       1. 不限于豆瓣的简介，影评
       2. 是否分析了演员与电影类型的关联关系？
       3. 是否分析了演员与演员的关系？
       4. 是否对简介和影评进行词云分析？
       5. 等等

问题：



解决方案：



附加实验： 机器学习（20”）（通过调用sklearn库实现，上课稍微提过，没做不扣分，做了酌情加分,但总分不超过100”）

1. 实现对获取的电影数据的统计分析（10”）

* + - 1. 可以考虑类型、语言、地区或演员等特征维度
      2. 可以考虑对简介，影评进行语义分析出来的结果进行统计
      3. 绘制相关图形

2. 实现某种分类算法（随意），用于测试某种分类（10”）

* + - 1. 例如通过简介或影评对电影的类型分类；
      2. 例如通过演员的组合分类电影，亦或是反过来；
      3. 其余

**五、实验考核方式：**

1、实验结果（100’’\*0.4）+汇报表现(100’’\*0.3)+实验报告(100’’\*0.3)共计100分，占最终成绩的45%

2、考核方式：

2.1 实验结果：完成上述五（六）个题目，每题分值如题后所示，总分不超过100”；

2.2 汇报表现：能流利回答检查人提出的问题，共计100”；

2.3 实验报告：按照格式（20”），参考文献(10”)，题目分析（25”），过程描述(25”)，结论分析（10”），总结(10”)等共计100”。

**六、其余事项：**

1、实验时间为17-18周，每周二下午5、6节课和周四下午5、6节课，地点为新校区信息楼202、203；

2、实验验收时间和方式在qq群里另行通知；

3、实验报告上交截止时间待定。

**附件1.封面（正式版，删掉此行）**

****

《Python程序设计》

实验报告

|  |  |
| --- | --- |
| 学生学号： |  |
| 学生姓名： |  |
| 指导教师： | 邹逸群 |
| 学 院： | 自动化学院 |
| 专业班级： |  |

20\*\*年\*\*月