

Mengsha Liu

年龄：24 岁

邮箱：mengshaliu87@gmail.com

个人总结

拥有头部互联网大厂（字节跳动、腾讯、阿里、华为）核心 AI 算法与大模型研发经验，具备从预训练/微调、强化学习到复杂业务落地的全栈能力。

- LLM 智能体 (Agent) 研发：擅长训练大模型完成自主推理、任务规划与工具调用的多步骤任务，已在搜索与游戏两类场景走通从 prototype 到线上生产的完整链路。
- 游戏场景大模型落地：在腾讯 AI Lab 参与《王者荣耀》开放世界项目，训练面向剧情推进与玩家动态状态的 NPC 对话与日程生成模型，并持有开放世界 NPC 相关专利。
- 多模态理解与数据闭环：精通 RAG 检索增强与图文/图表多模态理解，擅长搭建「挖掘 bad case—数据增强—反向纠错」的自动化闭环，持续迭代模型能力。

教育背景

新加坡南洋理工大学 - 信号处理与机器学习 - 硕士 2024.08 - 2025.06

中山大学 - 智能科学与技术 - 本科 2020.09 - 2024.06

荣誉奖项：GPA 年级 Top 1%，国家奖学金（全校 top 1%）、华为智能基座奖学金、中山大学一等奖学金。

核心课程：自然语言处理、计算机视觉、高级程序设计、深度学习、机器学习

工作经历

字节跳动 | 大模型算法工程师 base 新加坡 | 2025.07 - 至今

- DeepSearch Agent：训练可自主调用检索工具、自主改写 Query 的搜索智能体，面向用户长尾 Query 完成「检索—改写—再检索」的多轮决策，召回单轮模型无法解决的长尾商品。
- 数据自迭代 Agent 流水线：每日自动从训练集中挖掘 bad case，一路经增强流构造同类难样本，一路经纠错流在存量数据中检索相似 Query 反向修正标注，双路并行提升标注准确率与模型判别能力。
- Query 理解：负责大规模搜索场景下的 Query 纠错、高精度意图分类与商品召回算法，熟悉强化学习训练、大模型预训练与微调、海量多源异构数据的清洗。

华为 | 搜索算法实习生 base 新加坡 | 2025.01 - 2025.05

- 参与核心搜索算法的研发与优化，探索思维链 (CoT) 在搜索系统中的应用，具备扎实的底层策略算法调优经验。

腾讯 AI Lab | 大模型算法实习生 | 2024.03 - 2024.08

- 主导生成式 AI Agent (智能体) 的底层研发，优化大模型在开放世界动态环境下的复杂多步骤任务 (Skills) 规划与执行能力。
- 为《王者荣耀》开放世界项目训练角色对话生成大语言模型，基于剧情推进与玩家动态信息生成个性化 NPC 对话，并研发 NPC 日程生成方案 (已申请专利)。

阿里巴巴达摩院 | 大模型算法实习生 | 2023.06 - 2023.12

- 提出基于 CoT 的多模态大模型信息提取方案，显著提升复杂长文本与图表的解析精度。
- 主导基于 60 万图像文本对及 800 万条指令问答数据集的多模态大模型微调。

产出成果

论文：

- “ChartThinker: A Contextual Chain-of-Thought Approach to Optimized Chart Summarization”, International Conference on Computational Linguistics (COLING), 2024. Accepted. 第一作者
- “LCCR: Long-Text COVID-19 Chinese Rumor Detection Dataset”, Neural Computing and Applications, 2023. Accepted. 第一作者
- “Multimodal Large Model Diagnostic Technology for Alzheimer’s Disease”, Lancet Digital Health, 2023. Major Revision. 第二作者

专利与软著：

- 第一发明人：一种结合语言生成模型技术的对话生成未来框架。
- 第一发明人：一种基于大语言模型的开放世界游戏 NPC 日程生成模型。
- 软件著作权：基于证据检索和预训练模型的长文本假新闻检测系统。

专业技能

大模型与 AI 技术：熟练掌握 LLM 预训练与微调 (SFT / RLHF)、强化学习训练、RAG 检索增强生成与 AI Agent 搭建；精通多模态与异构数据处理；具备将推理、规划与工具使用能力落地到搜索与游戏场景的实战经验。

开发技能：Python, C, C++, MATLAB, SQL | PyTorch, TensorFlow, Linux, DeepSpeed。