实现思路

Python爬虫

一、数据获取与存储

1. request、BeautifulSoup

request是python爬虫最基本架包，利用它在加入header能够很好的伪造成浏览器请求，是避免爬虫被封的有效手段之一

BeautifulSoup能够帮助我们快速的找出想要爬取的网页内容

1. time,json

本次爬取的网站有一定的反爬措施，因此使用time在程序爬取一页数据后睡眠一段时间，会避免爬虫被封。

Json帮助我们加载返回的数据，快速查找到想要的信息

1. Pymysql

连接mysql数据库进行数据的存储与获取

二、数据分析

1、snownlp

调用情感分析库snownlp来分析评论，返回0-1的评价，1表示好评，0代表差评，数值越接近1表示越积极，越接近0表示越消极。

2、matplotlib

用matplotlib分别绘制出单个商品的情感分析，和所有商品的情感分析和价格折线图。