

QG 工作室周记

姓名： 许继元

组别： 数据挖掘 年级： 19 级

周次： 第 3 周

生活随记

这周的主要任务是完成中期考核项目，在多维数据可视化方面，由于之前没有系统学习 `Matplotlib` 和 `Seaborn` 等绘图库，因此花了一点时间查阅资料寻找多维数据的可视化方法，一共找到了平行坐标、`RadViz` 雷达图、`Andrews` 曲线这三种可视化方法。不足之处是这三种方法只能处理输出变量为离散值的情况，也就是分类问题，而对于回归问题，这三种方法不可行。

27 日抽出时间完成数学建模校赛的第三题，也就是电影票房模型的构建，最后得出的结果还不错，预测两部电影的票房也在正常范围内，模型平均误差达到 3.04 左右，接着就把论文剩下的部分补充完毕，提交论文。

28 日和 29 日陆续提交了中期考核项目的原始数据集分布、机器学习算法以及数据挖掘结果可视化。另外这周也有花时间继续学习 `Web` 开发的教程，边学边做出一个简单的微博客系统。

由于完成了中期考核项目算法部分的提交，所以抽出时间学习了 `GBDT` 这一机器学习领域中浅层模型的优秀模型，也是各大数据挖掘比赛中经常出现的框架，以及使用 `Latex` 完成其数学推导并记录到 `CSDN` 博客上。

考核 `Deadline` 的最后一个晚上，我花了两个小时那样做项目演示视频，以前就想过好多次学习 `PR` 视频剪辑，但是一拖再拖，这次终于有机会学习了，果然 `Deadline` 是第一生产力，我花了两个多小时就学会了基本操作并且做出了演示视频，虽然还是有很大的进步空间，不过对于这个制作速度我觉得还是可以的。

8.2 日下午进行了中期考核项目答辩，这是我们第一次多人协作完成项目，又由于是线上合作，所以还是有许多不足之处，从需求分析和任务分配以及拓展性等方面都需要改进，不过最后师兄的鼓励还是给予我们很大的信心，希望下次能做得更好。答辩完之后就整理了一下 `GitHub` 的代码，并编写了 `README` 文件，给中期考核项目画上句号。

吉他方面，这周也在继续着指弹曲的学习，进度虽然不是很快，但是每天坚持练习，也会有进步的。

学习 开发 比赛 情况	<p>学习上，在数据挖掘方面，认识了三种多维数据可视化的方法；在 Web 开发上，学习了 Flask 的拓展模块 Flask-WTF、Flask-Login、Flask-Mail 以及实现文件上传功能。在机器学习算法上，学习了 GBDT 这一各大数据挖掘比赛中经常出现的框架。另外还学会了 PR 视频剪辑，例如视频剪切、添加字幕、视频叠加、场景切换以及 BGM 的处理等基本操作。</p> <p>开发上，完成了中期考核项目的算法模块、协作完成文档模块以及演示视频的制作。把刚学 Web 开发时做的简单服务器慢慢完善成一个微博客系统，并把代码放到 GitHub 上。</p> <p>比赛上，构建了校赛论文的电影票房模型，配合队友完成论文并提交，完成数学建模国赛的报名。</p>
一周总结	<p>这周是中期考核的一周，所以会比上个星期忙碌许多，每天感觉很充实，但有时候也感觉没学到什么东西，可能是没有做好科学的计划吧，不过自己也是一个不喜欢计划约束的人。中期考核项目最后的成果虽然有许多不足之处，不过效果还是不错的。最后也花时间自己配置了一下，试着在本地测试，然后在 GitHub 上整理了一下代码，给中期考核项目画上句号。</p>
存在问题 未来规划	<p>存在问题：</p> <ul style="list-style-type: none">①中期考核项目任务分配不合理②数学建模学习以及绿色计算大赛的准备不足 <p>未来规划：</p> <ul style="list-style-type: none">①简单学习 Redis 数据库②配置 IP 代理池③开始系统地准备数学建模以及绿色计算大赛④练习 CET 6 套题
导师评价	