

基本情况

姓名: XXXX

电子邮箱: XXXX@qq.com

毕业院校: XXXX 大学

移动电话: XXXX

专业方向: 软件工程

年龄: 26

IT 技能

技能名称	熟练程度	使用时间
Hadoop	精通编写 MapRduce 程序, 善于抽象业务逻辑, 掌握 hadoop 的安装及运行原理。	3年
Hive、Hbase	熟练掌握、能安装运维以及编写相关代码。能设计 hbase 表结构, 实现存储与查询, 优化 HHive 查询问题。	2年
机器学习、Mahout	理解机器学习概念掌握基本的机器学习算法, 熟练运用 mahout 中提供的机器学习算法。能够读懂一些 R 语言设计的東西。	2年
Spark	熟练掌握 Spark 的运行原理及代码编写, 熟练使用 scala。	1年
Storm	熟练。运用 storm 进行过数据清洗和实时统计分析。	1.5年
OpenStack	能搭建复杂网络环境下的 OpenStack, 建设公司的私有云。	0.5年
PLSQL	精通。基本 SQL 增删查改、游标、动态游标、动态 SQL、DB_LINK、EBS 财务方面标准 API 使用。能使用 Oracle 进行千万级别数据的查询优化, 能熟练使用 PLSQL 编写极其复杂的存储过程, 能实现 oracle 多实例数据库之间的数据推拉和相互业务间的逻辑调用。	4年
数据库	熟练使用 Oracle、mysql。能独立进行关系型数据库的设计建模, 解决过 Oracle、MySQL 上亿条数据统计计算慢的问题。	4年
Java	熟练使用 java, 掌握 SSH 框架。独立设计实现过聊天服务器后台应用, 以及实现过 java 项目的缓存和定时任务模块。	4年
杂项	熟练掌握 Sqoop、Flume、Kafka、Redis、Memcache、Nginx、FastDfs、Netty、Protobuf。掌握他们的安装使用及运行原理。	

工作经历

公司	职务	时间段
上海 XXX 息技术	EBS 技术研发、大数据研发	2011.9-2013.6
XX 网	大数据技术研发	2013.6-2015.10

自我评价

实际操作、动手能力、能够解决各种棘手的问题。逻辑思维能力强, 设计思路清晰, 自主学习能力强。特别是在汉得的工作中锻炼出了较强的抗压能力, 沟通能力。工作态度认真负责做事严谨。

项目经历

● XXXX 网络（用户实时推荐系统 2014.4-2015.3）

项目描述：Flume+Kafka+MapReduce+Storm+Spark+HDFS+Hbase+Redis+Solr

拉手有很多历史数据，包括用户数据、商品数据、商家数据、用户交易记录、用户位置轨迹信息。这些数据极具价值，对其进行挖掘可以做到更好的用户体验。故而决定利用这些数据做一个推荐系统。

职责描述：

- 1、数据清洗，清除作弊、刷单、代购等噪声数据。这些数据的存在会严重影响算法的效果，因此要在第一步的数据清洗中就将这些数据剔除。
- 2、对用户画像进行补全
- 3、实时用户行为分析
- 4、对每个用户进行个性化推荐
- 5、热点地区计算
- 6、热门商品计算

技术描述：

- 1、利用 Storm 进行数据清洗。
- 2、利用现有数据中用户信息比较完整的记录训练一个贝叶斯分类器，利用这个贝叶斯分类器进行用户画像补全。
- 3、根据用户出现的地理位置，采用协同过滤的方式计算用户的相似度。
- 4、利用 spark 分别计算基于用户的协同过滤和基于物品的协同过滤以便将这两者进行结合使用。
- 5、根据用户的历史消费记录或者浏览记录通过 Spark 程序挖掘出某一个区域内的区域消费热单和区域购买热单。
- 6、热门商品分为区域购买热单和区域消费热单，当新的线上用户请求到达时，根据用户的几个地理位置对相应地理位置的区域消费热单和区域购买热单进行加权，最终得到一个推荐列表。

● XXXX 电信(定向广告投放及基站监控项目 2013.2—2013.5)

项目架构：flume+kafka+HDFS+Storm+Hbase+MySQL

项目描述：

随着智能手机的普及，用户通过移动终端上网的越来越多，使得电信公司积累了大量用户日志，同时使得电信公司的基站性能受到了考验。贵州电信希望将产生的日志信息利用起来，而且需要实时监控各个基站的压力情况。

责任描述：

- 1、实时监控基站，提供快速网络调优预警
- 2、统计那些地点为用户常态地点
- 3、统计那些终端限制了用户对业务的需求，那些终端值得推广
- 4、得出重点推荐用户

技术实现：

- 1、通过现有用户登记用户数据利用贝叶斯分类器训练出一个模型，运用这个模型对用户

画像进行补充。

- 2、运用 MapReduce 程序计算用户的常住点
- 3、利用用户来往电话记录运用 PageRank 算法来计算出用户的重要程度，我们称之为 peopleRank 值，筛选出 peopleRank 值比较高的用户。利用这些用户来计算出以这些用户为中心点的二级关系圈，这些重点用户就成为这个圈子的代言人了，重点向这些用户进行广告推荐。
- 4、通过 Storm 来实时处理用户与基站的日志得出每个基站承载的用户数目

● XXX 网络（拉手交易量实时监控系統 2013.11—2014.3）

项目描述：

需要实时监控拉手各个大区的交易额变化情况，用户在线情况

项目架构：Flume+Kafka+Storm+Redis

技术实现：

- 1、利用 Flume 收集主站 PHP 服务器上日志信息传入 kafka
- 2、利用 Storm 流式处理出入的日志信息，将用户的交易日志进行统计计算
- 3、将计算好的数据存入 Kafka 中，将 kafka 中的数据导入 redis 和 Hbase 中

● 拉手网络（私有云搭建 2015.6-2015.9）

项目描述：

公司有很多需要运用大数据进行计算分析的项目，为了各个开发之间不相互影响而又能节约资源，决定搭建自己的私有云计算平台。

职责描述：

- 1、研究 Openstack 技术
- 2、计算搭建私有云平台的硬件需求及成本
- 3、搭建 Openstack

● 拉手网络（大数据计算平台搭建项目 2013.7—2013.10）

项目背景：

2013 拉手的拉手在团购行业所占市场份额比较大，活跃用户和用户交易量都很大，为了将这些数据利用起来和解决交易大表日益膨胀的问题和保持良好的竞争力拉手决定建立自己的大数据计算平台。

职责描述：

- 1、根据现有的业务量计算出未来三年拉手大数据平台所需要的基础硬件配置
- 2、搭建 Hadoop、Hive、Hbase

3、将现有业务数据导入到大数据计算平台

● 上海汇众汽车(汇众 U 平台——企业制造管理和内部 OA 项目 2012.6—2013.1)

项目架构：SpringMVC+Hibernate+Memcache+Netty+Nginx+MySQL 集群

项目描述：

由于汇众的员工数量巨大，汽车零配件业务的不断发展原有的管理沟通方式已不太适合汇众当下的规模要求。为了公司的发展需要公司决定上马新项目。项目包含以下：模块仓储管理和展会管理；货运流程管理；决策分析；OA 系统。该项目已于 2013 年 10 月进行了第一次上线。

责任描述：

该项目制造业模块主要是集成 SAP 的基础部分然后结合一套自己设计开发的界面，这方面由其他同事负责。我主要负责 OA 系统中内部消息推送服务端模块，该模块要求如下：跨平台，实现 PC、IPC 和移动端的无障碍通信。

- 1、实时性
- 2、高并发
- 3、后期可扩展性

技术实现：

- 5、整个项目是基于主题是基于 JavaWeb+Netty 的，项目全部部署在 tomcat 中并且支持横向扩展。
- 6、采用基于 protobuf 自己设计了一套通信协议，以满足跨平台的功能。
- 7、使用 memcache 存储客户端与服务器之间的链接关系，然后通过 web 服务来调度客户端链接服务端资源分配的问题。
- 8、运用 SpringMVC 和 netty 向外部暴露两套接口，满足不同平台的消息推送需求。
- 9、运用 nginx 的反向代理解决高可用问题。
- 10、在 web 项目中启动一个自主资源监控线程实时监控自身资源使用情况并写入 memcache。