|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 梁川川简历 | | | | |
|  | | | | |
| **姓 名** | 梁川川 | **出生年月** | 1993.04 |  |
| **电 话** | 18511866554 | **邮 箱** | job\_lcc@sina.com |
| **求职意向** | 大数据开发工程师 | **学 历** | 本科 |
| **工作年限** | 2015.03 至今 | | |
|  | | | | |
| 技术点 | | | | |
|  | | | | |
| * 熟练使用 Spark Core、Spark SQL和Spark Streaming，理解Spark、Spark Streaming任务运行流程，有对Spark程序的调优经验 * 熟练使用 Kafka，理解High Level和Low Level API * 熟练使用 Hive，熟悉Hive工作流程，有对Hive进行调优的经验 * 熟练使用 HBase，熟悉HBase工作流程，能独立设计RowKey * 熟练使用 MapReduce，理解Yarn任务运行流程，理解MapReduce shuffle流程 * 熟练使用 Hadoop、HDFS、MapReduce、Yarn，能对MapReduce程序进行调优，能独立完成集群环境部署 * 熟练使用 Sqoop，抽取业务数据到数仓，了解DataX * 熟练使用 MySQL、MongoDB 能独立设计业务表及优化 * 熟练使用 Redis常用命令、持久化方式、信息发布与订阅、主从复制(主要实现读写分离业务) * 熟练使用 JavaEE技术、Spring、SpringMVC、MyBatis、Dubbo、Echarts等 * 熟练使用 Git、Git Hub * 熟悉 Linux命令及项目部署 * 理解 JVM内部结构、GC区域及原理 * 了解 Elasticsearch、Logstash、Kibana * 熟练使用 Java、Scala | | | | |
|  | | | | |
| 工作经验 | | | | |
|  | | | | |
| 2016-09至今 北京东奥华宇国际教育科技有限公司 大数据开发工程师   * 负责登陆分析系统、商城分析系统的设计、开发工作 * 负责部分实时计算数据的开发维护 * 参与数据模型体系构建及数据主题设计和开发，搭建离线、实时数据仓库   2015-03 2016-08 北京天诚安信科技股份有限公司 Java开发工程师   * 定制化安全产品-统一信任管理平台，客户需求分析，方案写作等工作 * 系统模块开发，测试 | | | | |
|  | | | | |
| 项目经验 | | | | |
| **项目名称：东奥-商城分析系统** 2018.03 - 至今 | | | | |
| **项目描述:** | 该项目主要是对东奥商城网站的流量进行离线和实时的统计分析  系统分析课程和图书销售、学员会话统计、学员页面跳转率、促销大屏直播等通过JavaEE展现给运营人员实现销售方案变更，展现给业务部门实现产品不断迭代更新  最终达到用大数据技术来帮助提升公司的业绩、营业额以及市场占有率的目标 | | | |
| **使用技术:** | Flume、Kafka、Spark、Hive、HBase、HDFS、Echarts | | | |
|  |  | | | |
| **技术要点:** | * 使用Spark累加器，对学员Session占比进行累加，在Driver端取出数据计算占比 * 通过广播变量优化程序，提升执行效率 * 自定义排序Case Class，使用sortByKey对热门课程进行排序 * 使用Spark Core实现用户访问Session统计以及页面之间的单步跳转转化率 * 使用Spark Streaming算子reduceByKeyAndWindow统计3小时内课程和图书销售情况 * 使用Spark Streaming算子updateStateByKey统计促销活动，用Echarts向运营展示大屏直播 * 使用随机key实现Spark Core双重聚合 | | | |
|  |  | | | |
| **责任描述:** | * 课程和图书销售渠道模块数据分析 * 学员会话模块、学员页面跳转率统计 * 数据倾斜优化 | | | |
|  | | | | |
| **项目名称：东奥-登陆平台分析系统** 2017.09 - 2018.02 | | | | |
| **项目描述:** | 一款APP上线后，如会计云课堂、题库宝典、PC豪华版等，公司需要了解学员对这款APP的详细使用情况  项目由离线数据处理系统和实时数据处理系统两个部分组成，对新增学员数、活跃学员数、沉默学员数、各平台学员占比等进行分析统计  公司根据APP使用人群画像和学员使用反馈对APP实行针对性的改进 | | | |
|  |  | | | |
| **使用技术:** | Flume、Kafka、Spark、Hive、HBase、HDFS | | | |
|  |  | | | |
| **技术要点:** | * 登陆平台进行日志埋点，学员登陆使用Log4j记录登陆信息 * 多层Flume架构，使用sink group * Flume TaildirSource读取数据模块修改，解决Log4j按天滚动后，日志文件名发生修改，Flume默认文件名变更重新读取文件，从而导致数据重复 * HBase预分区，rowKey重新设计解决热点问题 * Kafka分区分配策略修改为 * 手动维护Kafka offset，使用连接池技术代理Kafka连接的创建等工作 | | | |
|  |  | | | |
| **责任描述:** | * 参与项目分析以及数据埋点工作 * 编写Shell脚本定期将采集日志进行ETL并把数据导入到Hive中 * 分析业务指标，例如新增用户、活跃用户、沉默用户、启动次数、版本分布等 | | | |
|  | | | | |
| **项目名称：东奥-数据仓库** 2017.03 - 2017.08 | | | | |
| **项目描述:** | 公司业务不断发展，业务数据、学员行为数据逐渐庞大，为业务智能，提供指导业务流程改进、成本、以及质量控制  在服务端进行数据埋点，采集业务数据、用户行为数据，对数据进行分层采用更多宽表化设计，提升公共指标的复用性，减少重复加工  为公司决策性分析系统构建了基础 | | | |
|  |  | | | |
| **使用技术:** | Sqoop、Oozie、Flume、Kafka、Hive、HDFS | | | |
|  |  | | | |
| **技术要点:** | * 使用Sqoop接入业务数据，Flume、Kafka接入学员行为数据 * 使用Ozzie调度Hive脚本，对不同数据层数据进行计算清洗构建 * 参与数据分层ODS（操作数据层）DWD（明细数据层）DWS（汇总数据层）ADS(数据应用层) DWD层构建维度、事实表，DWS层对数据进行轻量汇总、宽表设计 * 开启Hadoop JVM重用机制 * 根据实际业务选区分区方式、Parquet为存储方式、Snappy为压缩方式 * 开启Hive 严格模式，限制部分查询条件避免不合理的HQL执行 * 根据具体业务数据量评估采用全量同步或增量同步 | | | |
|  |  | | | |
| **责任描述:** | * 对Hive进行调优 * 定义Hive UDF、UDAF函数 * 协助数据分析体系的数据建模、统计及分析 * 参与数据仓库系统的优化及计算性能提升 | | | |
|  | | | | |
| **项目名称：东奥-答疑系统** 2016.09 - 2017.02 | | | | |
| **项目描述:** | 为了给学员更好的学习体验，全面提升学习效果，东奥针对目前已有的课程推出答疑系统  项目由MySQL作为后台数据支撑，Redis缓解数据库压力，使用Dubbo作为分布式服务共其他业务系统使用  解决学员听课中的疑问，看书中的疑问，做题中的疑问，让学员在问题交流中迅速提高 | | | |
|  |  | | | |
| **使用技术:** | Spring、Spring MVC、Redis、Dubbo、MyBatis、Solr、MySQL、UEditor、Echarts | | | |
|  |  | | | |
| **技术要点:** | * 使用Redis缓存业务数据 * 分析查询慢、数据量大的表适当添加索引优化查询效率 * 使用Solr对学员提问内容进行索引构建 * 修改IK分词器源码支持粗细粒度分词 * 使用Dubbo为不同业务提供分布式服务 * 使用MyBatis generator生成实体表 * 使用Ajax异步刷新页面 | | | |
|  |  | | | |
| **责任描述:** | * 根据业务需求与其他系统协作开发并实现各产品功能 * 负责接口对接、测试 * 参与后端研发中涉及到的技术框架选型及相关实现 | | | |
|  | | | | |
| 自我评价 | | | | |
|  |  | | | |
| 深度互联网技术，对技术保持高度的敏感性和关注度  每周一本技术书籍 | | | | |