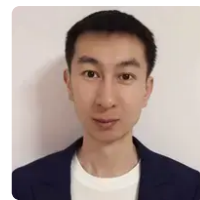


曲鹏飞

男 | 年龄: 35岁 | 18842436740 | qpf1009@163.com

13年工作经验 | 期望城市: 沈阳



个人优势

- 深入理解MapReduce与HDFS的核心原理, 具备大数据离线ETL架构设计与实施经验, 能够高效完成数据清洗、转换与加载任务, 提升数据处理效率。
- 熟练掌握CDH生态组件(包括Hive、HBase、Sqoop、Oozie、Hue、Flume、Kafka、Spark、Spark Streaming等), 具备从数据采集、处理到调度的全流程开发能力, 成功支撑多个大数据平台项目落地。
- 熟练使用MySQL等关系型数据库, 并具备Docker容器技术的实战经验, 能够有效提升系统部署效率与环境一致性。
- 熟练运用Python、Java、SQL等编程语言进行数据开发与业务逻辑实现, 具备良好的代码规范与可维护性意识。
- 熟悉Django后端框架, 具备一定的全栈开发能力, 可灵活应对多类型项目需求。
- 具备较强的问题分析与解决能力, 善于沟通协作, 具备团队管理与协调经验, 能够推动项目高效推进。
- 对新技术保持高度敏感, 具备良好的学习能力与总结习惯, 能够在实践中快速掌握并应用新技能, 确保项目高质量交付。
- 掌握AI工具编程, 提高开发效率。

工作经历

沈阳美行科技股份有限公司 python开发工程师 2023.09-2026.02

1、高精度地图数据开发与处理

- 基于Python及Shapely库, 设计并实现地图数据处理流程, 完成对原始数据的清洗、聚类、曲线拟合及语义信息更新, 提升数据质量和结构化程度。
- 开发自动化工具链, 将多源异构的嘈杂地图数据转化为标准化、高精度的地图数据产品, 支持下游应用及可视化需求。
- 持续优化几何计算与空间分析算法, 提升数据处理效率与准确性。

2、智慧停车平台系统开发

- 使用Django框架设计并开发停车业务后端系统, 涵盖车位管理、计费规则、数据统计等核心模块。
- 参与系统架构设计与数据库优化, 保障平台在高并发场景下的稳定性与可扩展性。
- 协同前端团队完成接口对接与业务逻辑实现, 推动平台功能迭代与用户体验提升。

北京四维图新科技股份有限公司-沈阳软件中心 大数据开发 2020.08-2023.09

- 负责长城大数据项目中 spark streaming、Hive、Azkaban 等代码开发与维护。
- 负责基线用户画像标签系统中 sparksql和 flinksql 的设计与开发, 及开发任务的协调分配。
- 负责丰田呼叫中心的后端项目的 java 开发工作。
- 调研、搭建基于 DSS 作为数据中台的相关工作。

紫勋智能科技(北京)有限公司 高级数据开发 2018.07-2020.07

内容:

- 负责搭建公司本地Hadoop架构集群, 设计大数据ETL 流程架构;
- 负责开发Flume、MapReduce、Hive、Kafka、Oozie代码;
- 使用Python开发应用程序和爬虫程序;
- 使用Python开发包含语音和视觉的人工智能应用。处理语音及图像及视频数据, 进行离线分析等;

5, 使用Python同Zato、Flink、机器学习的研究与开发工作。

业绩:

担任《AI + 智能交互》项目的技术负责人,可以与其他部门人员有效沟通,协力完成项目。

北京银河摩尔信息技术有限公司 大数据ETL工程师 2017.07-2018.07

- 1, 设计、搭建 GPDB 数据仓库。
- 2, 参与搭建 CDH 集群, 优化集群。
- 3, 利用 oozie 实现任务调度, 执行、开发 sqoop、hive、mapreduce、Python、shell等 action 功能, 实现数据的 ETL 流程控制, 完成定时计划、重跑机制等功能。
- 4, 修改 Python 的机器学习模型, 使其与 hdfs和 hive 相结合。

沈阳文远发有限公司 Hadoop开发工程师 2014.11-2017.05

- 1, 参与部署、开发 hadoop 集群, 对相关组件进行测试。
- 2, 负责通过 Mapreduce和 Hive 清洗、统计客户的投资收益情况、交易频率习惯、交易品类偏好等。
- 3, 负责利用 MapReduce和 Hive 统计客户的交易方式、挖掘潜在客户等, 利用 Hive 分析某段时间、某个地点的客户增长率。

前海中金贵金属东北运营中心 数据库管理员 2013.07-2014.11

- 1, 负责公司 Mysql 数据库的日常维护、备份。
- 2, 负责创建、整理客户信息, 电话数据, 员工信息, 分析客户增长率、员工工作效率等信息给相对应的部门。
- 3, 协调电话资源的分配。
- 4, 完成领导交办的其他任务, 积极主动配合、协助其他部门的工作。

项目经历

高精度地图 python开发工程师 2023.09-2025.03

该项目是一站式高精度地图众源平台。把汽车终端采集到的车道线路标等地图相关语义数据, 绘制成高精度地图。

项目职责:

- 1, 依据产品定制的工艺文档及讨论, 用Python开发, 制作车道线, 路标等相关语义内容;
- 2, 通过QIGS观察到的路线数据及可能情况, 对异常数据提出设计建议, 完成代码开发;

长城大数据平台 大数据开发工程师 2021.07-2023.06

该项目是为长城汽车研发的大数据平台项目。项目部署在金山云服务器, 接收到的车辆信号数据, 通过 SprakStreaming 程序对不同的车型进行解析过滤处理, 将处理好的数据分发到不同的下游 kafka, 之后数据会被进行分析, 包括分析车辆的故障, 驾驶行为情况等。

项目职责:

- 1, 由车辆信号专家给出不同车型的解析策略, 检查解析错误是否有误, 然后将解析策略配置到 Redis中;
- 2, 优化 SparkStreaming 程序对数据解析的效率;
- 3, 维护项目稳定运行;

基线用户画像标签系统 大数据开发工程师 2022.02-2023.02

该项目是公司内部依据企业发展战略, 专为商业运营分析所开发的一套用户画像标签系统。该项目基于开源的 DSS 项目, 通过编写 SparkSQL 代码, 执行定时任务, 将数据处理后为不同的用户打上不同的标签。

项目职责:

- 1, 负责沈阳区域任务的开发情况;

- 2, 开发以 Scala 为基础的 SparkSQL 代码;
- 3, 参与设计数据储存到 hbase与 hive与 clickhouse 的数据结构;
- 4, 调研 DSS 相关组件内容;

AI+赋能交互项目 技术负责人/Python开发 2018.12-2020.07

为了消费者更好的感受机器人服务, 增加智能对话功能, 通过聊天增加机器人的问答, 增强客户的兴趣度, 从而促进客户的购买欲望。收集客户对话数据, 分析数据, 生成常出现料库。

增加人脸识别模组, 根据不同状态, 像消费者反馈不同表情, 使机器人更具备互动性。

项目职责:

- 1, 拆解项目各功能模块, 定制子功能的软件开发思路, 开发部分功能。
- 2, 搭配 Python、nodejs 服务环境, 开发基于 visionSeed 模组的视觉 nodeJS 服务。
- 3, 开发通过 Python 开发集成科大麦克阵列、百度 AST/TTS、图灵语义分析接口的智能语音服务。
- 4, 开发表情、机器人头部云台、语音的集成逻辑实现。
- 5, 采集语音、视觉数据, 最终落实到公司本地集群, 分析各方舱售卖情况与人流情况, 提高客户流量的转化率。

精细化营销项目 大数据ETL工程师 2017.09-2018.05

项目简介:

集成某村镇银行的业务数据报表系统, 为其构建CRM管理平台, 对客户实现银行产品的精准营销。

项目描述:

多家银行组成的联盟, 使用同一个数据仓库, 银行每日只能获取业务报表数据。中小银行想通过对自身的数据信息分析, 增加存量客户对银行的贡献度。通过对不同的客户群体, 实现精准的营销策略。

责任描述:

- 1, 设计结合银行业务系统, 基于FS-LDM模型的GPDB数据仓库。
- 2, 设计开发数据的ETL流程, 通过oozie进行控制。
- 3, 开发oozie中的action内容, sqoop、hive、mapreduce、Python、shell等。
- 4, 优化CDH集群, 减少hdfs的小文件数量, 增加oozie的JVM溢出。

教育经历

辽宁工业大学 本科 信息管理与信息系统 2009-2013

在校期间, 专业课主要方向为ERP, 涉及技术包括SQLServer2000, MySQL, ProwerBuilder, JAVA, 统计学等。