

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： 2020.07.30

生活小记	<p>7.29 日，对于输出变量为连续值的高维数据集，没有找到一次性将其可视化的方法，于是采用分别选取一个特征与输出变量绘制其散点图分布的方法，就这样解决多元线性回归的原始数据集分布；而对于多元线性回归的数据挖掘结果分布也同样采用上述手段。下午抽出时间学习了 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 完成了 GBDT 的数学推导，并试着完成一个简单的回归问题案例，整个学习和数学推导的过程都记录在 CSDN 博客上。晚上完成了多元线性回归算法的原始数据分布以及数据挖掘结果可视化部分并合并到 GitHub 的 master 分支。睡觉前继续练习吉他指弹曲。</p> <p>7.30 日，由于不确定前端是否能处理多维数据的可视化，所以在原始数据分布上，利用 Python 做了图片备份，如果前端能顺利实现，那肯定比 Python 实现的要美观，毕竟可以使用 JavaScript 实现动态之类的效果。在实现的过程中也和前端组的小伙伴进行了讨论，在数据可视化以及算法部分也给前端的小伙伴讲清楚需求，以便项目的顺利进行。晚上继续利用时间学习 Web 开发，以及中期考核项目代码的修改和优化。睡觉前抽出时间练习吉他，然后听听音乐。</p>
学习 开发 比赛 概要	<p>学习上，抽出时间学习了 GBDT 这一打比赛常用模型，使用 Latex 数学公式完成了 GBDT 的数学推导，并把过程记录在 CSDN 博客上。在 Web 开发学习上，学习了利用线程加速发送电子邮件、利用 MySQL 数据库的一对多关系实现了用户发帖功能、利用 MySQL 数据库的多对多关系实现了用户之间的关注和取关功能。</p> <p>开发上，在中期考核项目中完成了多元线性回归算法模块并合并到 GitHub 的 master 分支。</p> <p>比赛上，和队友一起完成数学建模校赛论文的剩余部分，准备论文的提交。</p>

感想收获	<p>生活上，在家里的日常生活就琐事比较多，而且舒适的环境确实也带来了开发效率的降低。不过在家里也有许多便利条件。总的来说，做好时间管理，还是有很大进步空间的。</p> <p>学习上，抽出时间学习了GBDT模型，该模型的思想不算太难，学习的过程中也算顺利。这两天也有继续学习Web开发的知识，感觉学到了不少。</p> <p>开发上，和前端组和设计组的小伙伴进行了沟通，弄清楚项目需求。自己也完成了算法的初步提交。</p>
存在问题 (备注)	