

# 杨睿

✉ yangruipis@163.com  
☎ 188-1823-1378

📄 github.com/Yangruipis  
📖 cnblogs.com/yangruiGB2312

## 🎓 教育经历

### 上海对外经贸大学

2012.09 – 2019.03

研究生 (保研)

数量经济学

本科

统计学 (商务统计方向)

GPA:3.31, 专业排名:2/54

## 🏆 荣誉和奖项

“华为杯”全国研究生数学建模大赛	二等奖	2017.09
中国大数据创新行动大赛	一等奖 (奖金 5 万元)	2016.12
美国大学生数学建模竞赛 (ICM)	二等奖	2016.02
上海市优秀毕业生、校十佳学术标兵、爱建企业奖学金一等奖 (全校 Top10)		

## 💼 工作经历

### 上海通金投资有限公司 (通联数据旗下私募), 量化研究员

2017.04 – 2017.10

- 策略复现与效果评价 (python+R), 股指期货跨期套利策略开发调试 + 实盘运作 (年化收益 12%)
- ctp 接口 Python 版本再封装与测试, 大大简化了策略模拟盘和实盘操作 📄 ctp-python

### Kantar Media CIC (中国领先的网络口碑咨询公司), 数据挖掘

2016.03 – 2016.10

- 数据采集: Facebook 爬取 (python), In、Nice、Lofter 爬取 (WireShark+python)
- 分词器开发 (C#+python, Jieba 内核), 对接搜狗细胞词库, 分词后直接绘制想要的词云图。
- 文本挖掘软件开发 (两个月, C#), 基于公司业务需求, 将微博文本根据树状规则, 匹配到对应的产品上, 并判断其情感度 (朴素贝叶斯 + 公司语料库)

## 📁 项目经历

### “东证期货杯”全国大学生统计建模大赛, 信贷违约模型

2018.01 – 2018.02

- 负责利用 python 实现完整机器学习流程, 包括数据入库、特征工程、模型选择 (LR,NB,CART,RF,GBDT)、模型与融合 (stacking)
- 难点 1. 数据不平衡 (SMOTE+Tomek); 难点 2. 需要降低授信成本 (引入 Cost-Sensitive Matrix)

### 个人机器学习库开发

2017.12 – 2018.01

- 个人项目 📄 simpleML, 开发语言: python
- 仅依赖 numpy 实现了常用的特征工程、模型评价、分类与聚类算法, 且接口统一, 代码工整

### 大学生挑战杯, 商科统计软件 VISAP 开发

2015.01 – 2015.06

- 团队项目 📄 VISAP, 开发语言: C#
- 负责统计功能 (空间作图、方差分析、回归分析、时间序列等), 抽样算法开发 (ARS, ARMS)
- 负责爬虫对接, 包括: 国家统计局宏观数据、财经股票数据等一键获取

## 🔧 技能 & 自我描述

- python 两年开发经验, 熟练使用 sklearn, pandas 等库, 熟悉 XGBoost, Keras, 面向对象编程
- C# 两年开发经验, 熟悉窗体程序开发, 了解多线程编程
- 熟练掌握 MySQL, Hive 的查询和优化、R 语言数据可视化, 曾用 Java 实现基本数据结构与算法, 了解 Spark、Docker、TensorFlow, 一年半 Linux 使用经验, Emacs 使用者。
- 数理统计基础扎实, 熟悉计量经济, 多元统计, 时间序列, 贝叶斯统计等方面的知识。
- 熟练掌握机器学习相关算法及原理 (SVM, LR, DT, RF, NB 等), 了解深度学和推荐系统基本算法 (RNN, LSTM, SVD, ICF 等)

热爱数据与编程, 有大量统计建模经验, 对数据敏感, 擅长用数据分析解决业务问题。乐于挑战, 追求新技术。英语读写流畅 (CET6:541), 爱好有摄影 (校摄影大赛二等奖)、骑行 (270km/3 天) 等。