

2015 年《大学计算机基础》大作业要求

2015-12-20

作业 1：空气质量指数统计分析

作业参考：随着北方地区雾霾的逐年加重，空气质量指数（AQI）成为了人们经常关注的环境质量指标，借助图表对空气质量情况的统计与分析，能够直观地体现一段时间内空气质量的变化规律。本实验希望同学们综合本学期学到的数据结构、算法以及 Matplotlib 模块的绘图功能，完成对某地市一段时间内的空气质量情况分析。

作业要求：

1、在中华人民共和国环境保护部数据中心的网站中，可以查找到各城市每日的空气情况记录

http://datacenter.mep.gov.cn/report/air_daily/air_dairy.jsp?&lang，收集某个城市近半年至一年的 AQI 指数和首要污染物数据。

2、编写一个程序，存入所有数据，并计算每周平均 AQI 指数，绘出变化趋势的折线图。

3、根据 AQI 指数与空气质量等级的关系，得到每天对应的空气质量等级，用饼状图显示各月以及全部时间范围内空气质量等级的相对比例。

术语和定义：空气质量指数 air quality index(AQI)， 定量描述空气质量状况的无量纲指数。		
0-50	优	级别：一级
51-100	良	级别：二级
101-150	轻度污染	级别：三级
151-200	中度污染	级别：四级
201-300	重度污染	级别：五级
>300	严重污染	级别：六级

4、对各月的首要污染物进行统计，用柱状图显示各月各污染物的出现天数，并统计出近半年至一年首要污染物的前三位，用柱状图显示。

5、设计 GUI 界面，使用者可以在界面中通过点击按钮或输入查询条件，查询到 3、4 问题中所列举的统计信息。（**不完成本项任务者不能参加申优**）

6、课程学习总结：课程收获、难点分析、教师授课评价、助教评价、课程进一步改进建议。

作业 2：RSA 基本算法的实现

作业参考：

1、课程中关于 RSA 算法的描述

2、Matplotlib 资料

作业要求：

1、根据 RSA 算法，设计一对强质数，位数不少于 4 位，确定加密使用的密钥和公钥。

2、设计加密和解密程序。

3、设计 GUI 界面，实现明文输入，密文显示，以及解密后明文输出，以及加密、解密的操作。（**不完成本项任务者不能参加申优**）

4、作业报告要求给出实现方案、程序流程图、程序算法分析和结果分析。

5、课程学习总结：课程收获、难点分析、教师授课评价、助教评价、课程进一步改进建议。

作业 3：地图导航的实现

作业参考：

1、课程中关于地图的数据建模以及算法描述

2、地图最短路径算法，包括 Floyd 算法、Bellman 算法、Dijkstra 算法（选用该方法不能申优）

3、Matplotlib 资料、NETWORKX 资料

作业要求：

1、选取以沙河校区为中心的一副地图，路径出发点为北航沙河校区东门，

终点自行确定，要求顶点不少于 20 个、路径不少于 30 条。

2、根据实际地图，确定抽象数据模型及其结构，形式化描述地图。

3、采用合适算法，完成最短路径的寻找，并给出路径，路径权值可以包括距离、时间，如果同时计算出距离最短和耗时最少路径，加分。

4、设计 GUI 界面，将地图用图进行描述，采用有向图，描述其中的顶点和路径，以及里程的权值，要求顶点为中文，并在图中采取差异颜色标注出最短路径的结果。

5、与百度导航的地图进行比较。

6、课程学习总结：课程收获、难点分析、教师授课评价、助教评价、课程进一步改进建议。

作业 4：聚类中的经典算法 K-均值（K-means）的实现

作业参考：

2015 年 11 月 23 日至 25 日，在国家会议中心（北京）举办了世界机器人大会。习近平主席为大会发来贺电，指出以机器人科技为代表的智能产业蓬勃兴起，成为现代科技创新的一个重要标志，中国将机器人和智能制造纳入了国家科技创新的优先重点领域。

人工智能（AI）是机器人的技术核心，机器学习（Machine Learning）又是人工智能的一个重要研究分支，主要研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。机器学习的经典基础问题是分类与聚类问题：分类是根据一些给定的已知类别标号的样本，训练某种学习机器，使它能够对未知类别的样本进行分类，这属于 Supervised Learning（监督学习）；而聚类指事先并不知道任何样本的类别标号，希望通过某种算法来把一组未知类别的样本划分成若干类别，这在机器学习中被称作 Unsupervised Learning（无监督学习）。

本实验内容围绕聚类中的经典算法 K-均值（K-means），分为以下 4 个步骤：

1. 通过查阅网络资料，理解聚类的概念和 K-means 的基本原理。
2. 在二维坐标系下，分别以（1,1）和（2,2）为圆心，以 1 为半径，随机生

成 100 个属于两个圆范围内（不一定在圆周上）的点，将这 200 个点的坐标，存储在一个文件中（存储格式自定义），作为测试数据集备用。

3. 设计并实现自己的 K-means 算法（画出算法流程图，并写出源码），并用实验内容 2 中的数据，测试算法的正确性，最终结果用图形表示（散点图）；

4. 从 UCI 数据集中（<http://archive.ics.uci.edu/ml/>）下载并整理真实的 Iris 数据集，使用自己的 K-means 算法对 Iris 进行聚类。

作业要求：

- 1、给出算法的实现的构思和流程图。
- 2、给出聚类率以及错误样本分析。
- 3、用散点图将结果进行图形显示。
- 4、设计 GUI 界面，提供数据加载、聚类错误样本显示（列表）、绘图（按钮）等展示手段。（**不完成本项任务者不能参加评优**）
- 5、课程学习总结：课程收获、难点分析、教师授课评价、助教评价、课程进一步改进建议。

关于大作业的说明（所有同学）

- 1、从以上题目中任选一道题目进行实现，要求独立完成。
- 2、作业不允许抄袭，一旦发现抄袭，则取消成绩。
- 3、请仔细阅读实验要求，实现程序功能，并认真、详细地撰写实验报告。务必在**作业提交截止时间**（即 **2016 年 1 月 10 日 23: 00**）之前提交大作业到课程中心作业栏目的“**大作业**”。

具体提交要求：请以你的**学号**起名创建一个文件夹（不要有中文字符!），将实验报告、程序源代码和其他相关文件放入该文件夹中，压缩后提交。

课程申优通知（申优的同学）

1、课程申优的同学，除提交大作业到课程中心作业栏目的“大作业”外，还需**额外**提交电子版材料与纸版材料，并参加申优答辩。

2、电子版材料提交要求

申优截止日期（**2016 年 1 月 10 日 23: 00**）之前，在课程中心作业栏目的“**课程申优报名**”，提交如下材料：

（1）大作业的实验报告（**Word 文档，不接受其它类型文档**）（文件名为你的学号）

（2）大作业的答辩 **PPT**（文件名为你的学号+作业题号，如：“**15041020_作业 1**”）

（3）大作业的完整源代码（.py 文件）（文件名形如 **work1.py**）

（4）答辩有关的其他材料（你认为有必要的材料或者是程序运行必须的文件）

（5）“**2015 年《大学计算机基础》大作业申优报名表**”（Excel 表，文件名的括号中为你的学号）

此表一定要填！空表见附件。

3、纸版材料提交要求

在答辩当天提交按照下列顺序装订好的材料（订书机订一下即可）

- 2015《大学计算机基础》大作业实验报告_封面（见附件“2015 年《大学计算机基础》大作业实验报告_封面.doc”，填写学号等信息）
- 大作业的实验报告正文
- 大作业的完整源代码（此部分作为附录）
- 最后附上一张白纸（以备记录答辩情况等）

4、答辩时间和地点

答辩时间初步定在第 18 周周三晚上（即 **2016 年 1 月 13 日 18:00~22:00**）（请

大家密切关注课程网站，如果有调整（如变更时间等），我们会及时通知）。请提前做好答辩 PPT（尽量图文并茂），答辩时间每人 5 分钟，简单介绍你的作业完成功能、实现方案、设计思路和涉及的知识点、创新点或增加功能、作品效果等，并做现场程序演示。

答辩地点在实七楼机房，具体房间待申优报名截止后根据报名人数确定，**请密切关注课程网站上通知**。

鼓励同学们创新，采用不同的方法完成设计，或增加作业要求之外的功能。