**Python大作业个人文档**

计婷 5140379049

**代码部分：**抓取天猫商品信息，GUI界面设计

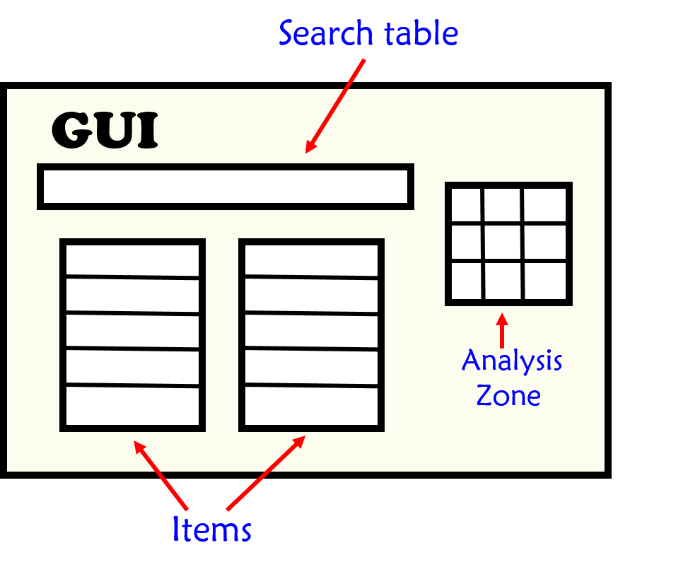
* 第三方库：bs4；urllib；json；requests；PyQt4；
* 实现效果：

后端：

输入是用户输入的商品名称，输出时一个包含商品list的json文件，每个商品中包含商品链接、商品名称、价格、评价、图片、销量六个信息。

GUI：

GUI界面初步设计如图：



其中，菜单栏包含File、Sort、About三部分。其中，File菜单里含有Save（保存搜索结果）和Exit（退出软件）两项，分别设置了快捷键；Sort菜单里含有Price、Rate、Sale，用户可以按照自己的需求选择，点击后会按照相应的顺序呈现出来；About菜单里含有Help（帮助）和About us（关于我们团队的信息）。界面上方是搜索框，输入点击后便可以得到相应的结果。背景、icon都选取了较为简单的图片。

* 实现方法：

后端：

1. 确定商品搜索网址。尝试搜索不同关键词，比较所得网址之间的异同，发现形式是基本相同的，为“http://www.amazon.cn/s/field-keywords=”+商品，所以拼出了搜索页面的URL。
2. 通过浏览器的“审查网页元素”功能，观察天猫网页的html元素信息。发现每个商品就是一个class=” a-section a-spacing-none a-inline-block s-position-relative “的div，所以可以利用bs4的find和get[]功能抓取具体信息。
3. 所有信息抓取完后，利用json可以直接dump()输出dict的信息，把item类改成dict，其中的key分别是source,price,rate,name,pic,sale，其中pic是保存在本地的商品图片名称。

GUI：

1.整体布局：设计了一个继承QtGui.QMainWindow类的Window类，其中包含了设置窗口尺寸、尺寸、背景、icon、名称、菜单等函数，采用QtGui.QGridLayout函数排版，避免了使用绝对位置的麻烦。

2.菜单栏：通过menubar实现菜单的建立，通过addMenu将三个菜单加到menubar，之后运用QtGui.QAction函数将每个菜单下具体的内容加进相应的菜单，其中运用了setShortcut函数添加了快捷键。通过connect函数将菜单的内容和调用的函数联系在了一起。Help和About us菜单运用QtGui.QMessageBox函数将我们想要表达的内容呈现了出来。

* 遇到的问题与解决方法：

后端：

1.拼接url时遇到中文的编码不对，导致url错误，无法抓取信息。原来是中文编码默认用的coding形式是cp936，而天猫的url用的编码用的是utf-8，所以加上encoding=utf-8即可。

2.刚开始尝试的是进入每一个商品的链接抓取具体信息，但是总是报Nontype can't find 的错误，后来查看网页源代码和审查网页元素进行对比后发现，许多在元素中出现的信息在源代码中找不到，说明并不是通过html直接呈现的，而是动态的，和后台有数据传送。然后回到主搜索页面进行对比，大部分信息都是静态的，所以选择直接抓取朱搜索页面的商品信息。但是虽然可以从主搜索页面找到评分页的链接跳转过去，但是评分的具体信息始终是动态的，所以只能选择采用默认值。

GUI：

1.刚开始以为QGirdLayout是按照自己的意愿划分格子，多次改变了搜索框和按钮的位置，仍然没有什么变化，也打不到我想要的效果。后来才发现，是向中间对齐的，因此刚开启的界面搜索框就在正中间，等有item出来了会自动把搜索框顶到最上方。

2.背景尝试了好几张，有的时候会一个界面好几张一样的图，有的时候就只有一张，研究了一下发现，背景图片的大小要和窗口的大小完全吻合才可以正好显示。因此，通过photoshop按照我们的窗口大小重新做了一个背景并运用上去了。

* 代码架构：

后端：

1.goods是5个conv的list

2.prepare函数负责拼搜索页面的url

3. processor 负责抓取，其中调用了5个帮助函数get\_name, get\_picture, get\_rate, get\_price, get\_source分别抓取5块信息

4. save\_img\_to\_file负责把图片存到本地

5. dump负责输出json

GUI：

1. Window类中包含\_\_init\_\_（构造函数），intiUI（初始化函数），menu（菜单函数），help（帮助函数），about（关于我们函数）以及button（按钮函数）
2. \_\_inti\_\_函数调用initUI
3. initUI函数包含设置窗口大小、GridLayout、grid 间距、背景图片、icon和名称，调用了menu和button函数
4. menu函数加了三个菜单栏及每个菜单栏下的内容，调用了help和about函数
5. button函数设置了搜索框和搜索按钮的名称和位置

**收获和感想：**

通过这次大作业，我学习了requests、beautifulsoup、PyQt4的用法，最后能较为熟练的运用它们，同时，也初步了解了亚马逊网站的设计，为将来学习网页制作也打下了基础。这次大作业也是第一次集体完成一个项目，我也从小伙伴们的身上学到了很多。最重要的是，在软件工程专业找到了自己感兴趣的方向——前端设计，为大三分方向定了目标——数字媒体与设计，希望将来也能在自己喜欢的领域不断学习、不断突破！