

# QG 工作室暑期实习生两日结

姓名： 许继元      组别： 数据挖掘      年级： 19 级      日期： 2020.08.04

|                      |  |
|----------------------|--|
| 生活小记                 | <p>8.3 日，上午安装和配置 Redis 数据库，并学了 Redis 数据库的基础内容，例如简单的命令以及数据类型等，接着安装了 Redis 数据库的可视化工具 Another Redis Desktop Manager。下午在 GitHub 上找到一个 star 最多的 IP 代理池，调度程序和 API 服务都编写好了，clone 下来之后配置到本地，以便后续爬虫知识的学习和实践。这两天弄到了阿里云半年的免费服务器，所以晚上计划把中期考核的项目学习着部署到云服务器，本来以为有教程很方便，不过由于自己对很多知识点和概念还是不清晰，所以遇到了蛮多困难，搞到了很晚才在服务器安装完 IIS。</p> <p>8.4 日，上午继续昨晚的任务，在服务器的 IIS 添加完网站，配置好安全组之后，出现了一个很奇怪的问题，在本机访问公网可以访问到前端，但是后台无法运行；而在服务器访问 127.0.0.1 的时候，也是无法运行后台，奇怪的是，在服务器访问 localhost 就一切正常，自己也对 0.0.0.0、127.0.0.1 以及 localhost 的概念不是很清晰，所以就先放着先，目前先完成主要任务先。晚上看了西瓜书的神经网络章节，完成 BPNN 模型的数学推导以及代码实现部分，且在手写字体数据集上测试了 BPNN 模型，并了解了其他类型的一些神经网络。睡觉前抽出时间练了半个多小时吉他。</p> |
| 学习<br>开发<br>比赛<br>概要 | <p>学习上，安装、配置了 Redis 数据库，并学习了 Redis 数据库的基础内容，命令以及数据类型等，并安装了可视化工具。学习了把 Web 项目部署到云服务器，但是遇到了一些问题，目前没有解决。</p> <p>比赛上，看了绿色计算大赛赛组一的解题思路 and 技巧直播回放，感觉今年的比赛内容和去年的差别有点大，今年会用到 ModelArts 这个开发平台，因为是全新的一个框架，所以得加快速度学习了。</p>   |

|                      |   |
|----------------------|---|
| <p>感想收获</p>          | <p>以前爬虫遇到的一个麻烦问题之一就是ban代理，这次配置好代理池之后，后面写爬虫应该会好点，不过由于这个代理池是去免费代理网站抓取的，所以效果应该没那么好，爬取中小型数据应该还算可以，但是爬取大型数据，可能还是得添加付费代理。</p> <p>对绿色计算比赛的时间和精力投入得不够多，接下来要好好准备。</p> <p>数据挖掘理论学习的进度有点慢，要加快速度了，在完成任务的同时要记得跟上文档、日结、周结和实验报告的编写，不然会越拖越慢、越拖越多。</p> |
| <p>存在问题<br/>(备注)</p> |   |