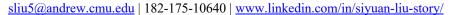
刘思源





个人介绍

机器学习工程师,拥有3年以上生产级AI系统经验

兼具 Agent 系统构建(主导构建 AIVC 自动化系统,提高公司生产效率)、ML 系统设计(构建基于机器学习模型的量化系统,超越纳斯达克 100 指数 15%的超额收益)与大数据处理能力(对 Spark/Hadoop 上的交易图挖掘算法进行优化,使处理速度翻倍)。擅长将前沿研究成果转化为产业价值,通过性能工程(将开源算法组件优化至连接子图分析加速 2 倍)与算法加速(TKDE 2025 发表的 k-hop 约束路径生成算法性能超越 SOTA 10 - 100 倍)。

教育经历

卡耐基梅隆大学, GPA: 3.85 / 4.0

匹兹堡, 宾夕法尼亚州, 美国

信息系统管理硕士(商业智能与数据分析方向)

08/2024 - 12/2025 (预计时间)

核心课程: 高级自然语言处理, 分布式系统管理, 机器学习

复旦大学, GPA: 3.58 / 4.0 数据科学与大数据技术学士

上海,中国

09/2019 - 06/2024

核心课程: 自然语言处理,人工智能,计算机可视化,数据库与实现,高级大数据处理

职业经历

多智能体研究系统, IOSG

05/2025 - 至今

- **主导设计开发端到端 AIVC 系统**,实现新闻周报、项目研究总结和领域研究报告的自动化生成。将分散于 Notion 和 Google Drive 的知识统一整合进**后端 RAG 知识库**,基于 FastAPI 搭建交互服务并部署于 AWS Cloud,实现公司过往知识沉淀的高效复用。
- **构建实时数据采集管道**:利用 RSS、Firecrawl、Grok 等多源数据接口抓取新闻、金融与社交媒体信息,结 合 LangGraph 多智能体框架实现检索增强和上下文推理,报告准确率达到生产级别。
- 引入 DL + HITL 专家反馈环路, 迭代优化生成结果; 采用 LoRA 对开源模型进行轻量级微调, 使 模型 F1 分数提升 20%, 报告冗余输出减少 80%, 新闻分类准确达到 99% 以上, 重要性判定能力达到 95% 以上。

量化分析实习生,上海双隆投资有限公司

06/2023 - 12/2023

- 设计并实现模块化因子择时回测框架,结合 IC/IR 动态调整因子权重,使投资团队验证策略的效率提升 40%
- 构建基于 LSTM 与 VAE 的模型,在不同市场状态下预测因子表现,相较基准模型提升碳期货预测准确率 20%。
- 识别并验证基于动量的投资策略,基于两年历史数据回测实现 15%超额收益,策略已被投资团队采纳部署。

数据工程实习生,0xScope

12/2022 - 05/2023

- 开发实时图挖掘算法,通过识别交易路径与图模式进行链上身份检测,使实体识别速度提升2倍、准确率提升10%,显著增强反洗钱智能流程。
- 在 AWS EMR 环境中重构并并行化基于 Spark 的 ETL 流程,处理 250GB+以太坊交易数据,使每日批处理延迟降低 10%,并配合 DA 团队满足下游任务的数据需求。

科研经历

研究助理 - 复旦大学(指导教授: 郑卫国、付彦伟)

项目 1: 用于可扩展子图分析的图算法优化

09/2023 - 06/2024

- 主导多个旨在提升大规模图分析算法可扩展性与效率的研究工作,从零实现 VEQ 与 CECI 子图匹配算法,并引入部分排序技术以加速候选节点枚举过程。
- 重构开源图计算库 LAGraph,引入顶点聚类与贪心剪枝策略,使连接子图检测性能提升2倍。
- 提出一种新颖的 k-hop 约束 s-t **路径图生成方法**,在真实图数据集上相较现有方法取得 10 至 100 倍加速。
- 相关成果支撑两篇论文发表于 TKDE 2025 与 SIGMOD 2023,并被某一线科技公司采纳落地。

项目 2: 基于增强技术的少样本学习方法

07/2021- 12/2022

- 提出一种全新的数据增强框架,以提高少样本学习在甲骨文字符识别任务中的准确率。通过图像配准与自由 形变(FFD)技术生成结构一致的训练样本变体。
- 将上述方法集成为统一模型 EASY,在多个基准测试中相较原始模型准确率提升 30%。
- 成果发表于 ACCV 2022,并被实验室广泛采纳用于其他少样本视觉任务。

学术论文

• Cai, Y., Liu, S., Zheng, W., Zhang, C. & Zhang, X. (2025). Generating k-hop-constrained s-t Path Graphs. *IEEE Transactions on Knowledge and Data Engineering, doi: 10.1109/TKDE.2025.3532318*TKDE 2025