实践1 Python pandas、sklearn基础实践

1、实验目的

(1) 掌握Python的基础数据结构，掌握pandas、sklearn的基础知识

(2) 掌握数据预处理、数据可视化、人工智能的流程

(3) 学会自己分析给定的数据

2、实验平台

操作系统：Windows或Ubuntu

软件：Python 2.7或Python 3.6搭配Anaconda+Jupyter或Pycharm等

3、实验内容和要求

（一）根据网页内容，练习基础知识

Sklearn：https://www.jb51.net/article/205965.htm

Pandas：https://zhuanlan.zhihu.com/p/393275819?utm\_medium=social&utm\_oi=623585954691158016

（二）按照提示的代码，完成1-1、1-2、1-3、1-4、1-5的人工智能数据分析过程

**将最后截图结果粘贴在文档中：**

1-1：鸢尾花数据集的分析与识别

1-2：汽车数据集的分析与识别

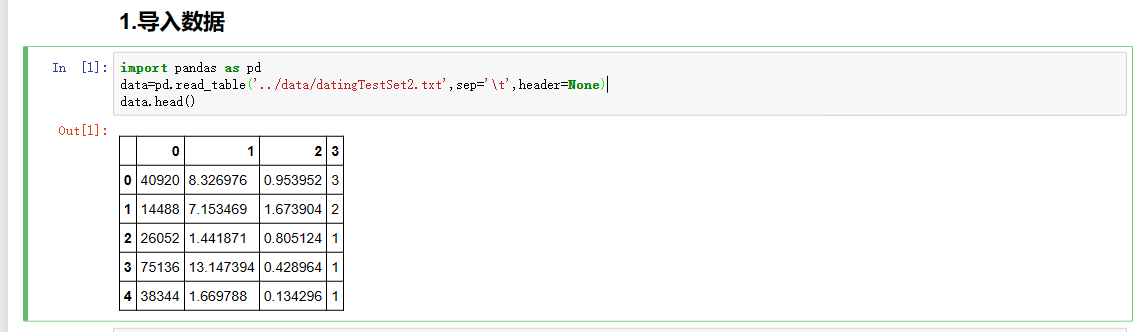
1-3：水泥数据集的分析与识别

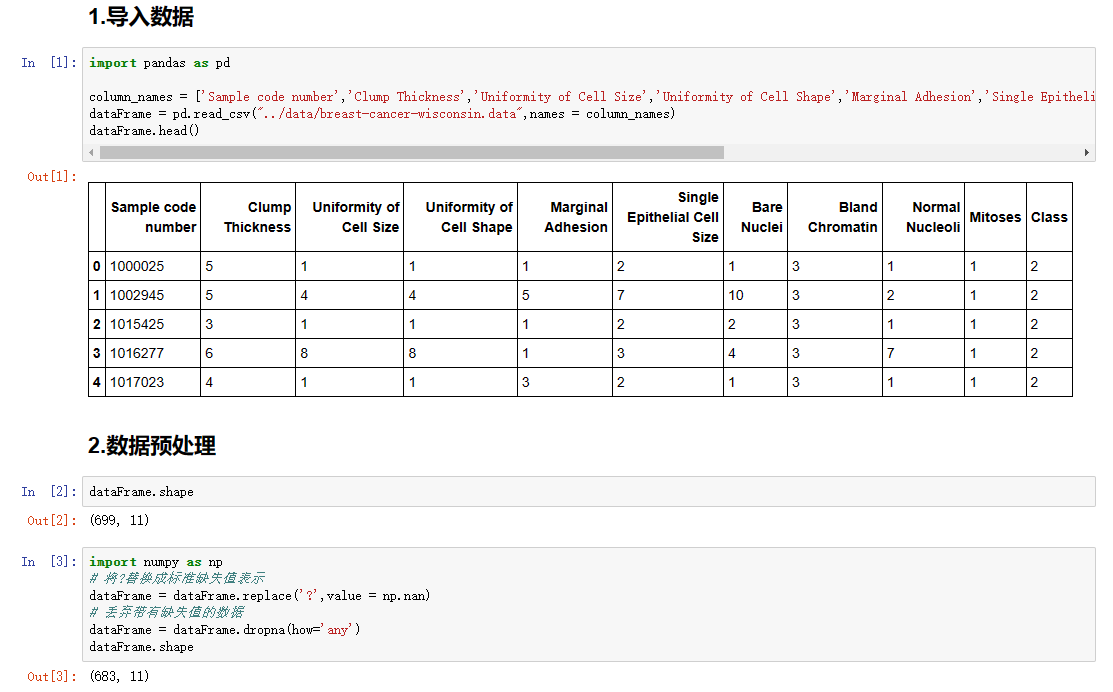
1-4：社交网络广告数据集的分析与识别

1-5：垃圾短信的分析与识别

（三）选做：练习处理给定的数据集

(1) 根据如下的方式读入测试数据集“datingTestSet2.txt”，参考示例代码，测试常用的人工智能算法，如：线性回归、k近邻算法、支持向量机算法等。**将你的运行结果粘贴在下面。**



(2) 根据如下的方式读入乳腺癌数据集“breast-cancer-wisconsin.data”，参考示例代码，采用逻,辑回归实现数据的分类，采用十折交叉验证该数据集的准确率。**将你的运行结果粘****贴在下面。**

思考：给定数据集以后，如何进行数据集读取、数据预处理、训练集和测试集划分，模型选择，结果展示？

# 4.实验报告

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 《人工智能技术应用实践》上机实践1报告 | | | | |
| 题目： | Python pandas、sklearn基础实践 | 姓名 | 周奕帆 | 日期 |
| 实验环境：mac | | | | |
| 实验内容与完成情况： | | | | |
| 出现的问题： | | | | |
| 解决方案（列出遇到的问题和解决办法，列出没有解决的问题）： | | | | |