

西安交通大学研究生专业实践总结报告

提示：研究生在提交总结报告前，请确认已登陆研究生院主页“学生服务”进行了网上登记。

研究生基本情况					
姓 名	张蕊峰	学 院	软件学院	专业学位 类型/领域	专业学位硕士/人工智能
学 号	3118311127	导 师	赵加坤	联系方式	18821715379
实践单位或学科竞赛基本情况					
实践单位名称	西安檬果信息科技有限公司		实践时间	2019 年 10 月 8 日——2019 年 12 月 31 日	
校外导师姓名	耿海滨	职 务	项目负责人	联系方式	13186168766
学科竞赛名称			起止时间	年 月 日—— 年 月 日	
指导教师姓名		学院		联系方式	
▲注：研究生可选择单独填写“实践单位名称”或“学科竞赛名称”或上述两者兼而有之。					
研究生专业实践总结					
单位 实践 或学 科竞 赛的 过程 及成 果	(企业、行业实践或学科竞赛的主要过程及工作成果，不少于 1000 字，可另附页及佐证材料)				
	实习内容介绍： 在这次实习中，我们做的工作是实现一个机器学习平台。机器学习平台是顺丰为 AI 学习者打造的一站式机器学习服务平台。为用户提供从数据预处理、模型构建、模型训练、模型评估到模型服务的全流程开发及部署支持。且机器学习平台内置了丰富的算法组件，支持多种算法框架，满足多种 AI 应用场景的需求。从而解决学习者在 AI 学习过程中的一系列难题，例如教程水平不一，教程和样例代码难以衔接，高质量的数据集不易获取，以及本地难以使用大量数据集进行模型训练。其自动化建模的支持与拖拽式任务流设计使得 AI 初学者也能轻松上手。				
	实习工作及过程： 在这个项目中，我主要是负责数据导入。数据导入要求既应支持线上传数据，又可支持多种异构数据源（HDFS、Hive、MySQL）的集成融合。在可视 UI 界面，用户在前端选择数据导入方式：本地上传数据或者数据源上传：1）本地上传：用户从本地选择文件格式（图像，音频，文本等），将相应的文件夹或文件上传到平台；2）数据源上传：用户可选择设置不同数据源，通过与其兼容的适配器接入平台。对于视频,图像(序列)等多媒体输入数据,平台采用 Ceph 进行对象存储；对于结构化半结构化的输入数据，平台拟采用 HDFS 进行存储，数据导入接口屏蔽了具体的实现细节，可支持不同数据源的接入。计算节点在执行计算时,通过调用数据导入接口导入数据。计算节点在运行代码时直接从数据服务器读取数据,数据服务器根据计算节点请求的				

	<p>刚开始的时候，我不知道如何下手，在老师和学长的指导下，我将自己的任务有分成了两大块：分别是本地上传和数据源上传，其中本地上传又分为文本、音频、视频、图片上传；数据源上传分为 HDFS 上传、MySQL 上传、HIVE 上传。通过商量我们小组使用 Django+VUE 来搭建前后端框架。</p> <p>在第四周，我开始正式做项目。这时，我才发现学到的东西运用到实践中是多么不容易，一个系统中需要考虑特别多的问题，有可能一个小 bug 就会造成巨大的损失。在这个过程中我也遇到了许多的问题，诸如：如何实现跨域，如何实现大文件的上传以及实现不同类型的文件存储。刚开始遇到问题的时候焦头烂额，无所适从，不知道从哪里下手，也不知道需要查什么资料，后来在学长的帮助下，我开始有了系统的分析问题的办法，慢慢的解决问题的速度开始加快。最终在实习中完成了如下工作：在数据上传中，使用 HDFS 进行存储，解决了不同结构的数据上传。在文件上传过程中，通过调研了七牛和腾讯云的切片分段上传功能，利用切片现了大文件上传，解决了上传大量的数据导致整个过程会比较漫长，且失败后需要重头开始上传的问题，提高了上传文件的性能和速度。</p>
个人的收获与体会	<p>（企业、行业实践或学科竞赛的个人收获、对实践单位和学校管理方面的意见与建议等，不少于 1000 字，可另附页）</p> <p>经过这段时间的实习，我觉得让我自己得到了许多的锻炼，这是我们在学校无法体验到的，因为各自的侧重面不同，虽然在学校也有实训课，也有模拟做程序，但是因为知道不是真枪实刀的做，所以心里面是怀有着应付的心态的，这种情况下能够学习到的东西自然就十分有限。而实习工作则不一样，它是真实的，我们的工作都是要对公司负责的，出了错误是要承担责任的，所以让我们在实习中必须要秉持着一定的责任心才行，因此特别能够锻炼我的实战能力。</p> <p>虽然学校里面用的都是国外翻译过来的专业书籍，但是到了工作上面我才发现其实我们学习的技术还是远远不能够满足工作需求的，比如我们在学校学习 Java，可能会教一两个框架，但是到了工作中，很多的公司并不需要这两个框架的程序员，现在流行的是其他的框架，这就是学校与社会的脱节了有一定的距离，所以到了实习上面，没办法必须要跟着公司的脚步走，刚开始的时候自己不会用这几个框架，就只能这前端用用 html，js 什么的，不过好在各种框架大致上还是差不多的，所以约起来也还比较快，所以这几个月自己实习的收获之一就是多会了几个框架。</p> <p>同时我们要有很强的自学能力。计算机知识日新月异，在一个项目中，他们经常被要求在很短的时间内学会用一种新的计算机语言进行开发设计，他们只能在自己以前的基础上努力自</p>

	<p>学尽快掌握知识。在这个信息爆炸的时代，知识更新太快，靠原有的一点知识肯定是不行的。我们必须在工作中勤于动手慢慢琢磨，不断学习不断积累。遇到不懂的地方，自己先想方设法解决，实在不行可以虚心请教他人，而没有自学能力的人迟早要被企业和社会所淘汰。</p> <p>这几个月的实习不仅仅是在技术上让我又不少的感触，在生活和习惯上的收获同样不小。在学校读书的时候自己总会有厌学校心态，明明知道读书是为了自己，可以在用功上还是不会特别努力，总是被动去学习，作业太多了心里也会忍不住埋怨老师。但是到了工作上面，自己会不自觉的去学习，遇到问题了不会像读书时那般不去管，而是主动去向别人请教，就算工作中遇到了加班了，心里有抱怨的同时也还在坚持做着，不会像读书时的自己自暴自弃，我想这就是成熟吧，到了特定的位置上面自己的心态会不自觉的发生转变。</p> <p style="text-align: right;">（签字） 2020 年 5 月 21 日</p>
校外指导或竞赛指导教师对研究生的评价	<p>（从工作态度、出勤、能力、成果等方面进行客观性评价。）</p> <p>。</p> <p style="text-align: right;">（签字） 2020 年 5 月 21 日</p>
实践单位对研究生的评价	<p style="text-align: right;">（部门章） 2020 年 5 月 21 日</p>
校内导师意见	<p>（从能力、成果、收获等方面进行客观性评价。）</p> <p style="text-align: right;">（签字） 2020 年 5 月 21 日</p>
学院审核意见	

（主管院长签字/章）

年 月 日