

宋 晓

上海市浦东新区

邮箱: malcolm@xsong.ltd ◇ 个人网站: <https://xsong.ltd/>

教育背景

华东师范大学,社会发展学院, 社会学, 上海 (均绩: 3.60 | 前 10%) 2016 年-2020 年

相关课程: R 语言数据分析、算法与程序设计基础 (Python)、回归分析与 Stata 应用、SPSS 应用

北京大学-密歇根大学学院, 北京, 暑期课程 2018 年 8 月

相关课程: 因果推论方法的研究设计和敏感性分析

中山大学, 社会学与人类学学院, 广州, 学期交换 2018 年 2 月-2018 年 7 月

相关课程: 高级统计与论文写作、社会模拟与 NetLogo 应用

学术研究

机器学习在社会科学定量研究中的应用 2020 年学士学位论文

- 独立作者。指出基于广义线性的传统定量分析方法具有滥用 P 值、不稳定性等问题, 造成了研究结果的不可靠性。机器学习方法具有更强大的预测能力与可解释性, 能够取代现有的统计方法得到广泛应用。
- 和机器学习算法如何应用于社会科学研究。

土地流转的福利效应与社会不平等: 来自 CFPS 的经验证据 2018 年-2019 年国家大学生创新训练项目

- 独立作者, 使用中国家庭追踪调查 (CFPS) 数据, 通过 Stata 和 R 进行数据清理和计量经济分析。使用无条件分位数回归和固定效应模型估计土地流转行为的福利效应和对社会不平等的影响。使用 ggplot2 软件包进行地理信息可视化。输出图表并最终形成约 12000 字的研究论文。

实习经历

中南财经政法大学 数据分析助理 2020 年 2 月-2020 年 3 月

- 使用 Python 的 pandas、numpy 和 scikit-learn 库进行数据清理、建模。使用决策树和随机森林算法建立信用卡评分模型。通过交叉验证 (Cross-Validation) 的方法训练模型, 对互联网法律文本数据进行分类。

艾瑞咨询 数据分析实习 2019 年 7 月-2019 年 9 月

- 使用 SQL 对业务数据进行查询和清理
- 使用 R、Python 对文本数据进行可视化

获奖与资助

优秀奖 第二届全国高校数据驱动创新研究大赛 2019 年 北京

特等奖 第十二届本科生社会科学论坛 2018 年 上海

作品: “颜值”真的重要吗? ——相貌因素对初婚风险的影响

二等奖 第二十六届“大夏杯”大学生课外学术科技作品竞赛 2018 年 华东师范大学

作品: 中国女性的婚姻幸福感——性别观念与家务劳动不平等的双重冲击

二等奖 优秀学生奖学金 2018 年 华东师范大学

二等奖 优秀学生奖学金 2017 年 华东师范大学

语言和计算机能力

语言能力 大学英语四级 614, 大学英语六级 549, 托福 103(阅读 29 听力 27 口语 21 写作 26), 中文 (母语)
GRE 321(语文 154 数学 167 作文 3.5)

软件技能 R, Python, SQL(包括但不限于 MySQL, SQLite, SQL Server), SPSS, Git,
L^AT_EX, Stata, MS Access, MS Office, Adobe PhotoShop