

# 个人简历

- 姓名：朱笑 性别：女
- 年龄：26 岁 工作经验：3 年
- 学历：硕士研究生 联系方式：17724705516
- 电子邮件：[karenmovingon@163.com](mailto:karenmovingon@163.com)



## 教育经历

2014.08——2015.09 香港城市大学——公共政策专业（海外留学）

主修课程：中国经济政策研究、中国外交政策等

2010.09——2014.06 嘉兴学院——工商管理专业（统招本科）

主修课程：经济学、财务管理学、统计学等

证书：证券从业资格证、基金从业资格证、会计从业资格证、大学英语 6 级、雅思 6.5 分

## 职业技能

- 熟练掌握 Python 核心编程方法，熟悉基本数据结构
- 熟练掌握 SQL 语言和数据库表体系，掌握关系型数据库 mysql、非关系型数据库 mongodb 和内存数据库 redis 的日常操作
- 熟练掌握 Numpy、Scipy、Matplotlib、Seaborn 等数据科学分析库的使用，能独立进行数据清洗、处理、分析与可视化操作
- 掌握 jieba 和 wordcloud 的使用，可以根据需求进行文本信息处理分析
- 熟悉 excel 及 VBA，可以对各项数据做各个维度的对比分析
- 熟悉朴素贝叶斯、决策树、线性回归、逻辑回归、聚类算法、PCA、推荐系统算法
- 熟悉爬虫基本库的使用，如 requests, urllib, bs4, xpath, selenium, 正则表达式等，能够独立进行爬虫项目的开发
- 熟悉 Scrapy 爬虫框架及网页抓取原理，可以高效地进行抓取海量数据、解析处理和入库等日常工作
- 熟悉 Linux 操作系统及 git 版本控制系统，熟悉掌握常用命令的操作

## 工作经历

深圳市鹏瑞信息科技有限公司

工作时间：2016 年 10 月-2018 年 09 月

在职职位：数据分析师

- 工作职责：
1. 收集、处理用户行为数据，挖掘用户行为特征，为产品、运营提供参考依据
  2. 关注数据变化，及时发现其背后的原因，整理成相关报告，提出业务发展及改进的建议
  3. 负责数据库数据的收集和整理，根据需求梳理数据，并用合理的方式分析和展示数据
  4. 熟悉各业务规则，与各个部门进行有效的沟通和协作，形成各类业务分析报表，配合其他部门完成数据统计和分析的工作
  5. 结合公司现状，定期分析产品及运营的痛点并提供有价值的解决方案

## 香港汇泽证券有限公司

工作时间：2015 年 09 月-2016 年 10 月

在职职位：数据分析师

工作职责：1. 根据公司整体经营发展需要和风控管理目标，开展分析及预测工作  
2. 建设并完善风险预警监控体系，管理日常报表体系，撰写专题分析报告  
3. 参与投资数据分析项目，协助分析主管进行数据提取、清洗、整合及后期的分析工  
4. 每日监控业务核心数据，对异常波动情况进行问题定位和深入分析

## 项目经验

### 【股票投资组合优化策略分析】

**项目描述：**爬取财经网站的股票数据，如股票的开票价、收盘价、成交量等等，将爬取的数据构建股票投资组合，根据金融理论里面的均值-方差理论分析并优化投资组合策略

**项目职责：**

1. 根据投资经理的要求，对标的股票的信息进行爬取，从东方财富、网易财经、新浪等财经网站爬取所需信息
2. 分析目标网站的 URL 和 response 内容，针对有反爬虫机制的网站，使用伪造请求头、设置 IP 代理等手段，进行 requests 请求，获取完整的响应信息
3. 使用 bs4、Xpath 对获取的响应信息解析，利用正则表达式提取有效信息，构建成为 json 数据
4. 将提取后的数据保存为 CSV 文件或存入 MongoDB 数据库，设置用户权限，方便统一管理
5. 利用 NumPy, Pandas 等数据科学库进行数据清洗，统计其对数收益率、方差和相关系数，用 Matplotlib 进行数据的可视化操作
6. 初始化股票资产组合里面各个标的的权重，然后通过蒙特卡洛模拟，产生大量随机的权重向量，并记录随机组合的预期收益率和方差，绘制其分布图
7. 优化投资组合，构建统计函数，统计夏普比率最大和方差最小的情况下，投资组合的整体表现，给投资经理提供策略建议

### 【信贷公司用户信用评分模型】

**项目描述：**根据公司风控要求，使用客户的信息资料及历史信用记录，利用机器学习里面的 Logistic 回归模型构建信用评分模型，预测客户的还款能力以及是否存在违约风险

**项目职责：**

1. 收集的数据来源于银行内部信息、申请表信息、征信机构信息，进行数据预处理，包括数据清洗，缺失值处理，异常值处理等，并且对数据集进行拆分，分为训练集和测试集
2. EDA 分析，获取总样本的大概情况，粗选预测变量，分析候选预测变量统计特征和违约率的分布
3. 变量分箱，基于特征工程细筛变量，利用 PCA 筛选出对违约状态影响最显著的一些指标
4. 建立和开发模型，先将变量进行 WOE 转化，选择 Logistic 回归模型进行分析
5. 模型验证和评估，主要是评估模型的区分能力、预测能力和稳定性，用 ROC 曲线和 AUC 来评估模型的拟合能力，并形成模型评估报告
6. 信用评分，根据逻辑回归的系数和 WOE 等确定信用评分的方法，选取评分卡参数，将 Logistic 模型转换为标准评分的形式

## 【猎头行业及职位招聘需求趋势分析】

**项目描述：**利用 Scrapy 框架爬取智联，拉勾等招聘网站的猎头职位信息，进行数据过滤、加工，分析猎头行业招聘需求和趋势，精准把握企业 HR 用人喜好，为企业设计招聘猎头的标准提供参考

**项目职责：**

1. 分析智联和拉勾招聘网站的 URL 结构和响应的内容，创建 Scrapy 项目和 Spider，在 Items 里面编写想要抓取的目标，如城市，薪资，工作年限等字段
2. 在 Downloader Middlewares 里面利用 selenium 模拟浏览器发送网页请求，分页爬取，包括每个职位对应新的 URL 里面的职位信息或职位描述
3. 使用 spider 里面的 parse 方法解析爬取的信息，获取所有正则过滤后的信息
4. 在 pipelines 里面设计需要存储的内容，将过滤后的信息保存到 MySQL 数据库中
5. 提取存入到 MySQL 数据库的数据，进行数据清洗，如将单条数据缺失值超过 3 个的删除等
6. 统计不同城市的猎头岗位和薪资的分布，综合对比分析教育程度及工作年限对薪资水平的影响程度，用 Matplotlib 进行数据图形化输出操作
7. 对提取的职位描述信息进行文本分析，利用 jieba 和 wordcloud 中文分词库，用 TF-IDF 量化出词频，分析出猎头岗位所具备的技能特征，总结出企业招人的喜好

## 【电子商务网站的用户行为分析及服务推荐】

**项目描述：**研究一家法律咨询企业网站的用户访问时间、访问内容及访问次数等分析主题，深入了解用户对访问网站的行为和目的以及关心的内容，采用协同过滤算法对用户进行服务页面的推荐

**项目职责：**

1. 从 MySQL 数据库中导入用户访问日志里面的具体数据，根据用户浏览的网页信息进行分类
2. 对数据的进行探索性分析，对原始数据中的网页类型、点击次数和网页排名等各个维度进行分布分析，获得其内在的规律
3. 从探索性分析的过程中发现与分析目标无关的数据，进行数据的清洗
4. 进行数据的转换，先识别翻页的网址，然后对翻页的网址进行还原，然后针对每个用户访问的页面进行去重操作，然后将数据划分为训练集和测试集
5. 选择 ItemCF 算法进行个性化推荐，通过 Jaccard 系数计算相似度，然后根据物品的相似度和用户的历史行为给用户生成推荐列表
6. 进行模型的评估，通过精准率（precision）和召回率（recall）去度量模型的准确率

## 自我评价

- ✧ 对数字敏感，有较强的数据洞察力，学习能力强
- ✧ 工作态度认真负责、严谨，具有团队协作精神
- ✧ 具有良好的逻辑分析能力、沟通能力
- ✧ 英语听说读写能力较强，通过英语六级，雅思 6.5 分
- ✧ 个人微信：A-smileL0L                      QQ: 375432567@qq.com