



丁语琪

📍 深圳 📞 15013858201 @ dingyq2023@gmail.com 🌐 技术博客: https://dingyuqi.com

|      |                                                                            |                 |
|------|----------------------------------------------------------------------------|-----------------|
| 教育经历 | 东南大学<br>计算机科学与技术<br>3.4 GPA<br>实验室导师: 杨绍富<br>实验室项目: 军科委分布式无人机编队 ORB-SLAM算法 | 2017-2021<br>本科 |
|------|----------------------------------------------------------------------------|-----------------|

|      |                                                                                                                                                                                                                                                                                                                                                                                                          |                   |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|
| 工作经历 | 深圳计算科学研究院<br>算法工程师 <ul style="list-style-type: none"><li><b>[后端开发]:</b> 基于微服务, 分布式的架构, 进行多个项目的后端开发. 开发技术涉及MySQL, PostgreSQL等关系型数据库, 常用的KV数据库Redis以及消息中间件和消息队列. 开发语言为Go.</li><li><b>[算法开发]:</b> 实现我国樊文飞院士的大数据分析系列论文的工程化应用, 参与分布式大数据图挖掘算法系统--钓鱼城系统的开发与维护. 日常工作涉及分布式算法, gRPC, HDFS, 图挖掘算法等技术. 开发语言为Python和Go.</li><li><b>[数据分析]:</b> 负责部分项目的客户数据的现场分析. 曾现场支持过某电信运营商和多个银行项目, 完成前期数据分析处理以及后续的POC支持.</li></ul> | 2021年07月-至今<br>深圳 |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |             |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| 项目经历 | 烟草均质化维稳系统<br>后端开发 <ul style="list-style-type: none"><li><b>项目背景:</b> 为解决烟草公司面临的库存缺失导致产品品质不稳定的痛点, 设计并开发了一个数据管理系统, 帮助客户在部分烟叶缺失的情况下使用有库存的烟叶进行配方替换且产品抽吸感受稳定.</li><li><b>[微服务架构]:</b> 使用Gin作为Web框架, Go语言开发. 整个系统包括: 基础数据管理服务、库存管理服务、生产计划管理服务和模型管理等5个子服务.</li><li><b>[分布式事务一致性]:</b> 业务上烟草生产计划的创建与执行涉及到对库存服务中数据的修改, 系统通过使用中间件Seata的2PC机制来保障分布式事务的一致性.</li><li><b>[Redis缓存]:</b> 在基础数据管理子服务中, 针对读多写少的应用场景, 通过Redis对部分数据进行缓存以提高读取效率. 开发过程中解决了Redis与后端MySQL数据库间的数据一致性问题.</li></ul> MySQL, Redis, 微服务, 数据一致性 | 2023年12月-至今 |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|

|  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |             |
|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
|  | 钓鱼城系统<br>大数据分布式图计算系统 <ul style="list-style-type: none"><li><b>项目背景:</b> 钓鱼城系统是对樊文飞院士关于图计算的系列论文的工程化实现. 目的是实现在大数据背景下基于图数据的关联关系的挖掘.</li><li><b>[微服务架构]:</b> 使用Gin作为Web框架, Go语言开发. 整个系统包括: 数据源管理、规则发现、模型管理等多个子服务.</li><li><b>[分布式算法]:</b> 为了实现在大数据情况下对内存的控制, 系统对图文件进行分割, 分多个子图完成计算. 由于图分割导致的子图边缘节点信息缺失在计算中会通过gPRC访问对应节点获取相应统计信息.</li><li><b>[数据转换]:</b> 由于实际应用中多数客户使用关系型数据库进行业务数据的存储, 所以钓鱼城系统支持在用户在数据管理子服务中配置数据库连接信息. 系统会自动将关系型数据库转换为图数据文件并将文件存储于磁盘用于后续的挖掘算法使用.</li></ul> 分布式算法, 微服务, gRPC | 2021年07月-至今 |
|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|

|  |                                                                                                                                                                                                                                                                                                                                                                                                                                        |                   |
|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|
|  | 银行审计异常检测模型<br>算法开发 <ul style="list-style-type: none"><li><b>项目背景:</b> 针对于当前内部审计系统预警方式单一, 异常值计算数学上不够严谨, 机构风险值计算不够科学等痛点, 为审计部门提供新的无监督异常检测算法, 帮助审计人员在日常工作中快速识别出经营机构各项业务中的异动, 及时掌握和处理经营指标的趋势、节奏以及结构的异常变化.</li><li><b>[数据洞察]:</b> 整理分析2016-2020年共8个业务体系的数据, 涵盖1200+指标, 数据量达到1,600,000+. 充分理解内部系统存储状况以及当前使用的预警系统的算法.</li><li><b>[构建模型]:</b> 300+个指标, 使用孤立森林和AutoEncoding等模型进行无监督异常检测. 并使用关联分析进行数据分析增加业务人员的分析维度.</li></ul> 孤立森林, AutoEncoding | 2023年09月-2023年10月 |
|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|

|  |                                                                                                                                                                                                                                                                                                                                                                                                           |                   |
|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|
|  | 运营商携号转网客户群预测<br>钓鱼城系统实际应用 <ul style="list-style-type: none"><li><b>项目背景:</b> 为解决用户流失严重的问题, 使用钓鱼城系统深度挖掘海量的用户行为数据, 预测未来可能携号转网或流失的客户群体并给出可能的流失原因.</li><li><b>[海量数据]:</b> 处理约50万个用户, 统计数据, 通话详单, 上网日志等累计超10T数据, 约113个维度.</li><li><b>[算法构建]:</b> 在充分理解业务的基础上, 构建业务的拓扑关系. 使用分布式大数据关联分析算法, 在时序图结构上挖掘关联规则, 锁定流失人群的行为模式.</li><li><b>[规则生成]:</b> 项目最终挖掘出多条关联规则, 在提高运营商的预测准确率的同时弥补了传统模型对结果解释性不足的缺陷.</li></ul> 频繁模式挖掘 | 2021年10月-2022年01月 |
|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|

|    |                                                                                    |      |
|----|------------------------------------------------------------------------------------|------|
| 证书 | 系统分析师<br>计算机技术与软件专业技术资格 (级别: 高级)<br>📄 https://www.dingyuqi.com/pdf/certificate.pdf | 2024 |
|----|------------------------------------------------------------------------------------|------|

|     |                                                      |                                                             |                                                                                         |                             |
|-----|------------------------------------------------------|-------------------------------------------------------------|-----------------------------------------------------------------------------------------|-----------------------------|
| 技术栈 | Python<br>●●●●○<br><br>Gin   Flask<br>Web框架<br>●●●●○ | Go<br>●●●●○<br><br>Kubernetes   Docker<br>基本的容器化部署<br>●●○○○ | MySQL   PostgreSQL<br>主流的关系型数据库<br>●●●●○<br><br>ETCD   Seata   Kafka<br>中间件的运用<br>●●○○○ | Redis<br>NoSQL数据库<br>●●●○○○ |
|-----|------------------------------------------------------|-------------------------------------------------------------|-----------------------------------------------------------------------------------------|-----------------------------|