

林玮雄

wx_lin@outlook.com

15161762373

[HomePage](#)

[Google Scholar](#)

[Github](#)

Research Interest: Foundation Models, Multi-modal Representation Learning, Data Optimization



教育背景

上海交通大学大学, 信息工程, 硕士

2022.9 - 2025.3

- 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

PMC-LLaMA

Star 507

Cite 84

- 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 89

Cite 1

- 将单语言模型 PMC-LLaMA 迁移到多语言

多模态大模型

PMC-CLIP

Star 95

Cite 29

- 挖掘医学论文中的图文数据, 开拓了医疗多模态大模型的新思路
- PMC-CLIP 作为图文基础模型, 提供了强大的泛化能力; 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 125

Cite 45

- 以 PMC-CLIP 的训练数据为基础, 通过提示 ChatGPT 增强医疗视觉问答数据集, 大幅度领先之前方法
- 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

数据优化

Dedu Patch

Under Going

- 对图片块的冗余进行建模, 实现去噪和收敛加速
- 减少模型训练和推理的显存消耗, 提升整体速度

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 400 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

2024.1 - NOW

上海人工智能实验室, 深度学习实习

2023.7 - 2023.12

微软亚洲研究院, 机器学习实习

2022.7 - 2022.10

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", *npj Digital Medicine* (IF 15.2).

[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