

个人简历

◆ 基本信息

姓名	王美华	性别	女
学校	华中农业大学楚天学院	专业	软件工程
出生年月	1992-8-6	学历	本科
英语水平	CET-4	联系电话	15807185224
求职意向	大数据开发工程师	电子邮件	15807185224@163.com

◆ 专业技能

熟练运用 JAVA 基础和面向对象编程思想；
熟练编写 Scala 代码，理解 Scala 函数式编程的思想；
熟练 Spark 的 RDD 原理，常用算子、SparkStreaming、SparkStruts、SparkSql 的编程模型；
熟练运用 Oracle、MySQL 关系型数据库；
掌握编写 javascript 脚本、html 静态网页；
掌握运用 AngularJs 前端框架；
掌握运用 Struts+ Spring+ MyBatis 开源框架；
掌握 Hadoop 以及 Hadoop 生态圈软件，熟悉 HDFS 存储、MapReduce 计算以及 Yarn 资源调度；
掌握 Hbase 数据库的使用，了解 Hbase 的常用二级索引；
掌握数据的 ETL 实现过程，以及 Hive 的使用，完成对数据主题抽取和多维分析；
掌握 Sqoop 框架，能够使用 Sqoop 对数据和数据库之间进行导入导出；
掌握 Flume 日志采集流程，了解 Flume 的负载均衡和容错机制；
掌握 Zookeeper 的运行机制，选举机制，以及通过 Zookeeper 实现其他软件的高可用；
掌握 Redis 内存数据库的基本原理；
掌握 kafka、logstash 数据采集工具的原理，实现流式数据的过滤和分析；
掌握 Flink，Azkaban 等其它大数据常用框架；
掌握 Linux 系统，熟悉常用的 Linux 的 Shell 命令，能够编写基本的 Shell 脚本；
了解 ELK 技术栈，了解 ElasticSearch，Logstash 的整合使用

◆ 教育背景

时间	学校	专业	学历
2012.9-2016.6	华中农业大学楚天学院	软件工程	本科

◆ 工作经历

工作时间	单位名称	职位名称
2016.7-2020.04	上海恩梯梯数据晋恒软件有限公司	Java 软件开发工程师

◆ 项目经验

项目一：优惠券分发数据分析系统

开发环境：IDEA3.1、Jdk1.7、SVN

系统架构：Scala2.11.8、Spark2.2、Hadoop2.7.1、Kudu1.7

项目简述：商户为了盈利，往往会发各种优惠券刺激消费者消费。本系统是通过研究用户原始数据，分析出各种指标和报表，用于分析出哪些商圈是消费者经常逛，人数多的，可以针对特定商圈发特定优惠券。又或者哪些商品是消费者买的最多的、优惠券使用率有多少等等。

1. 对第三方提供的原始数据(网络请求的 json 串)进行探索，得出规律。
2. 归纳数据特征，根据业务需求进行 ETL 操作，将清洗好的数据使用 Oozie 等工具进行调度执行，处理过的数据落地到 kudu 中。
3. 由于同一个用户每次的请求数据汇报的 id 不同，通过某些关联关系，识别出同一个用户。
4. 编写代码为每条数据打上对应的标签，落地到 Kudu，向第三方提供服务。

责任描述：

1.通过 zeppelin 工具，编写 sql，从结果的图形显示中，得出数据的大致规律。比如，原始请求数据中，哪种终端最多，哪种优惠券被使用最多次，最受欢迎，有效请求占比多少等等。

2.编写公共模块，提供对数据库的统一访问和配置信息（spark 集群心跳超时时间等）的读取

3.通过 map、filter 等算子，过滤分析出来的无效数据。

4.通过 ip2region 和 GeoLite 工具把原始数据中的 ip 信息扩充一组地域信息（省市，经纬度）。

5.编写代码，生成各类报表入库到 kudu，如按照地域统计数据集分布情况，优惠券使用情况等。

6.通过 ETL 阶段入库的经纬度信息，使用高德地图提供的逆地理 API 返回各个请求的商圈信息。

7.将获取到的商圈信息的 json 字符串转化成商圈对象，生成经纬度范围对应的商圈信息表入库。

8.给 ETL 后的每条数据，通过经纬度范围打上商圈信息的标签和其他标签入库。

9.通过 GraphFrame 图计算组件，绘制出关系图，相同 component 的记录为同一个用户。

10.通过 groupByKey、reduceGroups 算子将识别出的相同用户的标签合并，生成新表入库。

项目二：用户行为指标分析系统

开发环境：IDEA3.1、Jdk1.7、SVN

系统架构：Scala、Hadoop、HBase、Kafka、Flink、Canal

项目简述：该系统主要是通过第三方服务获取的用户行为（访问行为、购物行为、点击行为等），统计出 PV、UV 等指标。主要功能如下：

1. 将第三方服务实时获取的用户行为数据推送到 Kafka
2. 将第三方服务获取的离线数据从 Mysql 通过 Canal 推送到 Kafka
3. 实时分析模块消费 Kafka 数据，经过分析后，下沉到 HBase
4. 离线同步模块，消费 Kafka 数据，下沉到 HBase
5. 批处理分析模块从 HBase 取出数据，进行数据分析

责任描述：负责 1、2、3、4、5 功能核心代码的编写及调试。

项目三：优惠券分发系统

开发环境：Eclipses、Jdk1.7、SVN

系统架构：AngularJs、SpringMVC、MyBatis

项目简述：该系统分为前后端，前端负责管理优惠券，主要功能是，注册各种类别的优惠券，不同优惠券的使用场景不同。更改优惠券，控制只在某种情况下才能更改。查看优惠券，检索优惠券，删除

优惠券等功能。其中系统角色分为用户和客服，用户拥有所有操作权限，客服只有查看优惠券的权限。

该系统后端主要接受后端和第三方服务的 http 请求，通过不同请求获取请求信息，根据请求信息不同，实现了包括用户注册，注册新的优惠券，优惠券分发，优惠券被某个用户使用的记录，不合法请求的识别，删除优惠券的功能。

责任描述：

1. 负责要件分析，分析客户提出的功能的具体实现并且对有歧义的地方着重探讨和确认。
2. 根据要件编写前后端的式样书。
3. 根据前后端式样书编写代码。
4. 前端使用 Angularjs 框架，通过 watch 机制，监听非法的优惠券信息录入。
5. 通过 HTTP 请求后端的服务，获取优惠券信息和注册使用优惠券。
6. 代码编写完成后，进行单元测试的式样书和代码编写。
7. 完成单元测试，并做好证迹（单元测试的输入、输出、代码覆盖率等的截图）
8. 重复 6，7 两个步骤，完成集成测试和系统测试。

项目四：银行卡收单系统

开发环境：Eclipses、Jdk1.7、SVN

系统架构：Spring、IBatis

项目简述：该系统分为前后端，前端负责管理加盟店，主要功能是登录或者通过 csv 文件导入各种店铺。修改店铺基本信息和精算信息。查看店铺信息和删除店铺等功能。

该系统后端主要做店铺和银行之间的账单结算，大致分为以下步骤：

1. 批量读入服务端刷卡产生的数据文件，进行格式检查，当错误级别为文件级别的时候，当前整个处理流程不再往下进行，当错误只是明细级别的时候该条明细数据入错误明细表，其它正确的明细写入正确的明细表，不终止当前流程，继续下面的操作。
2. 根据明细表里每条数据用户使用的支付方式不同，卡总类不同算出交易额和服务费，根据每个店铺和银行的结算方式不同，确定结算日。
3. 对错误明细表里的数据通过人为改动进行再结算的处理。

责任描述：负责前后台核心代码和式样书的编写与单元测试。

项目五：自动化测试工具

系统架构：VBS

项目简述：由于银行卡收单系统太庞大，每次修改了部分功能在上线前都需要做一个全面的系统测试。对于繁复的测试流程，导致每一次的小改动，代码编写阶段花费的时间很少，但是测试阶段花费了大量的人力物力。因此而做的一个自动化测试的工具。

责任描述：负责整个工具的做成。使用公司内部提供的 WinActor 工具，根据业务需求，通过 VBS 编码，实现页面动态登入值，控制流程完成整个测试任务。

◆ 获奖或荣誉情况

“华三杯” 2015 全国大学生网络技术大赛湖北赛区二等奖
2015 年大学生数学建模国赛省三等奖

◆ 自我评价

- 1.为人随和，兼具团队协作精神。
- 2.拥有良好的学习态度和能Ⓒ力，积极的工作态度、较强的责任感。
- 4.能够积极的面对工作中遇到的问题，合理的处理工作中面对的压力。