

# 黄学勤

性别：男 年龄：25 电话 17608900886  
户籍：海南 现所在地：深圳 邮箱: 17608900886@163.com  
民族：汉族 出生日期：1999.06 工作经验:三年大数据开发

## 求职意向

意向岗位：数据开发工程师 当前状态：随时到岗

## 教育经历

2017-2021 青岛工学院  
软件工程 | 统招全日制本科  
1、软件工程专业大数据方向，系统性学习过大数据相关的理论知识体系以及软件工程理论知识体系，分布式计算框架的发展历程中的重要框架和技术都有过基于底层原理级别的认知学习  
2、具备技术视野前瞻性，关注大数据领域新兴开源项目，关注人工智能的发展近况  
3、业务解读和需求解构能力较强，基于对业务解耦和重构的项目优化思路负责过多个项目的性能优化，优化效果较之 sql 层面优化明显，曾受客户褒奖

## 个人技能

大数据技术：  
1、熟悉 Hadoop 框架的应用，熟悉 Hdfs 及 MapReduce 的底层原理，熟悉 Hive 的底层运行机制，具备使用 hive 和使用 spark、实现 MapReduce 的项目经验  
2、具备熟练运用基于 Sqoop、DataX、kettle 等 ETL 开发工具，以及 PowerDesigner 数据模型设计的开发经验。  
3、具备熟练掌握 Azkaban、ozzie、小海豚（Dolphin）等调度工具的使用经验。  
4、熟悉建模理论与方法、维度建模，能针对不同行业的需求正确划分主题，构建概念模型、业务模型、逻辑模型等。  
5、对常见企业级数据可视化方案都有了解，熟练使用基于 FineReport、FineBI 等数据可视化经验。  
6、具备优秀的逻辑思维能力，能对实际问题进行抽象、设计和编程，拥有良好的沟通协作及快速学习能力。

7、熟悉掌握 Oracle、MySQL 等数据库的底层原理与分别，参与项目中数据库数据载入、故障排查、性能优化、备份与恢复等项目经验。

## 工作经历

2021.6-2024.8

北京擎科生物技术股份有限公司广州分公司

大数据开发工程师 | 数据开发

## 项目经历

项目时间: 2023.02--2024.08

项目岗位: 大数据开发工程师

项目名称: 科技邦商城商品营收数据可视化大屏

项目架构: Hadoop+Hive+Oracle+Sqoop+Shell+FineBI

项目概述:

科技邦商城商品营收数据可视化大屏项目旨在为商城管理层提供直观、实时的营收数据洞察，以支持决策制定和业务优化。通过整合商城的各类营收数据，包括销售数据、订单数据、用户数据等，利用 Hadoop 进行分布式存储和处理，Hive 进行数据仓库构建，Oracle 进行数据存储和管理，Sqoop 实现数据的抽取和导入，Shell 进行脚本自动化处理，最终借助 FineBI 实现数据的可视化展示。

工作内容:

- 1、与运营团队充分交流意见后，整合需求文档，并进行数据仓库的设计；
- 2、业务数据通过 Sqoop 进行数据迁移，存储在 HDFS 上，开发 Shell 脚本并通过 Azkaban 定时调度，采用数仓工具 Hive 进行数据计算；
- 3、将 DWD 层数据在尽量贴源的前提下，做数据清洗、整合、标准化，并解决 Hive 优化数据倾斜等技术难题；
- 4、结合业务需求，确认 DWT 层部分主题的建立，将 DWD 与 DWS 层的数据按照业务需求进行加工创建对应的主题表；
- 5、根据需求文档在 ADS 层进行不同指标开发；
- 6、协助 BI 人员使用 FineBI 进行数据的可视化呈现，为公司决策层提供可靠的数据支持。

项目时间: 2022.09--2023.02

项目岗位: 大数据开发工程师

项目名称: 科技邦商城商品供需平台专项性能优化

项目架构: Hadoop+Hive+Oracle+PLSQL+Sqoop+Kettle+Shell+FineReport

## 项目概述：

科技邦商城商品供需平台随着业务的扩展，面临着数据量激增、系统响应变慢等问题。本项目旨在对该平台进行专项性能优化，以提高系统的稳定性、可靠性和性能，满足用户对高效购物体验的需求。通过采用 Hadoop 进行分布式存储和计算，Hive 进行数据仓库管理，Oracle 存储关键业务数据，PLSQL 进行复杂业务逻辑处理，Sqoop 和 Kettle 实现数据的抽取、转换和加载，Shell 进行脚本自动化处理，FineReport 进行数据可视化展示，全面提升平台的性能

## 责任描述：

- 1、性能分析与评估：对平台进行全面的性能分析，包括数据存储、查询响应时间、资源利用率等方面。确定性能瓶颈和优化重点，制定详细的优化计划。
- 2、数据库优化：优化 Oracle 数据库的结构和索引，提高数据查询和写入性能。调整数据库参数，以适应高并发和大数据量的处理需求。
- 3、数据处理优化：优化 Sqoop 和 Kettle 的数据抽取、转换和加载流程，提高数据处理效率。利用 Hadoop 和 Hive 进行分布式数据处理，提升数据处理的速度和规模。
- 4、系统架构优化：评估和优化项目架构，确保系统的可扩展性和稳定性。引入缓存机制，减少数据的重复读取和计算。
- 5、代码优化：审查和优化 PLSQL 代码，提高代码的执行效率。消除代码中的性能瓶颈和潜在问题。
- 6、性能测试与监控：进行性能测试，验证优化效果，确保系统满足性能指标要求。建立性能监控体系，实时监测系统性能，及时发现和解决问题。
- 7、协作与沟通：与开发团队、运维团队密切协作，共同推进性能优化工作。及时沟通优化进展和问题，确保项目顺利进行。

项目时间：2022.03--2022.09

项目岗位：大数据开发工程师

项目名称：商城日志数据离线数仓系统

项目架构：Hadoop+ Hive+ Sqoop+MySQL+FineBI

## 项目概述：

本次项目主要是搭建 商城日志数据分析离线系统，该系统可以商城日志数据进行多维度的分析，并通过网页显示成图表来帮助公司高层来进行决策，通过对用户的日志数据等数据进行一系列的建模和清洗转换，由用户留存率，沉默用户数，新增设备数等具体指标推演出客户使用的情况，挽回流失客户，以提升商城活动促销效果及产品销量。

工作内容：

- 1、负责用户留存率、活跃设备数、沉默用户数等 10 个日志数据指标，设计 Etl 流程；
- 2、使用 Sqoop 从 MySQL 中抽取数据到 Hive，并编写 Shell 脚本进行数据校验，有效进行数据过滤；
- 3、使用星型的维度模型进行数仓建模，负责 Ods、Dwd 和 Ads 的表构建、数据抽取、转换与加载；
- 4、编写 Shell 脚本并设置资源参数，使用 Azkaban 定时调度 30 个 Job，调度周期日调度；
- 5、使用拉链表解决历史数据变更问题，使用局部聚合加全局聚合的方法解决数据倾斜问题；
- 6、通过分桶采样提高开发和自测效率，通过分区、分桶、Orc 列式存储、索引提高查询性能；

项目时间：2021.07--2022.03

项目岗位：大数据开发工程师

项目名称：陇云业务平台项目升级和数据迁移

项目架构：Hadoop+ Hive+ Sqoop+MySQL+FineBI

项目概述：

业务开展顺利，数据增长，需要对整个系统进行核心数据部分的升级和改造，要将数据从 Oracle 数据库迁移到 Hadoop 平台以减轻企业数据负担。将客户,财务及其他相关系统的数据从 Oracle 迁移到 Hadoop 平台并且进行表格数据和需求指标的迁移。

工作内容：

- 1、使用 sqoop 和 kettle 完成数据的导出导入工作；
- 2、使用 hivesql 语言创建表格；
- 3、进行表格的分区和分桶的设计；
- 4、进行 hivesql 语句的自测和日常优化。

### 自我评价

- 1、对大数据开发充满热情且专注，热衷于学习新技术，学习能力强，对技术热点保持敏感，具备良好的团队合作精神，能够与团队成员有效沟通和协作，共同完成项目目标。
- 2、注重工作与生活的均衡，工作时可全身心地投入其中，专注且高效地完成任务；生活中喜欢独自发掘城市的细节，也享受与朋友一起打羽毛球、游泳、品尝美食的时光。