

易斌鑫

应聘信息

- 应聘职位：数据开发工程师
- 学历专业：本科/西安电子科技大学/数学与应用数学
- 基本信息：汉族/1993年3月/男
- 电子邮箱：ybxlife@foxmail.com
- 手机电话：187-1084-9538

工作经历

- | | | |
|-----------|------------------|--------------------|
| • 数据开发工程师 | 西安影视迅捷技术有限公司 | 2022年3月 ~ 2022年 5月 |
| • 软件开发工程师 | 北京外企德科人力资源服务有限公司 | 2021年6月 ~ 2021年11月 |
| • 数据开发工程师 | 迅捷联动科技有限公司 | 2018年1月 ~ 2021年 5月 |
| • 软件开发工程师 | 西安优云软件有限公司 | 2017年6月 ~ 2017年12月 |
| • 软件开发工程师 | 西安导学教育科技有限公司 | 2014年7月 ~ 2017年 5月 |

项目经验

大数据平台（2018年1月 - 2021年5月）

- 从0到1搭建并运维大数据技术体系，为用户画像分群、OLAP分析、推荐等业务提供服务；
- 设计实施数据平台技术方案（Hadoop、Spark、Kafka），开发离线ETL工具组件、搭建Airflow任务调度系统、设计元数据管理系统；
- 优化日志中间件Fluentd、Canal服务，搭建并维护Kafka集群，搭建Kafka Stream流式数据处理服务；
- 数据分层建模及离线数据分析模型（ROI、LTV）开发，完善数据仓库集成、即时查询及可视化功能；
- 基于Spark GraphX开发千万级N度社交关系推荐数据计算任务。

AB实验平台（2020年3月 - 2020年9月）

- 设计AB实验分流系统架构，实现实验分层、流量分域、白名单规则、用户受众规则；
- 采用Redis、Mysql、ElasticSearch技术开发实验管理服务，发布到阿里云自建K8s环境；
- 应用多实验分层，互斥实验分组和简单的流量百分比自动调整，完成流量及事件指标配置和统计；
- 减少业务开发的工作量，仅一次API调用，使产品业务优化探索效率提升一倍。

内容推荐系统（2019年5月 - 2020年3月）

- 开发特征工程程序，开发训练FM模型，提升AUC指标至0.9；
- 开发推荐服务端API，实现多路召回、重排序及冷启动策略，提高内容分发效率；
- 实现模型分布式、增量训练和小时级更新，通过Faiss、Redis提高Embedding向量检索速度；
- 基于BERT实现word2vec，增加内容文本I2I召回，提升内容主题相似度。

日志收集监控服务（2016年7月 - 2017年5月）

- 研究和实践ELK技术栈，使用轻量级Filebeat工具进行日志收集，减轻服务器压力；
- 编写Nginx、Tomcat的Logstash日志处理器，完成日志格式预处理，细化Kibana数据看板；
- 搭建CDH数据集群，实现业务数据集成处理，搭建Solr分布式索引服务。

K12题库搜索系统（2014年9月 - 2016年6月）

- 优化Lucene索引和OCR模糊搜索算法，提高图像处理效果和上传速度；
- 搭建Solr分布式搜索系统，完善索引词典建设，编写LaTeX分词算法，提升10倍搜索请求QPS；
- 实现HTML+LaTeX文本内容图片、PDF、DOCX格式转码服务，通过多任务队列实现并发快速处理千万任务；
- 参与翻译Solr使用指南，研究Pandoc源码，Patch开源项目。

我的技能

- 英语等级：CET-6
- Java、Scala、Python、Shell、Linux、Docker
- Hadoop、Spark、Kafka、Airflow、Canal、Spring

教育经历

- | | | |
|------------|--------------|-------------------|
| • 西安电子科技大学 | 本科 / 数学与应用数学 | 2010.09 - 2014.07 |
|------------|--------------|-------------------|

自我评价

- 热爱学习，不惧怕新技术
- 热爱开源，喜欢阅读技术书籍和博客，喜欢分享讨论