

# 顾冰竞

数据科学家 (风控, 反欺诈)

地址  
上海中国

电话  
+86(0)15901802846

邮箱  
gubingjing@hotmail.com

网站 & Git  
www.gubingjing.com  
github.com/sangaj



语言  
Chinese ★★★★★  
English ★★★★★

## 工作经历

03/19 - 03/20 营销风险数据科学家 OYO

利用业务经验, 离群值挖掘, 复杂网络, 生存分析等发掘及设计可配置的营销反欺诈特征、策略及模型;  
爬取网络公开数据建立基础字典表, 支持欺诈模型特征研究等;  
使用机器学习算法 (逻辑回归、随机森林、GBDT 等) 识别注册黑产用户;  
使用阿里云相关工具 MC,ADB,DATAV 等及 Hive 处理营销数据的 ETL, 可视化等, 设计搭建营销实时及定时数据的画像特征库并对关键数据做统计分析, 报表展示及可视化;

11/17 - 03/19 风险数据科学家 众安保险

利用复杂网络、社群研究、语义分析、密度聚类探索通讯录、位置信息、用户画像等要素, 结合统计分析设计反欺诈策略及模型;  
爬取网络公开数据建立基础字典表, 支持模型特征研究及合作方数据定价、验真等;  
使用机器学习算法 (逻辑回归、随机森林、GBDT 等) 进行贷前决策信用风险建模分析;  
使用并行算法提升数据清洗、模型验证的速度;

02/16 - 08/16 数据科学家 皇家史基甫机场

用 Spark,Python 和 R 处理清洗部署存储在 Impala 上的数据, 并为 Hackathon 和 Kaggle 的数据比赛准备数据;  
设计研究方法和区域来提高控制跑道附近区域飞禽数的效率, 用 R,Python 基于惩罚回归, 支持向量机, 神经网络等机器学习算法寻找最佳预测模型并可视化结果;  
利用 H2O 在 Cluster 上并行运算 R 和 Python 的各类预测模型算法;

10/11 - 08/14 数据分析师 群邑中国

创建, 优化收视率日报, 周报, 月报的数据库, 使更新效率提升 50%;  
用 R 和 weka 基于线性回归和神经网络创建黄金档卫视电视剧收视预测模型;  
用 C 和 Excel 宏编写数据清洗模块以提高工作效率;  
基于频道, 市场, 客群, 客户等分析收视率, 媒体市场, 广告花费等数据为决策提供支持;

04/10 - 10/11 大堂经理 中国农业银行

收集客户和市场信息, 识别、分类客户, 依据客户需求提供定制化理财建议及金融产品;

## 项目经验

09/15 - 01/16 婴儿啾呀模式分析 莱顿大学

清洗 Phonbank 数据库的婴儿啾呀数据, 通过纵向数据分析研究各语言间的差别及演化, 为莱顿大学语言学博士的相关论文提供数据咨询, 论文已发表于 Cascadilla Press;

## 教育背景

09/14 - 05/17 统计科学 硕士 莱顿大学

09/06 - 06/10 应用数学 学士 武汉大学

## 技能证书

数据科学 约翰霍普金斯大学/ H2O Coursera 认证证书

Practical Machine Learning on H2O;R Programming; Regression Models;Exploratory Data Analysis;Statistical Inference 等;

反洗钱师 ACAMS ACAMS 认证证书

AML Foundations;