

个人简历

姓名：李锦盛

性别：男

年龄：26岁

经验：5年

手机：15936794502

邮箱：15936794502@163.com

学历：本科

意向岗位：数据开发工程师

专业技能

- 编程语言**：熟练掌握 Java、Python、Scala 等编程语言，具备扎实编程基础；能熟练使用 Linux 操作系统，编写 Shell 脚本实现日常任务自动化。
- 计算框架**：精通 Flink 框架，熟悉 Flink SQL、CEP，掌握状态后端调优、检查点、水位线、窗口等原理；熟练运用 Spark Core、Spark SQL、Spark Streaming 进行大数据处理。
- 生态系统**：深入理解 Hadoop 生态系统，熟悉 MapReduce 工作原理与 YARN 资源调度，具备集群管理与文件存储经验。
- 数据仓库与 SQL**：精通 Hive 数据仓库设计与构建规范，能高效存储与访问数据；熟练使用 SQL，掌握 HiveSQL 开窗函数、UDF，擅长复杂查询性能调优。
- 数据库管理**：精通 MySQL、Oracle 等关系型数据库设计与维护，掌握分区表、索引、分片技术；熟悉 MongoDB、Redis、HBase 等非关系型数据库。
- 即席查询引擎**：熟练使用 ClickHouse、Doris 等，能创建物化视图，避免不必要计算与存储，保障数据准确性与完整性。
- ETL 与数据集成**：熟悉 Kettle，可完成数据集成、ETL 处理、数据转换清洗与迁移；了解阿里云数仓搭建，熟练使用 DataWorks、MaxCompute 组件。
- 数据湖**：熟悉 Paimon 和 Dinky，具备现代数据湖架构构建与管理能力。
- 监控工具**：熟悉 Prometheus + Grafana、Ganglia、EFAK 等监控系统，可监控 Flume、Kafka 等组件运行状态。
- 可视化工具**：熟练使用 Superset、FineReport、Sugar BI 等，能进行数据分析与可视化，绘制 BI 看板。
- 任务调度**：熟练掌握 Azkaban、DolphinScheduler 等调度器，实现数据处理工作流自动化管理。
- 机器学习**：熟悉用 Python 进行机器学习参数调优、性能评估，提升模型准确率、召回率、AUC 值等指标。

工作经历

2020/10 - 2025/07

河南奇点网络科技有限公司

大数据开发工程师

项目经历

2024/08 - 至今

智慧物流与供应链协同创新平台

项目概述：

该项目整合全省公路、铁路、航空、水运等多式联运数据（年增量达 8PB）、30 万+ 货运车辆实时轨迹、5000+ 仓储节点库存数据、20 万+ 企业供需信息及跨境电商物流数据，构建“数据驱动+智能协同”的现代物流体系。通过全链路数据治理与智能算法模型，实现货运供需精准匹配、多式联运路径优化、供应链风险预警及应急物资智能调度，助力河南打造全国智慧物流创新高地。

技术栈：Flume+Sqoop+Flink CDC+Kafka+Hadoop HDFS+Hive+HBase+MongoDB+ClickHouse+MinIO+Spark+Flink+Presto+Neo4j+PyTorch+Docker+Azkaban+Prometheus+Grafana+ZooKeeper等技术

项目成果：公铁联运中转效率提升 52%，集装箱铁水联运量同比增长 47%，郑州国际陆港中欧班列开行效率提升 35%，货柜周转率提高 28%；在防汛抗旱、疫情防控等场景中，应急物资调度响应时间从 48 小时缩短至 6 小时，精准投放准确率达 92%，保障 13 次重大应急事件物资供应，覆盖 90% 的货运枢纽节点，取得突破性成果

2023/07 - 2024/06

营运车辆实时监控与智能调度系统

项目概述：

随着公司专用车联网业务的快速拓展，旗下运营的工程车辆、物流车辆及网约车规模已突破 1.2 万辆，车辆实时位置、行驶状态、油耗数据及订单信息等产生的数据流日均增量达 800GB。公司决定构建基于实时 PySpark 的 "营运车辆实时监控与智能调度系统"，通过对车辆实时数据的秒级处理与分析，实现异常行为实时预警、运力供需动态匹配及路径智能优化，提升车辆运营安全性与调度效率。

技术栈 :PySpark+PySpark Streaming+Structured Streaming+ Kafka+ Redis+HBase+MySQL+MQTT+Flume+Python F lask+Vue.js+ECharts+ Fine BI

项目成果：

基于实时油耗分析与路径优化，单车辆日均油耗降低 8.5%，每月节省燃油成本约 42 万元；急加速、急刹车等危险驾驶行为减少 41%，车辆事故率下降 37%，司机安全培训针对性提升 系统沉淀的实时流处理框架已复用于公司网约车业务线，支撑了日均 20 万 + 订单的实时派单优化，用户叫车响应速度提升 18%；车辆异常状态（超速、设备故障）识别延迟从原来的 15 分钟缩短至 3 秒，预警信息推送准确率达 96.8。

2022/03 - 2023/05

电商平台大数据分析 with 智能运营

项目描述：

综合电商平台在河南省业务扩张，平台积累了海量电商数据，包括日均 200 万 + 订单数据、500 万 + 用户行为日志（浏览、收藏、加购）、30 万 + 商品信息及库存变动数据等。为解决传统运营模式中用户画像模糊、商品推荐精准度低、库存周转慢等问题，公司决定构建基于大数据技术栈的电商数据智能分析平台。通过对全量业务数据的采集、存储、计算与分析，实现用户精细化运营、商品智能推荐、库存动态优化及营销活动精准投放，提升平台整体运营效率与用户体验。

技术栈 :Flume+Sqoop+Hadoop+ NoSQL+MongoDB+Hive++Spark+Presto+Azkaban+FineBI+MySQL+ZooKeeper

项目成果：

日均处理数据量达 15TB，较传统方式提升 8 倍；Azkaban 调度的 120 + 个数据任务成功率达 99.6%，数据链路延迟控制在 2 小时内；基于 Spark MLlib 构建的用户画像系统，实现 1000 + 标签的精准打标，用户分群准确率达 92%；精细化运营活动转化率提升 40%，用户留存率（30 天）提高 18%；FineBI 构建的 30 + 个运营仪表盘，使业务指标获取时间从 2 天缩短至 5 分钟，支撑管理层制定 3 项关键运营策略调整，带来季度 GMV 增长 12%。

2020/10 - 2022/01

离线数据仓库构建

项目概述：

为支撑公司业务拓展,涵盖成品油流通服务、建设劳务对接、专用车辆、运营调度等多元业务场景,伴随业务版图的持续扩大与用户规模的稳步增长。这些数据涵盖成品油审批数据、建设劳务用工数据、专用车辆运营数据、用户行为数据等多维度信息。为充分发挥数据价值,实现精准业务决策、深度用户行为分析、运营策略优化及数据驱动的业务创新,公司决定自主构建一套高效、可靠、可扩展的离线数据仓库。

技术栈 :MaxCompute + DataWorks + RDS MySQL + DataX + QuickBI

项目成果：

基于该离线数据仓库构建的多业务数据智能分析平台在公司成功落地后,数据处理与分析效能显著提升。数据整合准确率提升至 97%,实现了跨业务线数据的高效关联与统一分析;业务运营决策响应速度加快 50%;自平台上线半年以来,企业端用户的业务办理效率提升 42%,政府合作单位的监管数据获取时效缩短 38%,有效推动了公司业务的数字化升级,为后续数据驱动的业务创新奠定了坚实基础。

拥有证书

影刀高级证书

FCA-FineBI

FCA-FineReport

FCA-FineDataLink

教育背景

2016/09 - 2020/07

黄河科技学院

本科 - 数据科学与大数据技术

自我评价

- 1.逻辑思维能力强,能快速解决复杂问题。
- 2.有着良好的团队合作精神,对新技术充满热情,持续学习以保持技术的前沿性
- 3.有复杂业务系统的抽象设计能力,能独立对接业务需求和解决业务与技术问题
- 4.及时反馈开发遇到的难点问题,沟通并寻求解决方案,按阶段有计划的完成开发任务