081809 预算2000左右

能从https://www.99808.com/products/?t1=206 从这个网站采集商品信息到我们自己的crmeb商城系统吗？

爬出所有商品信息到我们数据库，有发布接口，图片有水印，需要处理水印，图片放阿里oss

081826 预算5000左右

需求挺简单的：

1. 1000万条新浪微博帖子，然后爬每个帖子下面的100条高频回复以及这100条高频回复里面的所有追加回复

1. 爬取逻辑可以从微博的标签下面去爬帖子和回复，然后可以得到一批用户，再去爬这批用户的帖子和回复（至少1亿条，先按1亿条报）

081829.预算500-1000，https://xueqiu.com/6146592061/column 我需要这个页面每篇文章的文章标题、url、发布时间、阅读量、转发数、评论数、点赞数

081830 预算500 用爬虫等工具，看知网，关于一个议题，核心词：读者来信 大概大家关心啥，主要议题是什么，哪个作者写的等等内容

081832.预算 1000，就是给一张植株图片，用深度学习的方法检测到上面的关键点来提取骨架，然后输出关键点所在位置的坐标，时间上不是特别着急，半个月一个月都行，就是最好能在做的过程中偶尔和我说一下过程，因为每周开组会需要汇报进度

081908 预算500左右

1. 代码写出来了很乱；2.需要从代码中提取一部分公共信息简化代码也可用于其他代码中；3.其中SQL入库提取有问题；4.需要额外添加补充一些代码数据内容。主要用bs4编写的



081919 预算300-500 只爬虫小车科四必学一 [http://pc.kuaitongjiakao.com/http://pc.kuaitongjiakao.com/](http://pc.kuaitongjiakao.com/http:/pc.kuaitongjiakao.com/)

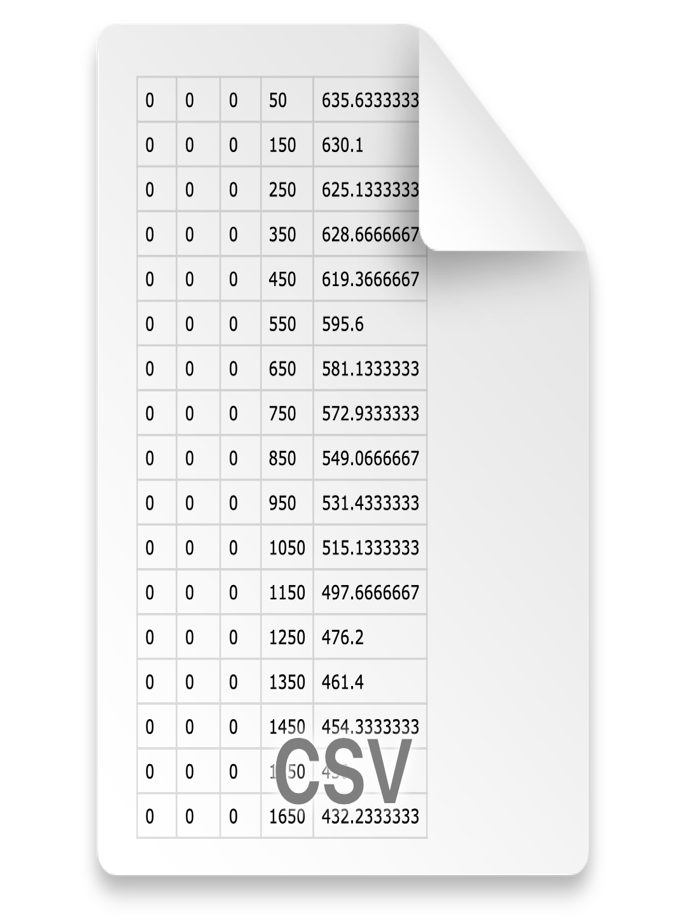
081921.预算500，从这个网站里面，把画横线的信息摘录下来，右上角有搜索栏，我只需要我关注的仪器，https://nrii.org.cn/



081931

预算500，时间2.3天

一个回归问题，4个输入参数，1个预测参数，160多条数据，我用catboost预测效果挺好的，R2=0.988。然后现在就想搞个深度学习模型，随便用什么，效果差不多就行，然后可以迁移到另一个数据集上面，就是要用迁移学习，就是弄好后要教我如何迁移到另一个数据集。



082006.，预算450，二、请运用Python完成如下数据爬取任务（二选一）（20分）

1.请从某一财经新闻媒体（如新浪财经或搜狐财经等）批量获取前15页的财经新闻，包含新闻标题、正文、作者及发布时间等相关信息，并存储至excel文件中。

2.请利用selenium在巨潮资讯网（或东方财富网、和讯网等）中批量获取2021年前5页显示的企业年报。参考链接：<http://www.cninfo.com.cn/new/commonUrl/pageOfSearch?url=disclosure/list/search&lastPage=index>

082007.预算600，爬微博，今天晚上，



082016

预算500-1000

需求，武汉市排名前十的小学4.5.6年纪和，初中一二三单元测试卷，还有期中，期末考试卷，还有模拟考的真题试卷，6年的大概2015-2022年！百度，知乎一类

082018.预算1000，写了个去抓一个网页的数据 但是现在那个网址好像用了ajax，翻页了URL也不变。 <http://www.cardhobby.com.cn/market/search?tid=10106fda-c5d8-4121-ab21-0428534eaf42，需要调试下，或者重新做>

