**杨睿(Ray)**

籍贯： 江苏南京

TEL: 188-1823-1378 E-mail: *yangruipis@163.com*

博客园：*http://www.cnblogs.com/yangruiGB2312/*

**教育经历 EDUCATION**

**2012/09/01—至今** **上海对外经贸大学 统计与信息学院 统计学**

* 专业成绩前2/56，平均绩点：3.31
* 校级优秀团员、校优秀学生、保研名额
* 2次校一等奖学金、3次校二等奖学金、1次校三等奖学金
* 爱建企业特种奖学金 一等奖（全校top10）

**获奖情况 AWARDS AND HONORS**

**2016/02** **美国大学生数学建模竞赛交叉学科竞赛(ICM)二等奖**

* 前后参加4次大型建模竞赛均有获奖，建模经验丰富

**2014/07—2014/09 暑期社会实践“国内外医患关系对比调查研究 ”**

* 上海市知行杯二等奖，校一等奖

**2014/10 校级 年度十佳学术标兵**

**2016/3 上海市 优秀毕业生**

**项目经历 Program**

**2016/01—2016/04 挑战杯项目：贝叶斯统计软件开发（基本完成）**

* 语言：C#，平台Visual Studio 2013，环境：.net4.5
* 负责所有基本统计功能搭建，精确计算、朴素贝叶斯、空间作图功能的开发，以及数据获取功能（动态爬虫）

**2014/10—至今** **上海市大学生创新活动计划项目：空间计量模型研究**

* 前沿学科，在核心期刊：旅游科学发表论文

**实习经历 Program**

**2015/06—2015/12 PayPal 人力资源分析部门**

* 大量数据对接PPT的自动化，人力资源实时配置

**2014/10—至今 Kantar Media CIC 技术部门+分析部门**

* 数据采集：独立完成Facebook数据爬取(python, C#)，为公司每笔项目每年省下约四万元（可展示）；抓取手机APP端数据（小红书, In, Nice, Lofter）；其他静态、动态网页抓取。
* 数据清洗与数据处理：熟练对接公司数据库，通过朴素贝叶斯等算法对海量电商数据去除非自然人贴； 开发速度极快的分词器(Jieba内核)，同时很好的对接搜狗细胞词库。（可展示）
* 数据分析：
  + A 文本挖掘：独立完成文本挖掘软件开发（开发周期两个月， C#，速度优化后远快于公司现有系统，在公司内部被广泛使用）（可展示）
  + B 微博低质量粉识别器（python, SVM算法）（可展示）
  + C 情感分析：朴素贝叶斯算法对帖子情感度进行判别，比传统的关键词命中法更加准确
  + D 微博账号影响力的评价模型建立：主成分分析，统计建模。
* 数据可视化：

数据报告撰写经验，为公司制作词云工具(python)，相关可视化作品(R)

**技能 SKILLS**

* **软件技能** 熟练python做爬虫，数据处理与科学计算，熟练掌握 C#做文本分析，机器学习算法与winform界面，熟练掌握R语言数据可视化技巧，熟练掌握stata做高级计量分析。

掌握SQL server, Matlab,等软件的使用

* **个人技术博客：**https://www.cnblogs.com/yangruiGB2312/
* **分析技能** 非常精通 统计分析方法，如回归分析，时间序列，多元分析等等

热爱数据科学，喜欢研究数据挖掘/机器学习知识，热爱贝叶斯理论

* **其他** 英语综合能力：CET-6:**541**