

李羿达

电话: 17600242671

邮箱: liyida@pku.edu.cn

地址: 北京市, 大兴工业开发区, 金苑路 24 号

Github: <https://github.com/idleyui/>



教育背景

北京大学 (北京), 软件与微电子学院, 软件工程硕士

2018 年 9 月-现在

北京交通大学 (北京), 软件学院, 软件工程学士

2014 年 9 月-2018 年 6 月

实习经历

阿里巴巴 (北京) 软件服务有限公司, 阿里云-基础软件部-操作系统-系统研发

2020 年 6 月-现在

实习开发工程师, 微处理器性能分析可视化

- 学习 Intel 关于微处理器性能自顶向下分析方法论与 Vtune 工具的使用
- 学习理解 PMU 原理, 分析相应事件的含义, 分析事件与性能指标的关系与计算
- 计划实现内存分析部分的功能

北京大学王选计算机研究所 (北京), 语言计算与互联网挖掘研究组

2019 年 6 月-2019 年 12 月

自然语言处理科研实习生, 多领域复述生成 & 可控复述生成

- 利用多个不同领域的复述数据集训练多任务模型, 提高模型在不同数据集上的表现

项目经验

搜索引擎 wiser-py

2019 年 5 月

使用 Python 实现「自制搜索引擎」书中的 wiser 搜索引擎

- 处理 30 万篇 Wikipedia 中文词条数据构建了一个简单的搜索引擎
- 通过分词和倒排列表构建倒排索引存入数据库, 使用 Golomb 算法对倒排列表压缩
- 通过对 query 分词, 在数据库中获取倒排索引, 并使用 TF-IDF 进行排序, 获取搜索结果

消息队列

2019 年 1 月

Java 课程大作业 (选题自 2018 年阿里中间件初赛)

- 利用多线程读写充分利用 CPU, 使用数据压缩减少读写的数据量
- 使用 ByteBuffer 进行高效的序列化, 减少 String 操作, 重用内存
- 利用 Linux 内存映射技术, 即 Java nio 包中的 MappedByteBuffer 加速文件读写

社交网络二度好友分析

2018 年 12 月

云计算课程大作业

- 利用三个 Hadoop 节点搭建分布式计算系统
- 通过 800 万条 Twitter 关注数据提取出 200 万条二度好友关系

专业技能

- 编程语言: 熟悉 Java, Python, 了解 C/C++, JavaScript, SQL, 了解基础的汇编
- 计算机基础: 熟悉链表, 树, 图, 排序, 搜索等基本数据结构及算法, 了解基本的操作系统, 网络, 数据库知识, 熟悉基本的 Linux 命令行使用
- 语言: 中文 母语; 英语 CET-6 514