

李逸凡

18811019056 | liyifan.steven616@gmail.com

教育经历

宾夕法尼亚大学

2021 年 9 月 - 2023 年 5 月

工学硕士：数据科学，工程与应用科学学院

美国 - 费城

- 选修课程：Statistics for Data Science, Computational Linguistics, Big Data Analytics, Algorithms & Computation

清华大学

2017 年 9 月 - 2021 年 6 月

管理学学士：信息管理与信息系统，经济与管理学院

中国 - 北京

辅修学位：数据科学与技术，软件学院

- 平均分：3.84/4.0
- 获 2019 年国家奖学金、2018 & 2020 年学业优秀奖学金等多项荣誉

实习经历

微软亚洲研究院 (MSRA)

2020 年 8 月 - 2021 年 3 月

软件分析组，实习研究员

中国 - 北京

- 参与一个云服务异常诊断算法的设计、优化与实现。作为论文第三作者，深度参与文献调研等论文撰写相关工作，并复现其中多篇相关工作进行对比实验。论文《HALO: Hierarchy-aware Fault Localization for Cloud Systems》已被 SIGKDD 2021 接收
- 与 Office 365 以及其他内部应用团队合作，基于该算法根据业务需求开发 Flask API 服务或 Kusto (Azure 数据资源管理器) 插件，帮助相关团队将算法整合进各自的报警处理流程。根据反馈，算法有效地缩短了业务团队工程师诊断并处理服务异常的时间。

字节跳动

2020 年 1 月 - 2020 年 8 月

大数据研发 / 数据科学实习生

中国 - 北京

- 在公司广告数据平台使用 Hive SQL, Spark 等技术深度参与数据仓库的建模、治理与建设，负责两张数据集市层 Hive 表及其数据链路的维护。主要对接抖音业务线的广告产品“Dou+”相关的数据需求，支持约 5 个产品新功能相关的数据分析，包括卡券、购物车和直播间推广等
- 通过时间序列数据挖掘方法，优化公司 A/B 测试平台的广告指标时序报警机制。将报警准确率从大约 60% 提升至 90%

研究经历

研究助理

2021 年 9 月 - 至今

导师：宾夕法尼亚大学沃顿商学院 Prof. Ryan Dew

美国 - 宾夕法尼亚

- TODO

课外活动经历

安永 NextWave 数据科学挑战赛

2019 年 4 月 - 2019 年 5 月

队长，地区冠军

中国 - 上海

- 带领团队在全球 2700 多名选手中获得中国地区决赛冠军
- 使用机器学习技术完成了居民活动路径终点的预测任务。为队伍贡献了 80% 的特征工程思路，使用 pandas 和 scikit-learn 独立完成了工作流框架的搭建，并与队友共同完成模型选择和参数调优
- 组织了一份清晰的展示报告，与队友一同阐述了整个探究过程的思路、对结果进行的分析，并提出实际应用的可能，给评委留下深刻印象并最终帮助我们获得冠军

CNKL-2-BibTeX

2019 年 2 月 - 至今

开源项目作者

- TODO

其他信息

- 技能：Python, Scala, Kotlin, Spark, SQL, Hive, L^AT_EX, C, Java, TypeScript, Node.js, 数据科学，机器学习