

宋辞

近三年行业机器学习，数据挖掘实践经验，法国工程师学校 **EFREI** 工程师学位（工程硕士学位），华南理工大学工学学士学位

工作经验

08.2018 – present

数据挖掘算法工程师 北京字节跳动科技有限公司 - 中国 北京

搭建投放渠道全链路风控系统

- + 基于网络，设备参数知识库等信息设计并上线了投放CPA渠道的实时风控系统。该风控系统能够实时反映渠道质量，辅助人工与自动化投放的决策
- + 建立了基于群组模型，分布检测等的离线作弊检测、质量打分系统。该离线系统能够过滤产品新增数目，辅助校准次留，抽取实时策略规则
- + 建立了渠道质量，作弊问题的解释系统。该系统能够辅助投放同事理解渠道问题来源，调整投放侧重，节省投放成本

11.2017 - 08.2018

高级机器学习工程师 北京玖富联银科技有限公司 - 中国 北京

- + 基于强化学习和遗传算法（进化策略）开发用于机器学习模型最优超参数组寻找的框架Finder
- + 开发智能运营系统。将机器学习算法（K-means, GBDT - XGBoost, ANN etc.）应用到运营工作中去，例如推广渠道质量（欺诈）监控，推荐模块（千人千面系统），卡券优化配置模块，用户运营（流失唤回）等
 - o 主要负责了推广渠道质量监控模块，主要通过分析设备行为数据（基于GrowingIO数据采集），聚类设备行为，行为标签价值量化等达到监控渠道质量
 - o 参与了推荐模块算法开发部分。开发了基于隐语义矩阵分解，行为数据轨迹的推荐模块
 - o 参与了卡券优化配置工作的特征工程（基于HIVE SQL）部分，对同事的建模工作提供支持
 - o 设计了基于强化学习的客户智能运营平台框架，并撰写了设计方案，开发计划

03.2016 – 11.2017

软件工程师 Atos Worldline - 中国 北京

12.2016 – 11.2017

- + 开发基于信贷业务全生命周期的通用机器学习平台Wisdom（基于Deeplearning4J）
- + 用户可以通过部署业务线达到从数据ETL（DataVec），特征抽象，模型建立，模型训练，到模型版本控制来实现将机器学习应用到贷前，贷中，贷后各个业务中去
- + 模型开发人员可以基于已经预定义的模型模板文件，通过drag-drop方式继续开发基于业务的模型
- + 部署基于Docker的持续集成平台，使交付工程实现一键部署

数据工程师 Atos Worldline - 法国 Bezons

03.2016 – 10.2016

- + 为英国铁路运营公司协会（ATOC）开发基于SpringXD以及GreenPlum的大数据解决方案
- + 平台能够采集分散在FTP服务器上的XML原始文本数据，通过数据流（SpringXD）的形式处理成结构化数据然后存储在分布式数据库（Greenplum）里

- + 对列车运行数据做数据挖掘，预测性分析。探索列车正晚点的内在规律，对列车的正晚点可能给出概率值。将以上结果部署在Tableau Server上
- + 系统包括数据ETL (SpringXD)，预测性分析 (SVM)，数据可视化 (Tableau) 等部分

04.2015 – 09.2015

软件工程师 **SARL EYELLE** – 法国 巴黎

- + Opencart框架的再开发工作。(Opencart是在线商店的CMS, 使用PHP MVC+Language开发)。实习期间开发了在线**实时订单打印技术**以及**在线支付模块**(基于源讯科技的支付解决方案)
- + 前端交互开发(主要从事jQuery和AngularJS的使用)

教育背景

2014 – 2016

硕士学位 – EFREI - 法国 巴黎

- + 专业: System Information Architecture
- + 学习内容摘要: 信息系统建模, AI, data warehouse, big data series

2010 – 2014

工学学士 – 华南理工大学 – 中国 广州

- + 专业: 信息工程
- + 学习内容摘要: 通信技术, 网络技术, 高等编程

技能

编程技能

PYTHON, JAVA, C, RUBY, SQL, NOSQL

数据处理

PANDAS, HIVE SQL, PYSPARK, MATLAB, R

硬件技术技能

LABVIEW, FPGA, DSP

语言

中文(普通话) 母语
英文 流利, 托业: 900/990(2015.4)
法语 常用交流水平

项目经验

强化学习实践

基于强化学习的 Sarsa 算法, A2C 算法实现的解决魔方问题 ([AlphaCube](#))

- + 抽象用于3阶魔方的状态空间, 动作空间, 以及奖励办法。
- + 使用Sarsa框架(算法)生成简单状态下的Q-Table的状态值。
- + 基于生成的Q-Table生成A2C算法的训练数据(两份数据分别用于训练Actor和Critic)
- + 使用A2C算法训练复杂状态空间下的魔方。
Sarsa, Policy Gradient(A2C), Reinforcement Learning

并行快速计算

使用 GPGPU 的并行计算能力加速计算。团队工作, 获学校 (EFREI) Enterprise day 推荐

- + 开发了一个支持远程部署在GPGPU上的平台
- + 提供基于OpenCL的接口, 使大型数据计算程序可以做到“即插即用”
OpenCL, JAVA, Android technologies

异构网络融合

基于 SIP 技术提供 WLAN 下的 VoIP 技术和移动网络通话服务的融合实验

SIP, Wlan, 3G, Android. 本科毕业设计项目

网页自动打印

网页在线来单实时打印技术, 无人值守状态下的订单来单打印系统

Ajax, PHP, MySQL, JAVA

更多信息

请登录 <https://cinqs.me/about/> 查看更多最新动态