**2020.7 至今 北京森域网络科技有限公司 大数据开发工程师**

1、参与基于大数据的数据仓库项目的前期业务调研工作。

2、参与数仓分层搭建工作。

3、负责离线数仓项目的各项指标的计算。

4、负责企业客户标签制作。

**2018.9-2020.6 天津智信慧通信息技术有限公司 大数据工程师**

1、参与相关业务指标讨论。

2、参与相关业务数仓搭建工作。

3、负责离线数仓项目的各种指标的计算。·

**2015.6-2018.8 华夏幸福基业股份有限公司 运营经理（数据分析）**

1、统计计算各园区月度、季度、年度预算申请及执行，各园区建设节点汇总，并分析数据，制作园区运营相关报表。

2、使用SQL 从数据库计算客户系统的客户的进展、各流程时间、推进进度、签约率等各项指标，制作周、月度等报告。

* 掌握Hadoop，理解HDFS、Yarn、MapReduce 基本原理、运行机制及在业务中的使用。
* 掌握基于Hive数据库进行数仓分层，理解数据仓库设计原理及常用的数仓实现方法。
* 掌握Spark 的Spark Core、Spark SQL、Structured Streaming 三大组件，并能使用其解决相关业务问题。
* 掌握Sqoop数据传送、Zookeeper分布式协调服务、Flume日志采集、Kafka消息中间件等大数据组件工具的使用。
* 掌握Oozie任务调度组件的使用。
* 熟练使用SQL语言。
* 掌握Python编程语言，能够使用Python 进行数据的ETL。
* 熟悉Linux常用操作命令，以及基于任务的Shell 脚本编写。
* 理解HBase的基本原理、列族、Rowkey设计原则，能够使用HBase 进行高效快速检索。
* 了解Flink 框架，FlinkSQL，及基本应用场景。
* 了解FineBI 数据可视化工具，并使用其进行数据的相关展示。

工作经历

专业技能

教育背景

**个人简历**

**项目一：****Acros 生化产品Smartbrain项目**

l **项目时间**：2020.7 – 2021.8

l **项目架构**：MySQL + ClouderaManager + Hive + Sqoop + Hadoop + Spark + Zookeeper + Oozie

l **项目描述**：基于公司多条产品线产生的业务数据，将客户信息、销售相关数据、仓储数据、商品数据、财务数据等相关数据，抽取到数仓，划分为客户主题、订单主题、产品主题、仓储主题、账款主题、物流周期主题等主题。分别按业务需求对数据进行处理分析，最终帮助公司制定相关销售策略、备货策略、生产计划，帮助完善仓储系统、收付账款周期、商品采购及配送到客户的物流周期协调，并作为宣传展示及其他宣传材料的数据基础，制作数据展示看板，供公司高层及时全面了解公司业务运营状况。

l **职责描述**：

1. 主要负责客户主题、销售主题、产品主题相关数据采集及分析。
2. 根据本项目的需求，数仓分为ATL、AML、ASL、ACL、ADM、AAL六层。
3. 用Sqoop将MySQL、PostgreSQL中的业务数据导入到HDFS中，映射成Hive表，采用T+1增量导入的模式。
4. 在AML层做数据轻度清洗、区分维度表和事实表，并根据业务需要制作拉链表。
5. 在ASL层，将各主题相关的表进行合并，形成相关主题宽表，供后续统计分析使用。
6. 在ACL层，分析计算相关业务需求指标，销售主题相关指标如：年度/季度/月度/每日，各大区/省份/城市，各品牌/产品线/自有品牌/主推品牌，直接客户/经销商等的销售额、销售毛利、利润、毛利率、利润率、销售同/环比增长率、报价单转化销售单率、销售额占比率、利润占比率、退货率、售后单占比、销售量等。客户主题包括各客户/不同行业/不同性质/直接/间接客户的销售额、利润率、购买的TOP10产品、销售额占比、购买时间段、付款周期等。商品主题包括不同产品线/不同品牌/自有品牌的TOP15产品、销售量、折扣率、售后率、平均到货周期等。
7. 在ADM层按月、季度、年粗粒度上卷，计算最终需要展示的指标，将结果存储在AAL层。
8. 使用Sqoop将最终将结果导入到MySQL中，供后续使用。
9. 使用Oozie调度，每日定时实现相关任务。

项目经验

3、使用SQL从数据库计算不同园区/企业，月度/季度/年度的租金总额、固定投资额、投资额、租赁面积、入驻率、空置

率等，并制作相关报表。

4、组织参与计划运营会议。

5、配合平台产业服务部门，共同推进园区企业入驻和项目落地，促进项目尽快达产。

**2011.7-2015.5**  **上海安谱实验科技股份有限公司 销售工程师 & 销售部总监助理(数据分析)**

1、开发新客户和维护大客户，完成区域销售业绩。年度销售额增长率维持在40%以上。

2、解决售前和协助解决售后问题，组织技术培训及讲座。区域客户满意度问卷调查反馈，满意度在95%以上。平均每年组织相关技术培训及讲座12场以上，每场1-3天不等。

3、经销商管理。完成区域内经销商合作谈判、达成合作。协助经销商完成招标、解决售前售后问题。

4、业务相关数据统计及分析。协助销售部总监，使用Excel进行月度、季度、年度业务指标统计及分析，包括区域销售情况，区域商品销售情况等各类关键指标。

5、数据可视化展示。根据数据分析结果，使用Excel和PPT制作数据可视化图表等，并制作PPT汇总材料。

**项目二：****Acros企业客户标签系统项目**

l **项目时间**：2021.6 - 2022.3

l **项目架构**：MySQL + ClouderaManager + Hadoop + Hive + Elasticsearch + Zookeeper + Spark + Oozie

l **项目描述**：Acros企业客户标签系统，是针对系统交易的企业客户，根据记录的基本信息、历史交易数据等提取相关信息，以及根据历史数据对未来做出预测，形成相关标签，再使用Spark技术栈，为企业客户打上相关标签。使用数仓ATL层数据导入ES作为制作标签系统的原始数据，使用Spark加载系统标签数据和ES中的企业客户基本信息、销售订单、出库单、付款记录等数据，开发了规则匹配类、统计类标签。为公司以后的客户策略、精准营销、产品配套主题等提供数据支撑。

l **职责描述**：

1、参与企业客户标签整个项目架构、指标体系搭建讨论。

2、使用ES-Hadoop方式，在Hive创建ES的外部表，使用HiveSQL将Hive数仓中所需要的数据导入ES。

3、实现了规则匹配类、统计类标签的开发，包括企业所属行业分类、企业性质、标准制定/参与资格分类、RFM、购买月度偏好、购买产品线偏好、购买品牌偏好、付款周期等在内的标签，共计150余个。

4、使用Spark读取ES制作标签的原始数据，以及标签元数据，进行标签的匹配计算，并将最终结果保存到ES中。

5、基于开发任务，抽取开发模板，后续标签基于模板开发，提高开发效率。

**项目三：Acros日志采集项目**

l **项目时间**：2022.4 - 2022.9

l **项目架构**：Flume + Kafka + Spark + MySQL

l **项目描述**：通过分析nignx日志，得到用户的pv和uv，解析ip地理位置、状态码、访问时间等信息；通过分析用户行为事件数据，得到浏览页面排行、页面停留时间、标准品COA证书下载量、下单总量等各类指标。从而判断网站受欢迎程度，为公司网站运营提供决策依据，同时也可以做异常用户筛选，保障网站稳定运行；判断受欢迎的产品、客户区域等各类情况，为网站的业务运营提供数据支撑。

l **职责描述**：

1、使用Flume采集网站Nginx日志数据及用户行为日志数据到Kafka。

2、使用Spark结构化流加载Kafka数据，并进行实时数据ETL，和各类字段解析及指标计算。

3、将统计及计算好的信息及指标统计写入到MySQL中，供后续数据展示使用。

**项目四：****银行****智能营销平台项目**

l **项目时间**：2019.4 - 2020.6

l **项目架构**：Oracle + ClouderaManager + Hadoop + Sqoop + Hive + Spark + Zookeeper + Oozie

l **项目描述**：“智能营销平台”定位服务于一线营销人员和营销运营人员，系统集整合化、实时化、智能化、自动化为一体，以大数据技术为核心，引入实时数据处理引擎，融合机器学习技术，整合客户数据，细分客户标签画像，聚焦精准营销，构建“客户标签画像 + 营销资源整合 + 精准营销”体系。深度结合互联网营销模式，构建高效的协同运营能力，

提高了营销的精准度及工作效率，打造极致的客户参与体验。

l **职责描述**：

1. 负责客户主题和贷款主题各类指标的计算，存入指标库供后续应用。

* 逻辑思维清晰，关注细节、做事踏实，具有较强的沟通、协作能力。
* 对新事物、新技术感兴趣，自主学习能力强。
* 具有较强文案撰写能力。

自我评价

2、从缓冲层获取数据，制作主题宽表，存入基础整合层。

3、根据业务需求，进行各类指标的计算，并放入指标库中。主要开发的指标有各类贷款客户数、网络金融产品客户数、逾期客户数、各项贷款余额、贷款余额增量、贷款加权平均利率等。

4、对相关脚本进行调度执行。

**项目五****：智慧园区大数据项目**

l **项目时间**：2018.9 – 2019.3

l **项目架构**：PostgreSQL + Hadoop + Hive + Sqoop + Oozie

l **项目描述**： 1. 针对入驻及孵化企业，整合、分析企业经营及发展数据，进行精细化管理和风险控制，针对企业实际情况，制定针对性的扶持和服务，进一步完善企业服务平台。同时数据分析将会作为园区多类宣传材料展示。2. 通过不同区域企业、不同行业企业、最终入驻客户和意向客户比率，入驻周期，接洽频率等数据的分析，为园区招商的方向、方式提供数据支撑。3. 针对物业管理相关数据，如缴费信息、车流信息、维修记录等数据做分析，用于指导制定相关费用标准，提高物业服务效率及质量，以及供总部分析决策使用。

l **职责描述**：

1. 主要负责园区企业经营发展和客户系统的数据整合分析。
2. 数据库分为ODS、DM、RTP三层。
3. 从园区平台数据库和客户系统数据库用Sqoop导入数据到ODS层。

4、在DM层对各项业务相关指标进行计算，并将计算好的指标导入RTP层。