

# 梁祖杰

(+86)189-0262-1904 | 邮箱 | 个人主页 | GitHub (800 stars) | Google Scholar (350+ Citations) | 1998-04

## 关于我

- 意向方向: LLM Agent 应用、大模型后训练 (Post-Training)、知识图谱
- 扎实的金融领域大模型 (LLM) 应用、智能体 (Agent) 项目产研经验, 主导/参与多项信贷业务中知识图谱/用户理解/风控建模的 LLM/Agent 项目落地
- 丰富的 NLP/LLM/VLM 领域研究经验, 已发表 10+ 顶会论文 (ICLR/ACL/SIGIR/EMNLP/CIKM), 在投 9 篇, 一作 7 篇, 二作 (实习生一作) 3 篇, 参与AIOS Foundation建设, ICLR/ACL/EMNLP/ARR/NLPCC 审稿人
- 22 届校招获多家公司 (蚂蚁/微软/美团/讯飞) 头部/SSP 计划, 蚂蚁校招 A+ 硕士 P6
- 曾合作导师: 姜大昕博士 (阶跃星辰 CEO, 前微软全球副总裁)、徐粲 (微软高级研究员, WizardLM 一作)、肖仰华教授 (复旦大学)、张永峰教授 (罗格斯大学)、胡海峰教授 (中山大学)

## 教育经历

中山大学 (985 工程)   信息与通信工程   保送硕士研究生	2019.09—2022.06
研究方向为自然语言处理、对话系统、多模态理解。	
武汉理工大学 (211 工程)   通信工程   工学学士	2015.09—2019.06

## 工作/科研经历

蚂蚁集团   高级算法工程师 (13 级)   信贷事业群-网商银行-人工智能部	2022.6—至今
LLM x Agent	

业务背景: 风控测, B 端信贷业务的用户的行业跨度大, 数据碎而薄, 通过大模型 Agent 建设标准的行业/产业认知 → 画像加工 → 建模 → 策略的生产链路, 实现小微企业 KYB 画像的规模化加工

- 产业认知 Agent: 基于 LLM 高分辨率的长文处理能力, 对高质量的研报、数据库 (数库/万得/AliOpenKG/百科) 数据进行结构化, 替代人工专家生成高质量产业知识库 (主体 → 维度 → 指标), 当前在产业链、农村场景加工136 个风控画像指标, 可用率70%
  - 分治框架提升短窗口 LLM 的长文吞吐: 提出自底向上的分治框架 Segment+ (EMNLP 24) 进行 context-window scaling, 提升了短窗口 LLM 的长文本语料吞吐能力, 在 NarrativeQA/HotpotQA 等长文本 QA 较 SOTA 提升10%+, 在 NIAH(Babilong) 较 SOTA+5.28%, 4k 窗口下效果与 16k 窗口相当或更优。
  - 改进注意力模式提升 LLM 长文分辨率: 探索如何设计最优静态稀疏注意力模式, 获得更好的效率与感受野的权衡, 通过转化为 DAG 边集寻找问题, 让每个 token 仅关注距离为 2 的幂次方的位置, 实现在有限层注意力网络达到最宽感受野, 在 RULER 长文本测评上, 32k 窗口下较其他静态稀疏注意力方法绝对提升+5.7%, 较 Full Attn 加速3x), 已投稿NeurIPS 25。
- 建模 Agent: 大量风控子 PD 模型建模场景, 迭代调优成本高, 通过 LLM 构建自动建模智能体, 输入指定建模任务需求/数据, 智能体通过完成数据探查、特征工程、建模、模型评测、自反思迭代若干步骤任务, 最终输出最佳算法模型, 目前在风控建模数据集 CALM 超越专家模型 SOTA +3.59%, 在场景内经营子 PD 建模中挖掘头部优质客户增量+15%
  - [Plan] 树状反思框架提升 Agent 的思维节点质量及搜索效率: 提出自省式 MCTS(I-MCTS) 框架, 通过树状 (父节点和兄弟节点) 反思结合混合奖励反馈机制, 提升了蒙特卡洛树搜索的效率以及节点质量, 在 AMLB/Kaggle 的 20+ML 任务中性能较 SOTA(MetaGPT) 提升+6%, 已投稿ACL 25
  - [Memory] 卡片盒笔记法实现动态智能体记忆机制: 非结构化的领域算法经验库 (如语雀/ATA 文档) 通过 RAG 这种偏相关性召回的方法难以提升 Agent 的性能效果, 提出 A-MEM(Agentic Memory) 框架, 通过卡片盒笔记法自主完成领域知识结构化, Agent 在遍历非结构化文档过程自主完成动态的经验构建 (新增、链接、删除、更新), 在 LoCoMo Benchmark 上, 该方法超同类 MemoryBank、MemGPT 精度+7.5%, 已投稿ACL 25 (获台大李宏毅2025 Agent Tutorial引用)
  - [Action]Rule-RL 实现 ML 领域代码生成优化: Agent 执行模块核心采用 CodeAct 框架, 利用 Deep-ML 等数据集基于 Code-R1 (VeRL) 框架探索了多种 rule-based 奖励信号 (语法/语义相似度/测试用例) 增强 ML 领域的 NL2Code 能力, 在 DS-1000 子集中测评 PASS@1, 将 qwen2.5-coder-7b 从44% 提升至67.5%, 超越 qwen2.5-72b (50.5%)

### LLM x 推荐系统

业务背景: 在存款理财营销业务中, 探索基于 LLM 的生成式序列推荐模型

- **跨层蒸馏加速生成式推荐模型**: 研究 LLM 的层数对序列推荐任务的影响, 并提出一种轻量化 LLM 模型 (<1B) 训练方法 SLMRec, 通过跨层知识蒸馏对齐教师-学生模型表征空间, 并结合多目标监督信号, 实现仅**13%** 的模型参数下性能逼近 7B 模型 (LLMRec), 训练/推理速度提升**6.6 倍/8.0 倍**, 被 **ICLR 25**接收。

## LLM x 知识图谱

**业务背景**: B 端信贷业务从供应链金融走向产业链金融, 对数据薄的小微企业进行精细画像, 极度依赖高质量产业链认知图谱设施。通过一套大模型图谱构建方案, 从 0 到 1 实现**132** 个产业细分链构建, 覆盖类目**3w+**, 覆盖关系**10.8w**, 识别超**2100w** 产业链上下游的小微企业, 目前是网商产业链金融 KYB 客群识别的**基础设施**。

- **统一知识抽取引擎实现属性规模化挖掘**: 金融领域业务场景多 (自证/票据/网站/..), 覆盖任务异构多样 (实体/二元组/多元组), 40% 场景为低资源冷启动。为此构建了 Ant-UIE 统一知识抽取引擎, 分别通过 PEFT 技术提出彩票提示微调 (**ACL 23**) 和层级提示微调 (**CIKM 23**), 缓解灾难性遗忘 (lifelong setting, F1 **+9.41%**) 和多任务跷跷板问题 (few-shot setting, GLUE AVG **+2.6%**), 支持**170+** 微观属性抽取, 平均准确率**90%+**, 赋能信贷 KYB 主营产品、产业身份等基础设施, 其中主营产品覆盖**4300w** 企业 (**84%**), 产业身份覆盖**4600w** 企业 (**91%**)
- **缓解 LLM 幻觉实现工业级图谱自动校验**: 利用 LLM 对大规模 (44w+) 的自动加工出来的领域图谱关系对校验提质, 结合高质量参考源, 提出不确定度知识边界 SFT 算法 (CoKE, 已投稿**ACL 25**) 缓解大模型幻觉 (LoRA tuning 下, Llama2-Chat-7B 在 TriviaQA 方法绝对提升**+10.8%**), 最终过滤低质关系对**20w+** (人机校验一致率: **86.2%**), 图谱知识校验效率提升**10 倍**。
- **标准化大模型评测基准建设**: 围绕真实金融业务场景及图谱推理前沿领域, 利用真实语料/合成样本方式构建多项业界大模型评测基准: 1) 主导蚂蚁集团 2023 外滩大会大模型赛题 “**AFAC23-金融知识抽取**” 设计 (10w 奖池), 为业界贡献首个金融领域多场景知识抽取 benchmark (**3k** SFT 样本, **2.6w** 行业研报/中标公示/保险条款语料), 共吸引 788 支队伍参赛; 2) 针对知识推理领域, 通过 self-instruct 等技术建设多项相关的大模型任务基准, 包含概念推理能力评估基准 CR-LLM (**ACL 2024**)、历史事件类比能力评测基准 PmP (已投稿**ACL 2025**)、以及引经据典评测基准QUILL (已投稿**ACL 2025**), 分析了大模型在复杂推理任务的行为和局限性。

**阿里集团** | 广告算法工程师 (实习生) | CTO 线-AI 国际事业部 2021.6—2021.9

- **计算广告 pCVR 校准纠偏**: 团队负责 LAZADA 广告主自动出价机制/策略/模型等产研, 我主要针对自动出价滞后有偏问题, 提出基于点击质量模型的 pCVR 校准, 建模密集点击行为信号缓解 CVR 稀疏, 实验提效显著 (CQM 模型 AUC **+15.86%**, pCVR 校准误差 GCN **-1.77**, 方案在 SEA 六国线上推全)

**微软集团** | NLP 研究实习生 | Bing 2020.10—2021.5

- **生成式对话研究**: 针对 Bing Chat 研究生成式 ChatBot, 探索搜索引擎的全新交互形态, 我主要负责生成式对话系统在多模态及推荐领域的研究, 期间创新性地提出了一种视觉常识 (Visual RAG) 的开放域对话 Agent (**ACL 2021**, 最早期的端到端多模态 VLM 对话 Agent 工作之一), 以及提出一种基于可训练的语言模版的对话式推荐系统提升域外泛化性 (**EMNLP 2021**), 获微软 SP return offer。

**中山大学** | 在校研一期间 2019.9—2020.6

- **多模态幻觉缓解**: 在多模态理解任务如 VQA 中, 模型由于模态偏差 (语言先验) 易产生幻觉, 在分布外的测试场景中性能差。我围绕模态偏差问题展开多模态训练范式的创新研究, 分别提出了反事实自监督对比学习机制 CL-VQA (较 SOTA 方法**+1.44%**, 最早期成功在反事实建模中实施对比学习的工作之一, 目前引用 **150+**), 以及单模态先验重加权目标函数 LPF (较 baseline UpDn 方法 (**+15.85%**)), 分别被**EMNLP 20**及**SIGIR 21**接收。

**华为** | 软件研发工程师 (实习生) | 消费者 BG 2018.7—2018.9

- 通过优化华为智能手机中故障自检测算法, 计算更准确的故障结果与维修建议, 获转正 offer。

## 荣誉奖励

- 蚂蚁集团-T STAR 优秀工程师, 2025
- 蚂蚁集团-信贷事业群优秀星青年, 2024
- ACM SIGIR 2021 Student Travel Grant
- ICCV 2019 Wider Challenge Track Four, invited for poster demonstration, Rank 5th, Korea [**leaderboard**]
- 中山大学优秀毕业生、学业优秀奖学金
- 武汉理工大学优秀毕业生、学业优秀奖学金
- 全国大学生节能减排科技竞赛, 三等奖
- 湖北省大学生电子设计竞赛, 二等奖

## 研究成果

### Publications

- 1. **Zujie Liang**, Weitao Jiang, Haifeng Hu, Jiaying Zhu. “Learning to Contrast the Counterfactual Samples for Robust Visual Question Answering”. **EMNLP**, 2020. [PDF] [Code] [Slides] (Citation **150+**)

- 2. **Zujie Liang**, Huang Hu, Can Xu, Chongyang Tao, Xiubo Geng, Yining Chen, Fan Liang, Daxin Jiang. “Maria: A Visual Experience Powered Conversational Agent”. *ACL*, 2021. [PDF] [Code] [Slides]
- 3. **Zujie Liang**, Huang Hu, Can Xu, Jian Miao, Yingying He, Yining Chen, Xiubo Geng, Fan Liang, Daxin Jiang. “Learning Neural Templates for Recommender Dialogue System”. *EMNLP*, 2021. [PDF] [Code]
- 4. **Zujie Liang**, Haifeng Hu, Jiaying Zhu. “LPF: A Language-Prior Feedback Objective Function for De-biased Visual Question Answering”. *SIGIR*, 2021. [PDF] [Code] [Slides]
- 5. **Zujie Liang**, Fan Liang. “TransPCC: Towards Deep Point Cloud Compression via Transformers”. *ICMR*, 2022. [PDF] [Code]
- 6. **Zujie Liang**, Feng Wei, Jie Yin, Yuxi Qian, Zhenghong Hao, Bing Han. “Prompts Can Play Lottery Tickets Well: Achieving Lifelong Information Extraction via Lottery Prompt Tuning”. *ACL*, 2023. [PDF]
- 7. Jingping Liu, Tao Chen (Intern), **Zujie Liang**, et al. “Hierarchical Prompt Tuning for Few-Shot Multi-Task Learning”. *CIKM*, 2023. [PDF]
- 8. Nianqi Li (Intern), .., **Zujie Liang**, et al. “CR-LLM: A Dataset and Optimization for Concept Reasoning of Large Language Models”. *ACL Findings*, 2024. [PDF]
- 9. Wei Shi (Intern), .., **Zujie Liang**, et al. “Segment+: Long Text Processing with Short-Context Language Models”. *EMNLP*, 2024. [PDF]
- 10. Wujiang Xu, Qitian Wu, **Zujie Liang**, et al. “SLMRec: Distilling Large Language Models into Small for Sequential Recommendation”. *ICLR*, 2025. [PDF]

## Under Review

- 1. **Zujie Liang**, et al. “I-MCTS: Enhancing Agentic AutoML via Introspective Monte Carlo Tree Search”. *ACL (under review)*, 2025. [PDF]
- 2. Wujiang Xu, .., **Zujie Liang**, et al. “A-MEM: Agentic Memory for LLM Agents”. *ACL (under review)*, 2025. [PDF]
- 3. Lida Chen (Intern), **Zujie Liang**, et al. “Teaching Large Language Models to Express Knowledge Boundary from Their Own Signals”. *ACL (under review)*, 2025. [PDF]
- 4. Nianqi Li (Intern), **Zujie Liang**, et al. “MultiLingPoT: Enhancing Mathematical Reasoning with Multilingual Program Fine-tuning”. *ACL (under review)*, 2025. [PDF]
- 5. Nianqi Li (Intern), .., **Zujie Liang**, et al. “Past Meets Present: Creating Historical Analogy with Large Language Models”. *ACL (under review)*, 2025. [PDF]
- 6. Jincheng Dai (Intern), **Zujie Liang**, et al. “TriggerGPT: Trigger Token Prediction for Universal Tool Learning”. *ACL (under review)*, 2025.
- 7. Wujiang Xu, Yunxiao Shi, **Zujie Liang**, et al. “InstructAgent: Building User Controllable Recommender via LLM Agent”. *ACL (under review)*, 2025. [PDF]
- 8. Jin Xiao (Intern), .., **Zujie Liang**, et al. “QUILL: Quotation Generation Enhancement of Large Language Models”. *ACL (under review)*, 2025. [PDF]
- 9. Lida Chen (Intern), .., **Zujie Liang**, et al. “PowerAttention: Exponentially Scaling of Receptive Fields for Effective Sparse Attention”. *NeurIPS (under review)*, 2025. [PDF]