



杨健

软件工程师 · 上海

☎ (+86) 18516191607 | ✉ ernestyj@outlook.com | 📷 Ernestyj

硕士 · 上海交通大学 | 学士 · 电子科技大学

个人简介

Current Site Reliability Engineer at start-up company Kasa. 7+ years experience specializing in the backend development, infrastructure automation, and computer hacking/security. Super nerd who loves Vim, Linux and OS X and enjoys to customize all of the development environment. Interested in devising a better problem-solving method for challenging tasks, and learning new technologies and tools if the need arises.

工作经历

PayPal · 全球风险控制

上海

数据工程师

2017年3月 - 2017年7月

- Mola 是百度内部广泛使用的 NoSQL 数据库，也为百度云提供全托管 NoSQL 数据库服务
- Built and deployed overall service infrastructure utilizing Docker container, CircleCI, and several AWS stack(Including EC2, ECS, Route 53, S3, CloudFront, RDS, ElastiCache, IAM), focusing on high-availability, fault tolerance, and auto-scaling.
- Developed an easy-to-use Payment module which connects to major PG(Payment Gateway) companies in Korea.

项目

- Implemented RESTful API server for car rental booking application(CARPLAT in Google Play).

百度 · 基础架构部

上海

研发工程师 · MOLA 分布式数据库

2017年3月 - 2017年7月

- Mola 是百度内部广泛使用的 NoSQL 数据库，也为百度云提供全托管 NoSQL 数据库服务
- 在 Proxyserver 中为百度云公网访问 MolaDB 服务添加标准访问日志，添加请求 key 记录，有效缩短后台故障排查时间
- 解决 Node 中分片恢复过程 Task 修改异步引发恢复错误问题，有效降低了 Mola 服务的故障率

百度 · 生态数据产品部

上海

实习数据工程师 · 精益-广告投放数据分析

2016年7月 - 2016年10月

- 精益对凤巢、Holmes 等广告数据进行监测与评估，建立漏斗模型(人群转化、流动、归因分析)分析广告营销活动效果
- 对上游广告源数据进行 ETL(Hive/SQL) 及数据拼接、编写 UDF(Spark)，优化 ETL 速率，并线上验证数据分析准确性
- 接入按查询/搜索关键词新建模型(Spark/Scala)的功能业务，广告主可以按关键词查看广告投放效果

荣石投资管理有限公司

上海

实习算法工程师 · 量化投资策略分析

2017年3月 - 2017年7月

- 基于 python 实现 Hurst 指数、SMA 双均线突破、分类预测择时策略
- 从 Wind 获取原始股票交易数据，并构建交易时间序列数据库(Cassandra)，作为公司内部数据源便于后续数据分析
- 提出基于最大信息熵的股票特征选取方法，对比主成分分析、奇异值分解等方法高出约 4.5% 的预测精确度
- 提出 SRA-Voting 混合二元分类模型，对比 SVM、随机森林、AdaBoost 高出约 3.125% 的预测精确度

技能

编程语言 C++, Python, Java, React

工具 Teradata, Excel, Tableau

英文 CET6/GRE/良好听说读写能力