

性别：男 年龄：28

电话：13599527847 户籍：福建

现所在地：深圳 开始工作时间：2018

邮箱：13599527847@163.com



求职意向

- 意向岗位：大数据开发工程师
- 期望行业：大数据
- 意向城市：深圳
- 当前状态：马上入职



荣誉证书

FCP-报表交付工程师证书、CET 4、全国计算机等级一级、FCA-FineReport 证书



自我评价

- 六年的大数据开发工作经验，具备优良的问题解决能力和创新思维，能够快速定位和解决大数据平台运行过程中出现的问题。善于从业务需求出发，运用大数据技术为业务部门提供有效的解决方案。
- 注重团队协作，能够与不同背景的人员有效沟通和合作。积极分享自己的技术经验，帮助团队成员共同成长，提高团队的整体技术水平。具有较强的抗压能力，能够在高压力的环境下高效工作，为企业的发展贡献自己的力量。



教育经历

2014.9-2018.7

福建医科大学

本科（一本）



相关技能

- SQL 语言基础扎实，具备 Oracle、 MySql、 等关系型数据库查询、优化和复杂指标开发的实战经验；
- PL/SQL 程序开发有一定经验，能够编写存储过程、函数；
- 掌握 Hive，能够编写 HiveSQL 进行大数据处理和分析以及复杂的指标开发；
- 利用 FineReport 和 FineBI 进行数据可视化，制作动态报告；
- 使用 Kettle、 DataX、 Sqoop 等 ETL 工具；
- 熟悉掌握 Hadoop 生态系统，包括 HDFS、 MapReduce、 YARN 等组件的原理和应用；
- 熟悉 Linux 操作系统，包括常用指令、环境部署、软件配置、具备 Shell 脚本编程能力；
- 熟悉数据仓库生命周期及各类建模理论，具备数据分析、挖掘、建模能力；
- 熟练使用 Navicat、 PL/SQLDeveloper、 DataGrip、 Azkaban 、 Dophin-schedule 等开发和数据库管理工具。
- 熟悉金融和公司理财知识，包括但不限于：利息、利率、货币政策、金融监管、国际资本流动、财务报表分析、折现现金流量估价、债券估值、股票估值、风险分析、套利定价模型（APT）。
- 了解医疗数据的特点和标准，如 HL7 (Health Level 7) 、 DICOM (Digital Imaging and Communications in Medicine) 等标准，确保采集到的数据的准确性和完整性。



工作经历

2021.10-2024.9

深圳市思无涯科创科技有限公司

数据开发工程师 | 大数据

2018.7-2021.9

深圳市思为软件技术有限公司

数据开发工程师 | 大数据



项目经历

项目一

2023.11-2024.09

深圳市“金租大数据智能风控”金融租赁项目

项目概述：本项目是“金租大数据智能风控”金融租赁项目，旨在整合金融租赁企业内外部数据资源，建立统一的数据存储和管理平台，实现数据的集中化、标准化和规范化，实现数据的存储、管理和维护，发现潜在规律和价值。

项目架构：Mysql、dataX、Maxwell、Zookeeper、Kafka、Flume、Linux、Hadoop、Hive、Dolphin-schedule

项目职责：

- 1、分析需求，参与需求评审，并且参与数据库的设计和讨论，与业务沟通，梳理需求逻辑，根据需求工单排期。
- 2、用 dataX、Maxwell、Zookeeper、Kafka、Flume 将 Mysql 中的业务数据导入到 HDFS 中实现数据的采集、清洗和存储。
- 3、数据仓库的分层和搭建，使用 hive 工具将数据分为 4 层，贴源层（ODS）存放未经过处理的原始数据；公共维度层（DIM）存放维度模型中的维度表；明细数据层（DWD）存放维度模型中的事实表，保存各业务过程的最小粒度的操作记录；数据应用层（ADS）构建分析的主题，存放各项统计指标结果。

项目二

2022.06-2023.10

宝健商城数据运营分析

项目描述：

该项目为宝健商城数据运营分析项目，宝健商城是实名注册海外淘购的一个平台。本项目通过对该平台用户浏览下单数据的获取，统计注册用户数，商品销量排名，商品加入购物车排名，商品差评率，下单数目统计（每日下单数，下单金额以及下单用户数），支付信息统计（每日支付金额、支付人数，支付商品数，支付笔数），用户留存率等来分析平台运营情况，根据用户人群、金额和来源情况的分析结果适度调整营销策略，促进平台长远的发展。

项目架构：Linux、Hadoop、hive、Oracle、shell、HDFS、mapreduce、kettle、FineReport/FineBi

项目职责：

1. 数据仓库的模型设计，积极参与需求分析和需求评审，参与需求和业务指标上的对接。
2. HIVE表的结构设计，如维度表、事实表。
3. 用 ETL 工具 kettle、sqoop 抽取相关数据到数据仓库中的 ODS 层，在 DWD 层建立事实表、DWS 层汇总每天用户行为、DWT 层汇总用户主题，最后在 ADS 层获取指标。
4. 配合数据分析。配合数据分析师根据需求和指标进行分析计算。
5. 配合 BI 工程师出具数据分析报表，为项目评估、考核提供依据。

项目三

2021.10-2022.05

深圳市医院患者出院审核系统数据仓库搭建及应用

项目描述：

深圳市医院为了提升出院患者医疗服务的质量和效率，并专注于开发出院患者医疗服务指标，旨在通过收集、整合和分析出院患者医疗服务的相关数据，为医院管理层提供决策支持，同时优化患者服务体检。

项目架构：Linux、Hadoop、hive、Oracle、shell、HDFS、mapreduce、kettle、FineReport/FineBi

职责描述：

1. 数据仓库模型。根据深圳市第二人民医院提出的需求和业务场景，参与数据仓库模型设计，参与需求和业务指标对接，数据来源于医院现有的 HIS（医院信息系统）。
2. 通过 ETL 工具，将分散在不同系统中的数据进行抽取、清洗、转换和加载，确保数据的一致性和准确性。
3. 构建数据仓库模型。使用 hive 数仓工具将数据分为 3 层，原始数据映射为原始数据表作为第一层的数据层，基于星

型模型构建维度模型作为第二层的控制层，负责存储经过清洗、转换和整合后的临床路径数据，为数据分析提供支持。根据需求和分析指标完成第三层的应用层。

4、配合报表开发。根据深圳市第二人民医院的要求，配合 BI 工程师对数据仓库中的数据进行深入分析，挖掘有价值的信息和规律，为医院管理层提供决策支持。

项目四

2019.12-2021.09

深圳宝安融兴银行理财管理数据分析

项目描述：该项目是针对公司旗下业务进行分析决策管理，对用户购买理财产品的情况进行数据分析，了解用户的购买习惯，推送类似产品以提高交易量。网银在售产品有包括货币基金、黄金、资管、保险类理财产品，通过产品类型、风险等级、客户购买金额、购买期限等多维度来分析，统计用户的购买类型、交易总量、购买期限、收益率等进行同比环比分析。通过报表进行可视化。

项目架构：Oracle、DataX、Linux、Shell、Hadoop、HDFS、mapreduce、kettle、Dolphin-schedule、FineReport
项目职责：

需求分析：分析需求，参与需求评审，并且参与数据库的设计和讨论。

需求开发：使用 sqoop 进行各业务系统数据库之间的全量抽取和增量抽取，对 hive 的 hql 脚本进行日常的优化用 hive sql 语句进行数据的清洗和转换。

稽核测试：根据调整前后进行交叉验证以及对数据量和结果集进行校验。

优化阶段：初步对比运行时效变化，对耗时较长的作业进行 Yarn 的分析判断，对数据倾斜进行加盐打散或者参数调整等方法优化。

代码管理：将开发的代码按照规范归档项目。

项目五

2018.07-2019.11

远光电子大数据搭建项目

一、项目描述：

企业随着业务的不断发展，积累了大量的数据资源。为了更好地挖掘数据价值，提高企业的决策效率和竞争力，远光电子决定启动大数据搭建项目。

二、项目架构：Oracle、MySQL、Hive、kettle、HDFS、yarn、Azkaban、shell、PLSQL、Navicat、DataGrip

三、项目职责：

负责从 Oracle、MySQL 等传统数据库迁移数据到 Hadoop 生态系统，并使用 Kettle 等工具进行数据的抽取、清洗、转移和加载（ETL）。

根据业务需求和技术特点，设计和实现高效、可扩展的数据仓库的高效运行，并使用相关工具进行监控和故障排查。