



教育背景

上海交通大学大学，信息工程，硕士	2022.9 - 2025.3
<ul style="list-style-type: none">指导老师：谢伟迪，张娅（CMIC 未来媒体网络实验中心）致远荣誉奖学金 * 4、一等学业奖学金	
上海交通大学大学，计算机科学，本科	2018.9 - 2022.7

项目经历

大语言模型 Shanghai AI Lab

PMC-LLaMA Star 566 Cite 70

- 负责 PMC-LLaMA 的训练, 并在预训练后进行指令微调
- 以书籍为载体对 LLM 进行知识注入, 在多个医学问答数据集上超过了 ChatGPT
- 通过难样本筛选进行数据剪枝, 加速指令微调

MMedLM Star 145 Cite 7

- 将单语言模型 PMC-LLaMA 迁移到多语言

多模态大模型 上海交大 - CMIC Lab

PMC-CLIP Star 123 Cite 64

- 用医学文献的图文数据, 训练出第一个全场景的医疗多模态大模型
- 作为图文基础模型, 以强大的泛化能力, 在图文搜索, 图像分类, 视觉问答等下游任务上获得第一

PMC-VQA Star 158 Cite 90

- 模型从医学图像生成诊断报告, 实现影像学的开集视觉问答
- 以特征层重建为代理任务, 提升视觉编码器对图像表征的提取能力

高密度知识预训练 阿里巴巴 - 拍立淘

PatchDedu Under Going

- 根据单模态先验模型对词元压缩实现知识密度提升
- 将数据筛选从样本级优化到词元级, 实现更精细的数据治理
- 同时实现多模态大模型的训练与推理加速

拍立淘 阿里巴巴 - 拍立淘

图搜相似商品 Under Going

- 提供图搜索系统的商品视觉表征模型
- 挖掘商品标题和商品风格描述的领域差异, 学习共有语义提升表征模型

交大柠檬

- 微信小程序”交大柠檬”, 深受交大师生喜爱的生活社区。总注册 10 万, 平均日活 5k,
- 负责二手交易板块后端开发. 每日人均打开 5 次, 次均停留时长 81s, 累计成交金额 450 余万
- 跨模态检索功能正在开发, 将支持用户上传的商品信息的自动补全



工作经历

阿里淘天，深度学习实习	2024.3 - NOW
上海人工智能实验室，深度学习实习	2023.7 - 2024.3
微软亚洲研究院，机器学习实习	2022.7 - 2022.10
交大柠檬，后端 & 机器学习	2022.3 - NOW
美国大学生数学建模比赛, Finalist Award 0.2%	2020.1

论文

[1]. **W. Lin***, Z. Zhao*, X. Zhang, C. Wu, Y. Zhang, Y. Wang, W. Xie, "PMC-CLIP: Contrastive language-image pre-training using biomedical documents", *MICCAI* 2023.

[2]. C. Wu*, **W. Lin***, X. Zhang, Y. Zhang, Y. Wang, W. Xie, "PMC-LLaMA: Towards Building Open-source Language Models for Medicine", *JAMIA* (IF 7.9).

[3]. X. Zhang, C. Wu, Z. Zhao, **W. Lin**, Y. Zhang, Y. Wang, W. Xie, "PMC-VQA: Visual instruction tuning for medical visual question answering"

[4]. C. Wu*, J Lei*, Q. Zheng*, W. Zhao*, **W. Lin***, X. Zhang*, X. Zhou*, Z. Zhao*, Y. Zhang, Y. Wang, W. Xie, "Can GPT-4V(ision) Serve Medical Applications ? Case Studies on GPT-4V for Multimodal Medical Diagnosis."

[5]. P. Qiu, C. Wu, X. Zhang, **W. Lin**, H. Wang, Y. Zhang, Y. Wang, W. Xie, "Towards Building Multilingual Language Model for Medicine."

其他技能

后端开发：SpringBoot, Flask, Mysql, MongoDB	入门
框架：Pytorch, Tensorflow	熟悉