



萧尧

(+86) 186-2182-3612 | yx2436@nyu.edu | charlie-xiao.github.io |  Charlie-XIAO |  in yao-xiao-200073244

🎓 教育背景

上海纽约大学 | 本科在读 | 荣誉数学 | 计算机科学 | **GPA: 3.89/4.00** 2020.09 – 至今

- **荣誉数学 GPA: 4.00/4.00**, 主要课程包括: 荣誉 {微积分、线性代数、数学分析、常微分方程、概率论、数值分析、抽象代数}、偏微分方程、建模与模拟、复分析*、随机分析*、概率论极限定理。
- **计算机科学 GPA: 3.97/4.00**, 主要课程包括: 离散数学、数据结构、计算机架构、算法、操作系统、开源软件开发、随机化算法、机器学习*、计算机网络。

🏢 工作经历

纽约大学柯朗数学科学研究所 / 上海纽约大学 | 助教 2021, 2023

- 于纽约大学柯朗数学科学研究所计算机部门, 担任 2023 年春季学期操作系统课程助教。
- 于上海纽约大学学术资源中心, 担任 2021 年和 2023 年秋季学期微积分课程助教。

复旦大学数据智能与社会计算实验室 | 实验室研究助理 2023.05 – 2023.08

- 构造近 30 万条指令数据, 并通过指令微调的方式训练在中国法律领域专精的领域大模型 DISC-LawLLM。
- 构建用于 DISC-LawLLM 检索增强框架及相应数据库, 包含 800 余部法律和 2.4 万余道法考试题。
- 构造主观和客观两方面的法律模型评测框架, 并测得 DISC-LawLLM 在法律领域性能超越基座模型 23%, 超越 GPT-3.5 Turbo 9%。

🔬 研究经历

大图的高效可视化方法 | 上海纽约大学本科生科研基金项目 | *DURF* 2022.07 – 2022.08

- 实现了 DeepWalk、Node2Vec 等图嵌入算法与 PCA、t-SNE 等降维算法, 将它们结合进行大图可视化。
- 提出了基于拉普拉斯特征变换与最短路径算法的图嵌入算法 SPLEE, 及根据图的特殊结构改进的降维算法 t-SGNE。
- 结合 SPLEE 与 t-SGNE, 可用于高达 300K 节点与 1M 条边的大图可视化, 可视化效果对比现有算法的组合提升约 10%, 计算时间小于 5 分钟 (现有算法组合耗时过长或内存溢出, 无法进行测量)。

📁 项目经历

Pandas / Scikit-learn | 开源贡献者 | 排名 #78/#50 2023.03 – 至今

- 在 pandas 中贡献了 24 处 (1592 行) 文档或代码的改进与修复。
- 在 scikit-learn 中贡献了 62 处 (4433 行) 文档或代码的改进与修复, 并被提名进入 Triage Team。

YouTube 页面定制插件 | 课程项目 | *CSCI-UA.0480.61* 2023.02

- 开发了一款火狐浏览器插件, 使用户能够切换 YouTube 界面的色彩主题, 并隐藏、重排、或定制多种组件。
- 撰写了该插件相关功能、使用方法、与贡献方法的文档, 并于火狐插件商店发布了该插件的 1.0 版本。

模拟经济体不平等化的过程 | 课程项目 | *MATH-UA.0395* 2022.12

- 通过经济体中的随机交易, 模拟了不同交易模型下经济体的不公平化过程, 并分析了不同的交易函数对模型的影响。
- 通过对经济体中终末财富分布的拟合, 得出现实社会中财富分布近似遵循 γ 分布或 β 系数分布。

模拟陀螺仪塔的性质 | 课程项目 | *MATH-UA.0395* 2022.10

- 用弹簧系统 (非刚体) 对陀螺仪进行建模, 并通过微分方程与欧拉方法对该模型的物理性质进行模拟。
- 模拟纵向叠放陀螺仪形成的陀螺仪塔, 并观察到中轴可弯折时, 陀螺仪塔具有与陀螺仪相同的陀螺进动性质。

简易 SQL 优化器 | 字节跳动第四届青训营 | 大数据方向 2022.07 – 2022.08

- 实现了简易的 SQL 解析器, 支持 Filter、GroupBy、Join 等多种算子, 并构建抽象语法树。
- 实现了谓词下推、列裁剪等基于规则的语法树优化算法, 以及 Join 重排等基于代价的语法树优化算法。

⚙️ 技术能力

- [1] 常用 Python 和 C 语言, 并熟悉 Java、JavaScript、PHP、C++、Julia、MATLAB 等其他多门编程语言。
- [2] 熟练使用 Git 与 GitHub 工作流。

🏆 荣誉奖项

- [1] 优胜奖 (Meritorious Winner), 2023 年美国大学生数学建模竞赛
- [2] 学年度院长嘉许名榜 (Dean's List of Academic Year), 2020 – 2021 学年, 2021 – 2022 学年, 2022 – 2023 学年

*研究生课程。