

田奇

大数据研发工程师 (可全职实习至 2018 年 09 月)

✉ 777tiangoole@gmail.com · ☎ (+86) 131-7724-7672 ·

🎓 教育背景

中国科学技术大学, 苏州研究院 2016 – 至今
在读硕士研究生 软件工程, 预计 2019 年 3 月毕业

华中农业大学, 武汉 2012 – 2016
学士 计算机科学与技术, Top10

👨‍💻 实习经历

美团云, 计算研发 2017 年 8 月 – 2017 年 12 月
Python/Go, CloudComputing, Tornado, RESTful, EBS, Qemu, Glance

- 日常值班负责虚拟机稳定性交付, 错误排查, bugfix, 新增 feature 等
- 负责美团云主机和 MEBS 云存储的对接任务, 完成云硬盘的自定义策略自动快照, 镜像模板跨地域拷贝等功能开发, 完成云平台 REST Cron 服务
- 负责开源镜像管理组件 openstack glance 和内部 mebs 分布式块存储的集成, Go 版 glance 分页查询优化
- 理解 REST 微服务设计, 熟悉异步框架 Tornado, 熟悉云计算 IaaS 业务, 熟悉云计算分布式架构

亚信智慧数据科技有限公司, 大数据运维分析 2015 年 5 月 – 2015 年 7 月
Hadoop, Hive, Linux, Shell

- 参与中国移动武汉大数据平台的建设工作, 维护 hadoop 集群
- hive 数据仓库数据入库错误排查, 主要通过 shell 脚本和 hiveQL 来统计实际入库数据, 然后和 mysql 元数据库记录对比来排查数据入库错误
- 针对需求写 hiveQL 抽取移动用户数据, 进行上网行为分析

👨‍💻 项目经历

Twitter 数据实时分析
Python, Java, Storm, Redis, Flask

- 通过 Twitter 提供的 API 获取 Twitter 数据, 构造 Storm spout 数据源;
- 使用 Python 中文分词技术, 将数据进行分词, 去除各种杂项词汇, 标点符号, 表情等, 构造出 Storm bolt 计算节点;
- 将 Storm 最后的计算结果发布到 Redis, 使用 Python flask web 从 redis 实时接受数据并展示 Twitter 实时主题词;

血常规机器学习 2016 年 11 月 – 2016 年 12 月
Python, Machine Learning, Flask, Neural Networks, Sklearn, Vue

- 从头搭建了 demo 的前后台, 前端使用 bootstrap 和 vue.js 等框架完成, 体会到了 MVVM 设计模式的简洁优雅; 后台采用 flask web 框架, json 格式做前后台数据传输, 快速学习了简易 python web.
- 采用 sklearn 构建了多种监督模型, 对血常规检测进行性别预测, 通过对数据的预处理, 归一化, 超参数调整, 最后采用 GBDT 特征选取, 提高模型预测正确率;
- 使用 C 语言实现了支持多隐层的神经网络模型, 加深对模型理解. 体会到软件工程的整个团队合作, 以及规范的文档代码的重要性. 项目的最终展示效果可以查阅[这篇博客](#).

⚙ IT 技能

- 熟悉语言: Java, Python, C, 了解 Go
- 熟悉系统: Centos/Ubuntu/MacOS/Windows
- 了解平台: Hadoop, Storm, 理解 Hadoop MapReduce 基本原理, 采用 Go 实现过简易 MapReduce 框架
- 熟悉工具: VSCode, Eclipse, Git, Markdown, Vim
- 机器学习: Neural Networks, SVM, LR, DecisionTree, RandomForest, AdaBoost, GBDT, CNN 等
- 并发技术: 多线程, 异步 IO

★ 获奖情况

甲骨文杯 *java* 设计大赛华中赛区二等奖
国家励志奖学金

2015 年 3 月
2013 年 10 月

♡ 其他

- 独立博客: [MetaHacks](#) 开源社区活动: [GitHub Coding](#)
- 外语: 英语 - 熟练阅读英文技术博客文档等, 可以简单的交流
- 良好的计算机科学素养, 扎实的数据结构算法基础, 喜欢琢磨优雅紧凑简介的代码