

# 个人简历

Programming is not just a profession,  
but also a pursuit of excellence.

## 基本信息

姓 名：李铭

年 龄：27 岁

毕业院校：河南大学民生学院

联系电话：13183271683

求职状态：已离职

性 别：男

工作年限：5 年

毕业时间：2018-07

电子邮箱：13183271683@163.com

期望职位：大数据开发工程师

## 专业技能

1. 熟练掌握 Java/Scala 编程语言，多线程、集合等基础框架，熟悉常用设计模式
2. 熟悉 JVM 内存结构、垃圾回收机制及其垃圾收集器
3. 熟悉 Linux 常用的操作命令，可以编写 shell 脚本调度服务
4. 熟悉 Hadoop 生态圈组件，熟悉 HDFS 架构和读写流程、YARN 工作机制、YARN 调度器调度策略、作业提交流程和 MR 的 Shuffle 工作流程及其参数优化
5. 熟练使用 Flume 进行日志采集，熟悉 Flume 的三大组件(Source/Channel/Sink)，自定义拦截器对日志数据进行拦截过滤
6. 熟练掌握 ZooKeeper 基础架构及半数选拔机制，能够使用 ZooKeeper 协调其它组件实现其它组件的高可用架构
7. 熟练掌握 Hive 的基础架构，熟练使用 Hive SQL 对数据进行统计分析，熟练使用 Hive 中常见函数，开窗函数，自定义函数，了解 Hive 中的一些优化，可以解决 Hive 中的常见问题
8. 熟练使用 DataX 在多种数据源之间进行数据同步，熟悉 MaxWell 实现增量采集的原理并使用 MaxWell 采集增量数据
9. 熟练掌握 Spark，使用 Spark Core 中的算子提交任务，熟悉 Spark 的任务提交流程、任务的调度切分、通讯架构、Shuffle 过程
10. 熟悉 Hbase 基础架构及其读写流程、RowKey 设计原则，能使用 Phoenix 配合 Hbase 进行数据操作
11. 熟悉 Redis 常用数据结构及其备份策略，使用 Redis 实现热存储功能，了解高并发环境下缓存穿透雪崩解决方案
12. 掌握 Spring Cloud、Spring Boot、Spring MVC、Spring、MyBatis 等主流开源框架
13. 掌握 VUE3/2、HTML、CSS、JavaScript 等 Web 前端技术
14. 熟悉 Kafka 消息中间件，利用 MQ 降低系统间耦合、实现流量的缓冲及削峰填谷的功能

## 工作经历

**北京三十网络科技有限公司** 大数据开发工程师（2020 年 3 月~2023 年 7 月）

工作内容：

1. 配合完成大数据相关项目规划和设计方案，生产环境的部署和实施
2. 根据业务需求对项目业务数据、用户行为数据进行多维度分析，得出用户行为与项目运营的各项指标
3. 将分析处理后的数据落地到 RDBMS 中，接入业务使用，为企业产品部门和领导层决策提供数据支持
4. 使用 Spark 高效分析处理用户行为数据并创建用户流失预测模型，减少公司客户流失
5. 大数据平台的日常维护和优化

## 项目经历

### 项目一：瑞祥用户画像中心

**项目介绍：**瑞祥用户画像中心是一个旨在分析和整合客户、信息员、业务员、设计师多角色的数据，以生成全面用户画像的平台。通过收集、处理和分析来自多个平台产生的业务、日志数据，为企业产品部门和领导层提供更深入的用户洞察，从而支持更精准的营销活动策略和业务决策。

**技术架构：**Spring Boot + Spark + MyBatis + MySQL + ClickHouse + Hive + Redis + VUE + Ant Design Vue

**主要工作：**

1. 项目前期参与需求分析、业务探讨以及功能模块的划分等工作；
2. 设计并实现标签管理、词云图以及用户分群模块，为团队提供高质量的数据分析和可视化工具；
3. 通过对业务需求的理解将三级标签转化为高效的 Hive SQL 查询，提高数据查询性能和资源利用率；
4. 设计自动化的定时任务，定期更新用户分群中的用户信息，确保数据的准确性，为业务模块开发团队提供数据支持，确保分析结果和业务目标紧密对齐。

**成果地址：**瑞祥用户画像中心 <https://userprofile.rxjyzs.com>（需公司内网访问）

## 项目二：瑞祥离线数仓平台

**项目介绍：**离线数仓平台为公司提供了高效可扩展的数据存储和分析解决方案，以支持敏捷的业务决策和深入的洞察发现，项目采取了数据分层存储策略，侧重于对历史数据的监控与分析，通过定制化的指标追踪，帮助业务团队快速识别趋势和问题，这使得公司能够更为迅速的做出反应和调整业务策略同时为报表系统提供了数据支持。

**技术架构：**DataX + MaxWell + Flume + Kafka + HDFS + Hive + DolphinScheduler + Superset

### 主要工作：

1. 参与搭建可扩展的离线数仓平台；
2. 负责设计和优化 Hive 数仓表结构，以支持多维数据分析和高效查询性能；
3. 设计并实现了针对公司业务特点的定制化指标，包括每月新客个数、每月签单金额、项目类型占比、每月登录用户数、各年龄层比率以及用户地域分布排名等；
4. 将经过分析处理的数据以定期任务的形式落地到 RDBMS 中，保障了数据的可靠性和持久性。通过与业务团队的紧密协作，保证分析结果能够无缝接入公司的产品和业务流程，为产品优化和决策指定提供数据支持。

**成果地址：**瑞祥指标报表平台 <https://bi.itchzz.com>（需公司内网访问）

**数据来源：**瑞祥官网 <https://www.rxjyzs.com>

客户端 PC 端 <https://user.rxjyzs.com>

客户端小程序 微信搜索小程序“室内设计装饰装修专属服务客户端”

业务员 PC 端 <https://swpt.itchzz.com>（需公司内网访问）

业务员小程序 微信搜索小程序“瑞祥顾问”

设计师 PC 端 <https://zapt.itchzz.com>（需公司内网访问）

设计师小程序 微信搜索小程序“设计师室内设计服务 IDS”

信息员小程序 微信搜索小程序“牛经纪”

## 教育经历

2014.9~2018.7 期间就读于河南大学民生学院的通信工程专业

## 自我评价

1. 具备敏锐的技术嗅觉，对新兴技术保持持续学习和深入研究的热情；
2. 出色的沟通和协调能力，具备团队合作意识，乐于接受各种挑战；
3. 具有较强的逻辑思维能力，拥有丰富的编程经验，追求不断学习和探索；
4. 始终以尽职尽责的态度对待工作，乐于应对具有挑战性的任务；
5. 积极与用户和同事沟通，快速解决项目中遇到的问题；
6. 通过 CET-6 考试，具备流利阅读英文资料和技术文档的能力。

