

刘悦

电话: (86) 18605813850

邮箱: yliumh@connect.ust.hk

出生年月: 1999/07

个人博客: <http://yliuhz.github.io/blogs>, 知乎: 悦大



教育背景

香港科技大学	计算机 (方向: 多模态大语言模型)	研究型硕士	2023-2026
浙江大学	计算机 (方向: 图神经网络)	专业硕士	2020-2023
浙江大学	应用数学	本科	2016-2020

项目经历

多模态大语言模型的引用可靠性	2024-2025
<ul style="list-style-type: none">描述: 大模型对多跳问题的多模态引用容易产生幻觉, 降低用户对问答系统的可信度目标: 提升多跳问答系统的答案-多模态信息来源引用的一致性, 例如自回归生成的引用文献对模型输出的支持性职责: 1. 基于组合图像检索任务和多模态问答任务的实验, 发现干扰信息对后端引用的不利影响; 2. 提出基于问题分解的重排序方案, 有效筛除干扰信息; 3. 对 Self-Ask 等问题分解方案进行改进, 提出零样本问题分解方案; 4. 基于子问题对正确文献的检索召回率定义奖励函数; 5. 基于多模态 QA 数据集构建包含 37K 条问题分解样本的 DPO 训练集, 使用 Llama-Factory 框架进行微调; 6. QDRAG 在同模型下对多模态引用可靠性提升 18.7%, 检索召回率提升 9.3%, 综合性能超越商用 Gemini 模型。	
多模态信息检索和问答系统的研究综述	2023-2024
<ul style="list-style-type: none">描述: 对多模态检索增强系统的主要模块的现有方法进行调研调研内容: 1. 跨模态检索, 从早期的 CCA 统计模型、基于排序 loss 和难样本挖掘训练的轻量表征模型, 到图文表征预训练模型, 如 LXMERT, CLIP 等, 和提取多模态大模型隐藏层的表征方法; 2. 重排序方法, 分为 pointwise, pairwise, listwise, setwise 4 类; 3. 信息增强方法, 如提示词拼接、分块 attention、最近邻后处理方法等; 4. 引用方法, 如检索后处理、SFT、DPO 微调方法等。	
基于政务知识图谱的检索增强系统 (浪潮合作项目)	2024-2025
<ul style="list-style-type: none">描述: 使用大模型抽取政务网页数据中的实体和关系三元组; 构建基于知识图谱和 RAG 的问答系统职责: 1. 使用 7B 模型标注 1.5K 条三元组提取数据, SFT 微调 3B 模型; 2. 参考 GoG 定义 thinking-searching-generating 三元组检索框架, 与 ColBERT 对文本块检索合并组成多路检索框架; 3.使用 8B Qwen3-Ranker 对文本块重排序, 类比 ColBERT 对千问重排模型在长文本上的效率提升 50%; 4. 使用 vLLM 部署问答系统。	

实习

小鹏汽车 机器人中心	多模态大模型算法实习生	(2025/09-2025/12)
(M)LLM 后训练论文复现, 基于私有清洗、训练、评测框架。		
香港科技大学 计算机科学与工程学院	科研助理	(2023/04-2023/08)
负责调通及部署大语言模型预训练语料的自动化清洗工具, 包括段落哈希、去重、语种识别、困惑度分组等。		
杭州海康威视数字技术股份有限公司	AI 算法工程师	(2021/10-2022/05)

负责在海康私有数据上研究虚拟数据对 YOLOX 模型训练推理的影响，涉及小目标、打架行为等真实监控数据稀缺的场景。结果表明，私有数据对少样本目标检测性能有所提升。

论文

[1] Liu et al. (first author). (2025). Faithful in Steps: Improving Generalization and Citation in RAG via Query Decomposition. Submitted to AAAI.

[2] Liu et al. (first author). (2024). Progressive Graph Construction for Attributed Graph Clustering.

[3] Liu et al. (student first author). Reliable Community Search on Uncertain Graphs. ICDE 2022 (CCF-A).

[4] Liu et al. (co-first author). Prediction of Maximum Absorption Wavelength using Deep Neural Networks. JCIM 2022.

专业技能

- 使用 vLLM、SGLang 等推理框架部署在 A100, 4090 等 GPU 上部署最高 72B 模型的本地批量推理、本地 server 服务
- 调用商用大模型 API, 包括 Google Gemini、字节 Doubao 等
- 使用 Llama-Factory、verl 等训练框架对大模型进行 SFT 微调、DPO/GRPO 偏好微调
- 熟练应用 Python、C++、SQL 等编程语言, 了解 Python 多进程编程
- 通过雅思 (7.0), 四六级等考试, 熟练阅读英文文献, 撰写过若干英文论文

荣誉奖项

- 香港学业奖学金、光谷奖学金、浙江大学优秀毕业生、浙江大学优秀研究生、浙江大学三好学生

发明专利

一种异常节点检测方法、装置及电子设备 (CN119961823A)

本申请的实施例揭示了一种异常节点检测方法、装置及电子设备，方法包括：根据数据库中的用户交易记录，构建交易数据集；基于结构相似度和语义相似度获取与所述交易数据集相匹配的参考数据集；将交易数据集和参考数据集进行特征拼接，得到目标训练矩阵；根据目标训练矩阵对图神经网络进行残差学习，得到节点嵌入矩阵；基于上下文注意力对节点嵌入矩阵进行异常评分，得到节点的异常分类结果；本申请通过利用少量的正常样本，在推理阶段动态地学习和预测其他节点的异常程度，从而实现了高效的异常检测。

无社区基数下的社交网络社区发现方法及装置、电子设备 (CN116775942A)

本发明公开了一种未知社区基数下社交网络的社区发现方法。该方法首先构建编码器，使用任意无监督方法训练编码器，将输入社交网络映射到欧氏空间，获得顶点的表征向量；计算顶点对的相似度矩阵，并对相似度进行排序，取相似度较高的顶点构成正样本，取相似度较低的顶点作为负样本，并分别赋予伪标签，构成训练集；构建社区关系预测模型，使用二元交叉熵损失函数训练模型，预测顶点对的社区关系，同时优化超参数；根据模型预测结果构建结构图，将结构图输入刘温算法得到最终的社区发现结果。实验表明，在不提供社区基数的条件下，该方法能够较准确地探测社交网络中的社区结构。

一种基于深度神经网络的有机荧光小分子光学性质预测方法 (CN113380337A)

本发明提供了一种基于深度神经网络的有机荧光小分子光学性质预测方法。本方法通过建立新的有机荧光小分子数据库，采用分子描述符和分子指纹提取分子信息，输入多层神经网络和卷积神经网络进行深度学习训练获得有机荧光小

分子光学性质预测模型，将待预测有机荧光小分子及其实验溶剂的特征信息输入训练好的有机荧光小分子光学性质预测模型从而预测其光学性质。本发明方法能准确地预测有机荧光小分子的光学性质(平均相对误差小于 5%)，从而提高有机荧光小分子的开发效率。