

唐艺文

(336) 918-3479 | tonytang0731@outlook.com

3600 Chestnut St., Philadelphia, PA 19104 | tonytang731.github.io/github-pages-react-app/

教育背景

宾夕法尼亚大学，工程学院 (3.73/4.00)

费城，宾夕法尼亚州

理工科研究生 – 数据科学

May 2021

维克森林大学 (3.74/4.00)

温斯顿-塞勒姆，北卡罗来纳州

荣誉理科学士 – 数学

May 2019

荣誉文科学士 – 经济 (1st / 124 于经济系毕业生考试)

辅修 – 计算机科学

- STARR 研究旅行补助金 & 暑期研究奖学金

技能资质

编程语言: Python, Java, PHP

数据库: MySQL, MongoDB

数据分析: R, Stata

前端语言及工具: JavaScript, HTML/CSS

数据可视化: Tableau

微软办公软件: Excel, PowerPoint

工作经历

Two Six Capital

旧金山，加利福尼亚州

数据科学实习生

06/2020 – 至今

- 为开源包 twosixtools 与 Lego 设计及编写方程与测试
- 建造数据管道以提供产品化的数据分析及自动数据可视化服务

Cognitive Computation Group, University of Pennsylvania

费城，宾夕法尼亚州

研究开发员

01/2020 – 至今

- 实现 8 个由 CogComp NLP 支持的文本注释服务的交互式 Demo
- 发布了 2 个高级注释工具并辅以完善的 JSON 框架的 API (Wikifier, 名称实体相似性)

微信事业群，腾讯

深圳，广东

数据分析实习生

06/2019 – 08/2019

- 与金融科技团队进行财务报告分析；从数据库和外部来源检索及提取信息
- 主持了有关机器学习主题的讲座，并编写了时间序列分析模板

普华永道深圳事务所

深圳，广东

审计实习生

07/2017 – 08/2017

- 收集并分析了平安保险子公司 30 多个报告中的数据
- 审计 A + H 股票报告中的几个板块，并与普华永道上海的 Fintech 同事进行了合作

项目实践

Classifying Hate Speech Using a Two-layer Model

06/2018 – 05/2019

- 论文发表在学术期刊《统计与公共政策》上，在北卡罗来纳州的 AISC 会议上获得杰出表现奖
- 建立了 30 多个测试用例以确保文本清理程序的有效性
- 方法中纳入了多种机器学习算法和特征提取技术
- 设计基于词典的特征将预测准确度从 58.25% 提高到 66.75%

Photogram

09/2019 – 12/2019

- 使用 React 和 UIKit 复制了 Instagram 的主要功能和 UI
- 应用 Github 创建项目管理分支策略并跟踪问题/用户故事
- 使用 Jest 进行单元测试，通过 Selenium 进行功能测试以及运用 Travis CI 进行持续集成

Wharton Analytics Fellow, 沃顿商学院

09/2019 - 12/2019

- 与 BB&T 的人力资源部门直接合作，并通过文本分析得出员工每周反馈与其表现的相关性