**一个基于python的小应用：天气播报器**

# 项目简介

Python的出现，使得很多程序代码语言被简化了。而爬虫技术便是其中受益很大的一个。Python使得网络爬虫变得更加简单，因此本次项目我便尝试用Python进行一个爬虫项目的开发。

爬虫技术和目标对象的网页复杂度相关，作为一个爬虫新手，我选择的是复杂度较小的中国新闻网，作为此次项目练习的目标网站。

该项目主要需要实现以下几个功能：

（1）通过爬虫方式，调取中国天气网的全国城市最低气温信息；

（2）选取最低气温前十的城市显示出来；

（3）用可视化的图片将最低气温前十的城市展现；

# 项目逻辑框架及代码实现



图1 项目逻辑框架

**（1）导入**

导入相关的requests、BeautifulSoup和matplotlib等包，为后续爬取网页，并读取信息，作图做好准备；



图2 导入的相关信息

**（2）解析网页**

解析网页时要有三个步骤：一是要将进入该目标网页进行访问；二是要将网页HTML中二进制字符串转化为中文，该步骤需要用到utf-8；三是要将所有的div中的值存在一个table中，下图2、3显示相关的代码。



图3 解析该网页代码

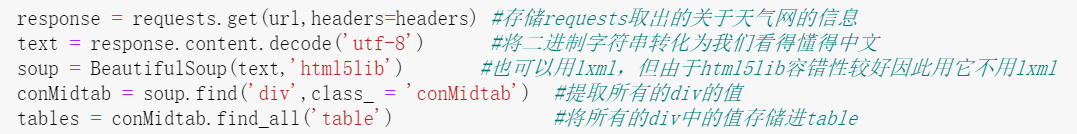


图4 存储相应div代码

**（3）读取并存储所需信息**

* **天气信息所在位置**：在中国天气网中看到这些相应的天气信息都在tr标签中，因此需要取出tr标签。且实际涉及到具体城市的是在第3个tr标签，因此我们从第三个tr标签中开始取数据。
* **具体城市信息所在位置**：本应用需要的是城市、最低温度和天气三者的信息，从中国天气网中看到这些信息存在td标签中，因此我们需要拿到td标签中的数据。



图5 查看所需信息所在位置

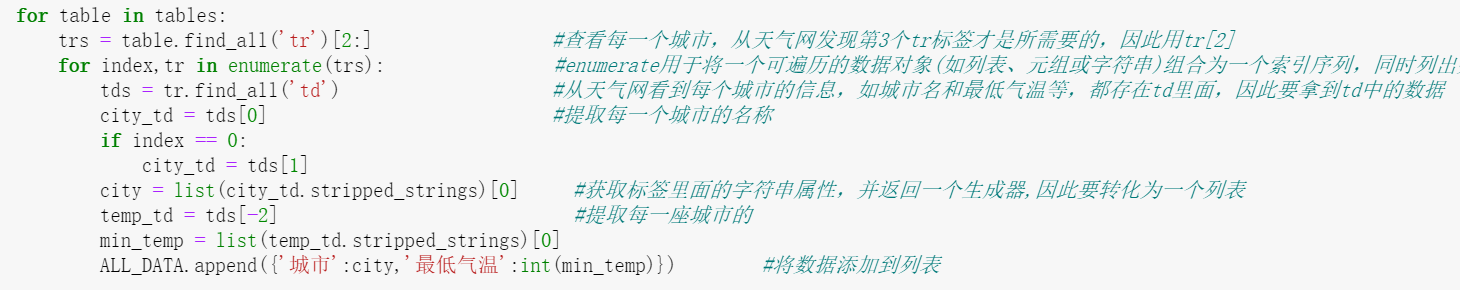


图6 提取城市、最低温度、天气的代码

**（4）展示所取信息**

首先利用sort函数对这些城市最低气温进行一个排序，然后利用matplotlib将前十的数据可视化展示处理。

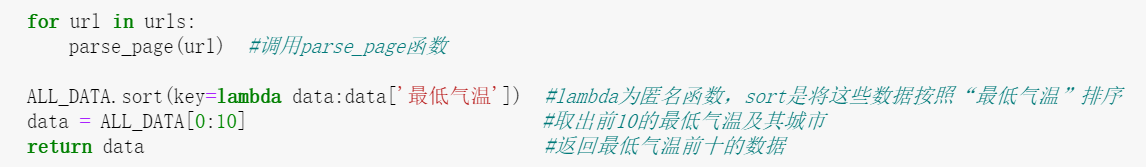


图7 排序的代码

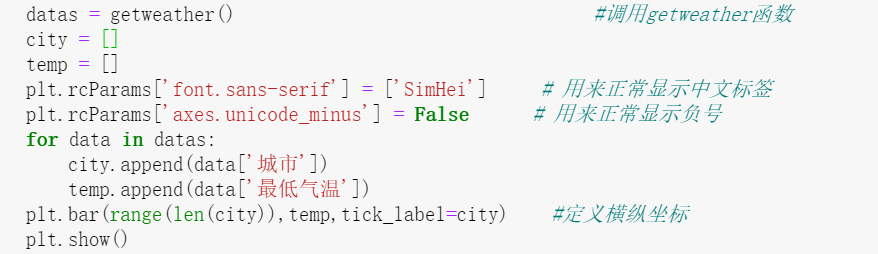


图8 数据可视化的代码

**（5）用GUI界面呈现**

最后用ttk中的Button和Lable作为界面的呈现的主要方式。

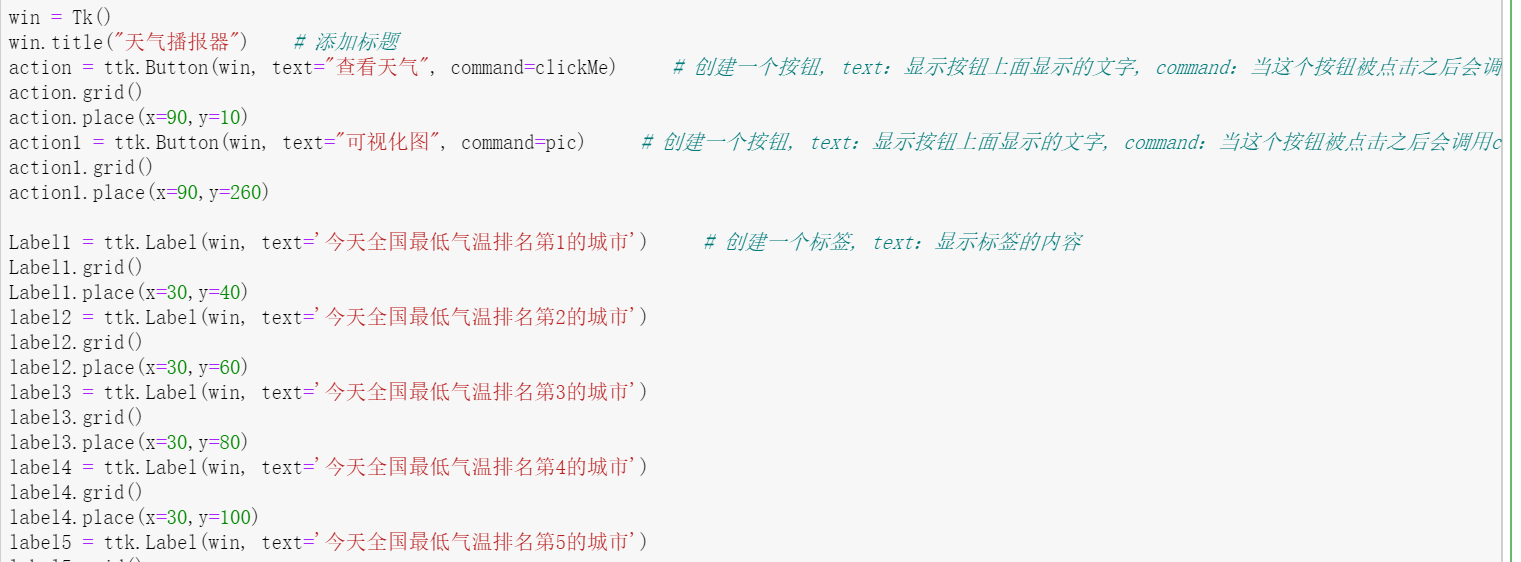


图9 GUI界面代码

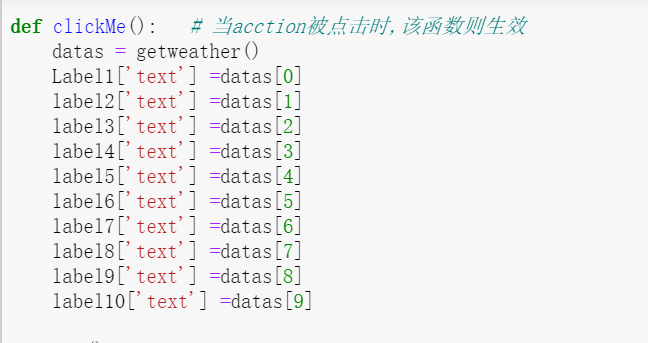


图10 GUI函数代码

# 运行结果

点击“查看天气”按钮，可以看到依次排列的前十的气温最低的城市。再点击可视化图形展示，可以看到给用户一个提示，去代码界面看可视化图形展示。返回代码界面，可发现如图13所示。

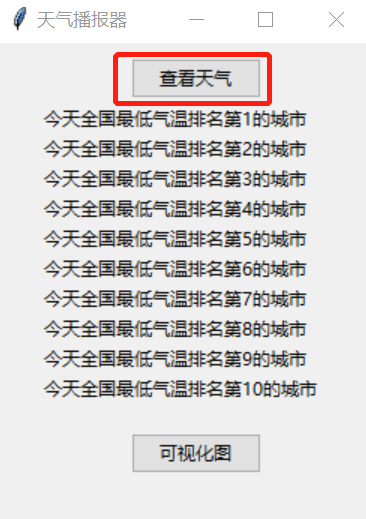


图11 “查看天气”

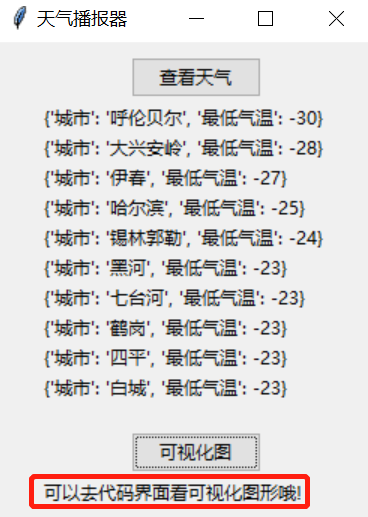


图12 “可视化”

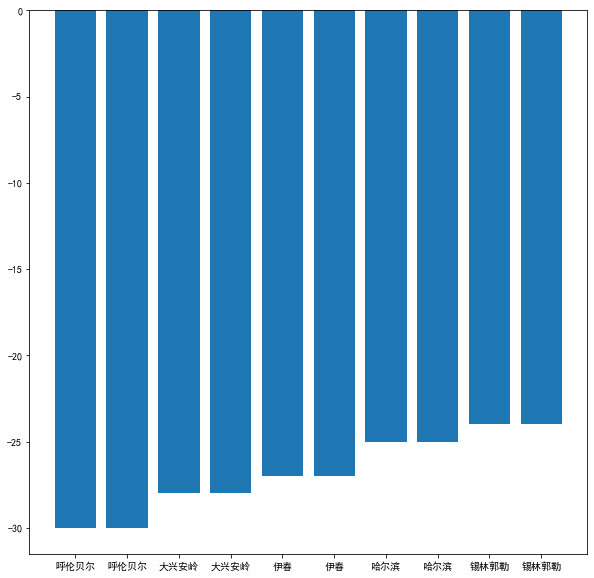


图13 前十的可视化图形

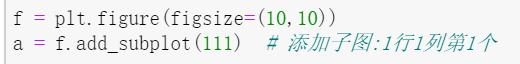
# 心得体会

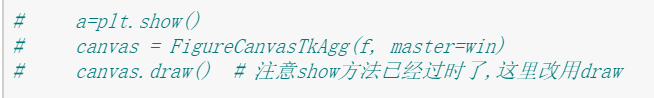
Python是最近几年一直比较火热的语言，作为教育信息技术学系得学生，应该需要对这些新型语言有一定的了解，这也是这门课对我们而言的一个实际意义。爬虫技术也是一直以来每种语言都会去研究的一个东西，python的出现大大简化了爬虫技术，这个课程给了我这种机会去尝试这样一个新鲜的事物，我觉得受益匪浅的。其实有看到教育领域有很多人利用爬虫技术去获取一些教育的数据，对于自己来说，以前可能想不到原来我也是会这么去做的，而且能够做得出来，我觉得开阔了视角，也学习到了很多，更多地是增强了自信吧~总体归纳来说，我有以下几点收获：

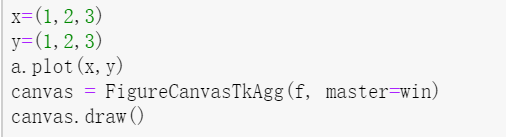
* 通过整个学期的课程，我对python的基本知识有了一定的了解，对python的基本算法、应用工具、应用情境都有了一定的接触；
* 通过期末大作业，我自己百度了解了爬虫的相关原理，并且学会了如何用python画图。同时，在这个过程中，我也整合了课堂上的相关知识，如sort排序算法，ttk界面，def函数等等，并且在通过和其他伙伴的沟通中，也了解到python调取本地时间、调取本地图片的一些方法。

# 疑惑及可改进之处

我本想通过查找资料，看看如何将matplotlib中显示的图片在GUI中显示，也调试了很多次，但都不成功，每次都是可以显示在代码下面的运行板块，但是不能显示在ttk界面中，如果以后有机会的化，希望可以将这个功能再完善一下。







# 参考网站

<https://blog.csdn.net/Bo_wen_/article/details/50868339>

<http://ddrv.cn/a/279633>

<https://baijiahao.baidu.com/s?id=1644927012536144580&wfr=spider&for=pc>

姓名：刘冠群

学号：51194108001

系所：教育信息技术学系