

## 教育背景

毕业学校：天津科技大学

学 历：本科

专 业：物流管理

## 求职意向

意向职位：大数据开发工程师

求职地址：上海

期待薪资：面议

## 工作经历

2018.07 — 2022.02 北京春雨天下软件有限公司 大数据开发工程师

2022.04 — 至今 上海天跃科技股份有限公司 大数据开发工程师

工作职责：

- 1、 负责公司 CDH 大数据平台的搭建、配置及调优工作
- 2、 数据 ETL 处理，参与数据仓库设计
- 3、 根据业务需求指标，高效率地数据查询、数据分析
- 4、 分析计算离线、实时指标，为公司产品部门和高层决策提供数据服务支撑
- 5、 数据分析平台的日常维护与优化

## 专业技能

1. 熟练 Flink 的运行架构、Flink 流处理核心编程；熟练 Flink Table API & SQL 的应用及优化
2. 熟练搭建 CDH 集群，及装载相关大数据组件
3. 熟练 Hive 架构原理、HQL、Hive 项目的优化及数据倾斜的解决方案
4. 熟练掌握 Hbase，理解基本原理，及 Rowkey 设计原则，针对不同场景设计表。
5. 熟练掌握 Hadoop 核心组件 HDFS、MapReduce、Yarn 及其框架原理、源码及 API 操作；熟悉 HDFS 故障排除、Hadoop 企业优化、HA 高可用
6. 熟练掌握 Hbase 的基本架构、工作原理、Hbase API 及优化设计和 Phoenix 操作
7. 熟悉 Kafka、Flume、Zookeeper 集群部署、架构原理及参数调优；熟悉运用 Kafka Eagle 对 Kafka 集群监控；熟悉运用 Ganglia 对 Flume 数据流监控

- 
8. 熟练使用 datax 工具，进行数据在数据库之间传输。
  9. 熟练掌握 MySQL 数据库，熟练相关操作，针对不同场景建表。
  10. 熟悉维度建模理论，根据需求进行建表。
  11. 熟悉 Spark 核心原理 SparkCore、SparkSQL、SparkStreaming 及应用开发;理解并熟练 spark 各种性能调优、背压及数据倾斜的解决方案
  12. 了解 Zabbix 集群监控及 Granaga 的基础使用;熟悉 Presto、Superset 的部署及可视化报表；熟悉 Azkaban 的安装部署及应用
  13. 熟悉 ClickHouse 及 OLAP 在线分析处理
  14. 熟悉 Redis 数据库存储系统及应用
  15. 熟练 doris 的部署，以使用
  16. 熟悉 mongo 简单使用

## 项目经验

### 项目一、天跃科技履职在线分析系统

1. 项目名称：天跃科技履职在线分析系统
  2. 应用技术：Flink、FlinkCDC、CDH 集群、dataX、Hadoop、Zookeeper、Hbase、MySQL、MongoDB、Kafka、Phoenix
  3. 项目描述：搭建 Flink 在线分析平台以解决一些对实时性要求很高的流处理场景，实现在线分析指标计算
  4. 技术要点：搭建 CDH 分布式集群及装载大数据相关组件，使用 dataX 将历史数据迁移到数据分析库，使用 FlinkCDC 实时监听业务库 MySQL、MongoDB 的数据变化，应用 Flink 状态编程数据去重统计有效数据、应用 Flink 侧输出流实现用户行为的动态分流、应用 Flink 两阶段提交统计需求指标，配置表广播流保存维度到 Hbase、加入 Redis 旁路缓存解决外部数据源查询、采用异步 IO 解决网络延迟。
- 个人职责：基于 Flink 架构，负责搭建在线分析平台，表结构设计，负责在线指标需求分析，负责平台的调优及维护

### 项目二、天跃科技警情预警智擎在线系统

5. 项目名称：天跃科技警情预警智擎在线系统
6. 应用技术：Flink、FlinkCDC、Hadoop、Zookeeper、Hbase、MySQL、Kafka、Redis、ClickHouse

---

7. 项目描述：搭建 Flink 在线分析平台以解决一些对实时性要求很高的流处理场景，实现在线分析指标计算

8. 技术要点：使用 FlinkCDC 实时监控报警设备数据，导入数据分析库，配置表广播流保存维度到 Hbase，应用 Flink 侧输出流实现警情相关数据的动态分流、应用 Flink 状态编程实现有效警情校验及数据去重，加入 Redis 旁路缓存解决外部数据源查询、采用异步 IO 解决网络延迟。

个人职责：基于 Flink 架构，负责搭建警情报警在线分析平台，表结构设计，负责在线指标需求分析，负责平台的调优及维护

## 项目三、

1. 项目名称：春雨医生在线指标分析系统

2. 应用技术：Flink · FlinkCDC · Hadoop · Zookeeper · Hbase · MySql · Kafka · SpringBoot · ClinckHouse ·

3. 项目描述：搭建 Flink 在线分析平台以解决一些对实时性要求很高的流处理场景，实现在线分析指标计算

4. 技术要点：应用 Flink 状态编程实现新老用户校验及数据去重、应用 Flink 侧输出流实现用户行为的动态分流、配置表广播流保存维度到 Hbase、基于模式序列判定用户行为跳出明细、加入 Redis 旁路缓存解决外部数据源查询、采用异步 IO 解决网络延迟。

5. 个人职责：基于 Flink 架构，负责搭建在线分析平台，负责在线指标需求分析，负责平台的调优及维护

## 项目四、

1. 项目名称：春雨医生在线指标分析系统

2. 应用技术：SparkStreaming · Scala · Hbase · Redis · ElasticSearch · Kibana · Kafka · FlinkCDC

3. 项目描述：搭建了基于 SparkStreaming 的在线分析平台，实现在线分析指标计算，采用 Redis 进行数据去重实现 DAU 指标分析、实现同一设备五分钟内三次不同账号登录，领取优惠券但没有浏览商品的预警日志分析、订单明细与订单双流 join 获取销售明细

4. 个人职责：配合搭建 SparkStreaming 的在线分析平台，负责在线指标需求分析，协助平台的调优及维护

## 项目五、

1. 项目名称：春雨医生离线数据指标分析系统

2. 应用技术：Spark · Scala · Hbase · Hive · MySQL · Sqoop · Kafka · Flume · Zookeeper · Hadoop · Nginx · ElasticSearch · Zabbix · Presto · SuperSet

---

3. 项目描述：该项目主要为公司决策制定提供系统数据支持,改进业务流程、降低成本。项目用户行为数据由前端埋点生成,经由 Nginx 及日志服务器流入 Flume,经 Flume 传输至 Kafka,Flume 消费 Kafka 并采用 HDFS Sink 存储到 HDFS 上;业务数据由 MySQL 数据库经 Sqoop 导入至 HDFS;应用 Hive 进行数仓分层及离线指标分析计算;应用 Azkaban 全流程调度、Zabbix

4. 个人职责：负责基于 Hadoop、Hive 架构离线分析平台的搭建，负责数据的 ETL、参与数据仓库建模，负责离线指标的分析计算、项目调优及集群维护

#### 个人评价

- 1、为人稳重负责，具有良好的团队协作能力
- 2、能够严格按照公司的制度规范自己，对工作认真负责
- 3、很好的学习能力，善于领悟、接受新事物和新技术