

赵翰夫

+1 - 814-954-0729 | zhao@hanfu.us

+86-1821-000-5667 | hanfu.us/resume

懂数据的包工头, 渴望一个数据科学/机器学习岗位

教育

哥伦比亚大学, 土木工程硕士	2014
宾夕法尼亚州立大学, 生物工程学士, 环境工程辅修	2012

独立项目

KDD2019 CTR预测模型复现	2019
-------------------	------

- 复现KDD2019 DeepGBM神经网络与决策树结合的深度学习模型.
- 使用 **PyTorch** 搭建因子分解机处理稀疏数据, 获得嵌入向量.
- 使用 **XGBoost** 搭建梯度提升树处理稠密数据, 获得高维特征交叉.
- 提炼提升树模型的神经网络表征, 加入之前的嵌入向量, 一起进行端到端的离线学习和新数据的在线学习.
- 结合特征预处理, 超参数优化等技巧, 逼近论文结果.

政府人事数据平台产品	2019
------------	------

- 使用 **Scrapy** 等采集框架处理网站和新闻流, 并清洗数据归纳出人物履历, 覆盖90%以上的副部级及以上官员.
- 以 **Neo4j** 图数据库突出人物之间的关系, 结合 **Flask** 和 **Dash** 制作成云端数据库产品, 使用 **Docker** 部署.
- 因其健全的数据和出色的时效性, 数据库的付费订阅深受智库和投资机构青睐.

数据科学与机器学习的培训	2018 - Present
--------------	----------------

- 为机构内部及友好单位进行 **R**, **Python**, 应用统计等数据培训.
- 以微积分, 线性代数, 概率论等基础内容为重点, 编写机器学习系列教程.

工作经历

中国城市规划设计研究院	2018 - Present
-------------	----------------

- 使用 **R** 和 **MongoDB** 搭建地理数据可视化平台, 直观展示城市规模的共享单车运营, 并以时空数据模型支持管理决策
- 参与北京城市副中心水资源与水环境总体规划; 武汉, 株洲等海绵城市专项试点的规划.

美国费城市水务局	2014 - 2018
----------	-------------

- 为15亿美金的暴雨基础建设更新项目提供技术咨询和项目管理, 帮助市政府顺利通过联邦环境署的5年阶段验收.
- 搭建雨水检测数据平台, 实现数据的自动收集, 处理, 备份, 输出表现评分以供决策.
- 项目管理首次使用新技术的试点工程实施, 协调五所合作高校的联邦环境署科研项目.