

旷视实习工作总结

时间：2019/10/23

我所在的岗位是城市大脑的 core 测试部门，主要负责使用 python 书写后台的自动化测试脚本，工作任务更详细的说是，连接后台图片库，调用算法分析图片，得到预测结果后结合标注文件，对结果进行分析，建立相应的评价指标。

刚到公司的时候，我对业务也不熟悉，感谢量哥和明哥对初来乍到的我，耐心地解释介绍工作业务。尤其感谢明哥，因为对于一个刚入职的新人来说，很多新的知识要学，要逐渐了解公司所使用的工具，框架，业务，算法等等丰富的内容。

起初，我在看公司前辈们写过的 python 代码，学习如何调用算法 API 接口，之后，开始学着自己去写脚本，建立虚拟环境，连接后台数据库，爬取图片，调用分析接口，异步处理分析数据，对数据进行清洗整理，结合标注文件和预测文件进行结果分析。

其中，较为繁琐的工作是清洗数据，在这项工作之前，我会和明哥商量讨论，建立起统一标准，比如预测的标签不相同，应当按照标注文件的标签来修改。然后，实验中需要注意很多细节，一不小心就容易出 bug，所以工作时一方面可以提升个人写代码的能力，另一方面，也可以锻炼人的心态。

当初还有一项工作是建立 nDCG 指标,这算是一个新工作，结合维基百科对 nDCG 的定义以及实际业务，自己重新按照分析结果重新建立了 nDCG 评价指标,并且能和明哥结果比对上。

印象还很深刻的是，最后做的一个项目工程，多属性分析。整个工程需要从头开始写，从分析，到数据清洗，到结果分析，最后绘出混淆矩阵，并可视化，展示 false positives。工程还需要实现自动化，比如文件是 sjs 还是 tongren，属性涉及到多个，分类有两种，多分类和二分类，所以编程时也要注意尽量使得代码自动化程度较高。自己也将项目集成到服务器端，在 linux 命令行中输入简单的参数便能一键执行。

上面很多的工作可能对于前辈来说，可能是轻车熟路，小菜一碟，但对于新人的我，还是从中学到很多，收获蛮大。在测试岗位，自己得以有机会更加了解公司整个的运作流程，了解公司的业务、框架和数据库，熟悉工作环境，而不是一开头便扎进了算法圈子中。在写代码时，一方面，自己进一步提升了写 python 代码的能力，另一方面，在工作时我也有逐渐在感悟，工程合作时如何减少人们的沟通成本，如何优化代码运行效率，如何尽力使得代码变得高效自动化和便于使用。

最后想说的是，很感谢量哥和明哥给我这样的机会，也在工作中不断关心我帮助我，我很喜欢公司这样的工作氛围。其实在我来这所公司之前，我就在关注旷视这家别具一格的公司，并且对三位清华创始人的故事颇感兴趣。用 AI 科技改变生活，始终是这所公司创始人的信念和口号，公司走过了 8 年时间，工作氛围自由但人们又崇尚务实，同事间友好而又能相互帮助，这些都是一个优秀公司的基因所在。今年 10 月 9 号，公司上了美国商务部的实体名单，怎么说呢，能让对手感觉到震慑力有时未尝不是一件好事。祝愿旷视这所公司，能肩负他的使命，越走越远。