

李子茉

性别: 男

出生日期: 2001/03

电话: 13349343833

邮箱: lizimo@nuist.edu.cn

个人作品集: [Link 链接](#)

研究方向: 多模态检索、LLM 内容理解

学历: 硕士研究生

毕业院校: 南京信息工程大学

专业: 数学

毕业时间: 2026 年 6 月

个人技能

- 开发能力:** 掌握 python、java、cpp 语言、具备 python(fastapi) 和 java(springboot) 的工程开发能力; 掌握 Pytorch 深度学习框架; 掌握 Transformers、Unsloth、verl、VLLM、SGLang 等大模型训练推理框架; 掌握 LlamaIndex、Langchain、Dify、ComfyUI 等 AI 应用开发框架;
- 研究能力:** 以全量 SFT 与 RLHF 训练过千亿参数量的模型, 对 LLM 训练的整个 Pipeline 理解比较深刻; 熟悉 DeepSpeed、FSDP、SP/TP 等分布式训练技术, 具有百卡集群的模型训练和部署经历; 精通 PPO、DPO、GRPO、DAPO、GSPO 等强化学习算法;

实习经历

vivo 南京研发中心

大模型算法实习生 2025.06 – 2025.09

- 参与研发 Vivo 的智能创作平台, 负责 LLM 的业务化对齐训练与工程化开发。

南信大影像技术工程有限公司

大模型算法实习生 2025.02 – 2025.05

- 主导开发面向气象服务材料写作场景的 LLM 应用链路, 包括 RAG 设计与优化、LLM 微调训练与部署。

项目经历

基于 LLM/MLLM 的图文广告推广链路

实习项目

- 项目背景:** 商务团队在与各类广告主合作过程中, 需要为 vivo 手机用户推送应用类广告。然而, 在传统的 AI 解决方案中, 其无法准确的捕捉 app 的核心投放意图, 导致生产的广告图文与 app 的相关性有一定的差异。
- 技术链路:** 接收广告主提供的 App 名称及一级、二级分类信息后, 首先调用意图识别模型识别出 App 相关的意图, 系统会优先尝试从自建的 AIGC 广告素材库进行匹配: 若匹配成功, 则返回广告素材图像。若未匹配成功, 则调用大模型自动生成图像 Prompt, 随后调用 Liblib API 完成广告图生成, 将素材回调给 AIGC 素材库。最后, 拿到的广告图交由 Qwen-VL-Plus 进行广告文案撰写, 从而输出完整的广告投放素材。
- 有效收益:** 行业识别的 Top1-Acc 达到了 99%, IRL(非法标签率)<1%; 意图-创意识别模型的相关性分数从 65 分 (SFT 之前) 提升到了 93 分 (SFT 之后)。生成广告图与 App 的一致性从 70% 提升至 97%, 生图的 badcase 率从 20% 降低至 0.4%; ctr 点击率从 2.5% 提升至 4%。

基于 RAG 与 NL2SQL 的气象服务材料写作系统

实习项目

- 项目背景:** 在气象业务场景中, 服务报告 (如呈阅件、强天气报告) 需要严格遵循模板结构, 并依赖最新的气象知识与历史数据。传统依靠人工撰写效率低, 直接使用大模型生成则常出现格式不符、知识错配、数据缺失等问题。因此, 我们以自动化、标准化和智能化报告生成为目标, 设计并实现了基于 RAG + NL2SQL 的智能体系统。
- 技术链路:** 核心架构基于 DeepSeek-R1:70B + RAG。针对直接使用大模型生成报告难以对齐多变模板的问题, 我们引入了 RAG。为提升检索精度, 我们对 embedding 模型进行轻量级微调, 并采用多路召回 + RRF 融合排序, 使正确类型与最新版本模板优先进入 top1-2, 气象知识优先进入 top3-5, 从而有效缓解模板错配与知识缺失问题。此外, 报告撰写需要历史气象数据支撑, 手动输入繁琐, 我们引入了 NL2SQL 能力, 基于开源模型进行全量微调, 可根据自然语言 query 自动生成业务相关的 SQL 查询, 从数据库中提取历史数据, 并与检索知识拼接, 提供给生成模型, 保证报告的数据支撑与知识完备性。
- 有效收益:** RAG 模块在多路召回 + RRF 排序 + 微调 Embedding 的机制下, Precision@5 提升 **87.2%**, MRR 达到 **0.79**; NL2SQL 子模块在真实问答集上 EX (执行正确率) 为 **82.5%**, EM (完全匹配) 为 **74.1%**, Top Logical Accuracy 达 **88.9%**。

科研与竞赛获奖经历

ADT-Net: A Text-Based Person Search Network Utilizing Adaptive Transformation 多模态检索领域-第一作者

2024 年”华为杯”中国研究生数学建模竞赛

全国二等奖 (13.05% 获奖率)