

刘兆东



Email: z14789@nyu.edu | Tel: (+86) 131-5629-6623

教育经历

上海纽约大学

上海, 中国

理学学士; 双主修: 计算机科学、数学, Cum Laude (荣誉毕业)

2022年9月-2026年5月

累计 GPA: 3.803/4.0 | 计算机专业 GPA: 3.946/4.0 | 数学专业 GPA: 3.818/4.0

- 海外学习经历: 纽约大学纽约校区, 2024年9月 - 2025年5月
- 相关课程: 多元微积分、荣誉线性代数、概率与统计、荣誉数值分析、离散数学、数学分析、常微分方程、随机过程、数据结构、算法、操作系统、数据库、机器学习、自然语言处理
- 荣誉: Dean's List 院长名单 (学年: 2022 - 2026)、Founders' Day Award (年度杰出学生荣誉奖)

卡耐基梅隆大学

匹兹堡, 宾夕法尼亚州, 美国

理学硕士; 计算数据科学 (Master of Computational Data Science)

2026年8月-2027年12月 (预计)

科研经历

研究助理, 上海纽约大学

2025年5月-2026年5月

多模态推荐系统的生成式检索

指导: 文弘毅教授, 计算机科学系、数据科学与工程系

- 开发了一个多模态框架模型, 用于预测亚马逊数据集上的用户行为模式。
- 设计并评估了模型架构, 实现了 VQ-VAE、RQ-VAE 和 OPQ 等量化方法, 以捕捉离散行为模式, 从而实现了 47.2% 的性能提升 (从 0.0339 提升至 0.0499)。
- 采用了受 BPE (字节对编码) 启发的分词技术, 从而提高了序列的连贯性和表示质量。

研究助理, 上海纽约大学

2024年5月-2024年8月

考虑交通信号灯时长的高效驾驶路线优化

指导: 陈志斌教授, 计算机科学系、数据科学与工程系

- 在 SUMO 平台上基于真实地图及相关约束条件, 构建了一个具备交通感知功能的驾驶路线优化模型
- 在真实交通信号数据集上构建的模拟模型, 与基线路由算法相比, 能够缩短出行时间。
- 针对电动汽车和燃油汽车的能耗特性应用相应的模型, 兼顾能源效率与行驶时间的平衡。

项目展示

微调预训练语言模型, 完成化学反应的逆合成高效预测

2024年9月-2024年12月

指导: 何河教授, 纽约大学数据科学中心, 纽约

- 利用 SMILES 标记法开发了用于化学反应逆合成的预测模型。
- 应用了大语言模型 (LlaSMol-Llama2) 对数据集进行重建和扩充, 使 Top-1 准确率提高了 26.2% (从 0.588 提升至 0.742), 同时减少了对实验室数据的依赖。
- 采用了精细的调优策略 (Encoder 参数冻结、MLP 适配器、LoRA) 以优化训练效率。

虚拟体育联赛经理数据库

2024年10月-2024年12月

指导: Salim Arfaoui 教授, 数据库导论课程, 纽约大学坦顿工程学院, 纽约

- 在 MySQL 中设计了包含 13 张数据表的关系型数据库, 支持多联赛 (橄榄球、篮球、足球) 的虚拟队伍运营, 涵盖用户认证、选秀调度、阵容管理与交易系统, 录入 390 余条基于现实世界的样本数据。
- 实现了符合 ACID 规范的事务管理、自动化触发器与存储过程, 通过索引优化查询性能, 并强制保证所有关系间的参照完整性。构建了排名算法、赛程编排与积分系统及免签机制。
- 以 Python 构建后端服务, 实现用户认证、数据查询与业务逻辑处理; 以 HTML/CSS 开发前端交互界面, 支持可视化阵容管理、实时积分展示及交易操作, 实现前后端与 MySQL 数据库的全链路集成。

多人协作钢琴应用

2025年3月-2025年5月

指导: Daniel Katz-Braunschweig 教授, 纽约大学坦顿工程学院, 纽约

- 采用 Java 全栈开发, 以 Java Swing/AWT 构建图形化钢琴键盘前端界面, 以 Java Socket 服务端处理多客户端连接, 实现实时音频同步、在线聊天及精确时间戳的录音回放功能。
- 基于 TCP 套接字网络与线程并发 (ExecutorService、ConcurrentHashMap) 实现服务端多客户端并发管理, 通过 Java Sound API 支持合成波形与真实钢琴采样等多种音色。
- 利用 Graphics2D 与 AffineTransform 实现动态节拍器动画, 将前端 GUI、服务端网络通信、文件 I/O 整合为一个完整的实时多媒体全栈系统。

工作经历

数据工程师实习, 麦当劳中国数据战略部, 上海

2025年12月-2026年4月

- 开发并验证了可扩展的数据库模式以及企业级的数据映射, 整理构建数据地图。
- 设计了一种基于菜单构成和用户历史交互的相似性推荐算法, 为活动策划、物资筹备提供参考。
- 开发 AI Agent 应用于数据检索, 显著提升了数据面板的分类和可视化效果。

领导力与活动

财务主管, PEER-Pal 社团, 上海纽约大学

2023年5月-2024年8月

- 负责管理学生主导的教育项目的财务运作及年度预算编制工作。
- 举办了与湖南省高中生的信件交流活动以及问题交换解答活动。

交响乐团长号声部首席, 上海纽约大学

2022年9月-2026年5月

- 领导长号声部, 协调排练事宜, 并负责乐器存放的后勤工作。

奖项

Mathematical Contest in Modeling, Honorable Mention 奖

2023年5月

- 对预处理后的大型数据集进行处理, 以剔除异常值并确保数据质量。
- 利用包括贝叶斯回归在内的多种方法构建了预测模型, 用于对区域市场数据进行分析。

技能及其他

- 编程语言: Python, Java, C, SQL, Stata, R
- 框架&工具: Git, PyTorch, Pandas, NumPy, LaTeX, Markdown
- 语言: 中文(母语), 英文(流畅)