**李澍楠**

男 | 户口-黑龙江-哈尔滨 | 现居住于上海

4年工作经验 | 本科

手机；18811624089

E-mail；li\_shunan@163.com

**求职意向**

**工作性质**：全职

**期望职业**：IT管理/项目协调

**期望行业**：IT服务（系统/数据/维护）

**工作地区**：上海

**目前状况**：目前处于在职状态

**专业技能**

1. 熟练掌握hadoop生态圈系统CDH组件，如Hive，Hbase，Spark
2. 熟练掌握流批处理框架flink运行时工作原理，设计开发测试流式数据处理算子。
3. 熟练使用java，python开发语言，springMVC，springboot，flask等服务框架。
4. 熟悉各种数据库和搜索引擎的使用，如Mysql，Redis。ES，Solr，clickhouse。
5. 熟练运用Docker，gitlab，Jenkins等集成kubernetes实现代码CI/CD部署。
6. 熟悉一些典型的算法，例：关联规则，聚合，逻辑递归，神经网络。

**工作经历**

**2020/5-至今 九章云极**

职务：中级研发工程师Ⅰ

**工作描述：**主要负责项目进度跟踪，跨团队沟通，资源协调，以及平台系统功能点设计，代码开发和系统调试。

1. 实现基于大数据环境下的数据抽取，处理，分析与计算。
2. 实现基于微服务架构下系统后端服务请求处理与响应开发测试。
3. 实现基于流批处理框架下数据的流式处理与转换算子设计，开发，调试。

**工作内容：**

1. 负责海量数据抽取，处理，存储。以及高并发场景下交易查询优化。完成投产部署。
2. 负责系统前后端接口设计，开发，调试，优化，与fate原生代码交互，算法运行。
3. 负责系统流式功能算子的开发调试，优化高并发，低延迟场景下功能算子tps等指标。
4. 负责跨组之间的项目沟通，进度协调，编写部署投产文档模板，实现稳定投产部署。
5. 负责调研并调试fate原生算法支持，多host下，统计分析（均值，方差，中位数），机器学习（逻辑回归，SecureBoost），神经网络（DNN）等。
6. 负责原生fate工作流和模型仓库管理模块开发，调试。如工作流管理，以及模型预测，
7. 负责日志模块管理，包括计算日志入库，联合任务凭证管理，存证发现追踪等。

**2019/2-2020/4 文思海辉金融-中央国债新一代系统建设**

职务：高级数据工程师

**工作描述**：着重处理企业内非结构化、半结构化以及海量的结构化数据。依托大数据处理、分析、挖掘手段，为管理分析类应用提供数据支持。

**工作内容**：

1. 基于Hadoop平台的系统管理，代码开发优化，以及开发文档的编写。
2. 利用统一调度工具Control-M开发配置大数据平台的工作任务调度。
3. 利用ETL工具informatica开发配置优化数据流管道。
4. 编写并调优Sql，解决大数据平台和数据仓库不同层级之间的数据对接。

**2017/8-2019/1 明码（上海）生物科技有限公司**

职务：·大数据软件开发工程师

**工作描述**：基于基因大数据的系统开发、维护、部署上线，保证医疗基因数据（VCF）高效读写、海量存储、实时分析等系统操作的稳定运行。构建基因医疗大数据的系统体系。  
**工作内容**：

1. 基于CDH版本Hadoop大数据系统的基础搭建、平台维护、代码调优。
2. 结合Jenkins、gitlab、docker容器开发配置代码高度集成，自动化部署管道。
3. 搭建ELK平台，分析app运行情况、出错概率、以及bug的修复情况。
4. 开发、维护、VCF基因数据处理分析管道流的代码开发，维护。调优。
5. Confluence编写平台搭建维护，代码迭代开发调优后证记文档。
6. 平台架构设计，比较各个数据库对VCF文件的搜索的速度，精选出最为合适的数据库作为相应 VCF文件的搜索方式。
7. 数据监控，增量数据监控管理。

**项目经验**

**项目名称：数据条线数据服务支撑**

责任描述；

1、高并发批处理计算代码优化，需求开发文档编写。

2、调优sql，解决大数据平台和数据仓库不同层级之间的数据对接。

3、统一调度工具Control-M配置大数据平台的任务调度。

4、ETL工具informatica配置大数据平台数据抽取。

**项目名称： 微基因项目优化体系**

责任描述；

1. 将oss康码基因数据（VCF）抽取到Hadoop平台的hdfs待预处理操作。
2. 转译VCF文件内容，提取相关字段数据，数据排序以及数据统计等相关数据清洗。
3. 针对不同阶段的代码优化，提高运行效率。
4. 利用hbase的builkload对hdfs上转译后的基因数据做hbase的高效上传。
5. 结合Kafka、Spark和HBASE特点，利用Azkaban调度工具，共同实现数据的海量存储，高效查询。

**项目名称：来因部落日志分析系统**

责任描述；

1. 部署开发调试ELK日志分析系统，将日志分为多个层，用户层、故障层、模块指标层等等。
2. 在elasticsearch中创建不同的索引存储不同的日志层。
3. 定时监控故障层的日志数据，反馈报告故障发生失败的原因，及时调整优化。
4. 利用Kibana生成界面对用户层的数据进行折线图，扇形图，柱状图等，图形胡的展示。
5. 针对不同模块指标，制定相应模块的调整方案。

**集成自动化部署**

责任描述；

1. 利用docker容器，gitlab，Jenkins配置代码集成自动化部署（CI/CD）。
2. 针对前端和后台的不同代码配置相应的dockerfile，docker-compose文件，将项目封装成docker容器，并实现打包部署上线。
3. 配置gitlab-webhook连接Jenkins实现项目的自动触发自动部署。
4. 编写pipline脚本，规范gitlab秘钥，各个项目全局凭据，提高项目部署效率。
5. 利用kong网关配置前端H5项目的访问域名，实现kong网关的灵活运用。

**语言能力**

英语 | 读写能力一般 | 听说能力一般

**自我评价**

1. 逻辑思维能力强，思路清楚，学习能力极强，对新技术有着强烈的好奇心
2. 对工作尽职尽责，乐于从事有挑战性的工作，对机器学习有极大兴趣
3. 乐于与用户以及同事和领导沟通，以便快速解决项目遇到的问题
4. 可以接受外派，能吃苦，抗压能力强
5. 有良好的数学学习能力，掌握MATLAB等数学软件