

冀轲

1ke-ji.github.io  Google Scholar Profile

keji@link.cuhk.edu.cn  (+86) 18338991628



教育经历

香港中文大学 (深圳)

博士, 计算机科学, 导师王本友教授。研究方向为 **NLP**、**LLM** 和 **Reasoning Alignment**。

深圳

2024.9 - 至今

东南大学

硕士, 软件工程, 导师汪鹏教授。曾获国家奖学金、东南大学优秀硕士学位论文和优秀毕业生。

南京

郑州大学

本科, 软件工程, 曾获一等奖学金和优秀毕业生。

2021.9 - 2024.6

郑州

2017.9 - 2021.6

研究兴趣

研究兴趣主要围绕: 在**数据高效学习与大模型自主对齐**。针对目标域少标注样本场景, **HierVerb (ACL' 23)** 和 **HierICRF (IJCAI' 24)** 通过重定义 post-training 优化目标, 将下游任务与基座模型特征空间对齐, 利用基座模型内在先验压缩目标解空间, 提升了单个样本的优化效率 (1-shot macro-F1 绝对值提升 74.79%), 针对大规模域适应场景, **HuatuoGPT-II (COLM 2024)** 提出一阶段域适应策略, 在 Mid-training 阶段利用动态采样融合指令数据, 解决了预训练与微调阶段的分布偏移问题。在大模型自主对齐方面, 我发表了 **UPFT (NeurIPS'25)** 来挖掘模型内在监督信号, 利用前缀自一致性作为弱监督信号提升模型目标域推理能力, **AL (ACL Findings'25)** 通过基于目标领域数据文档构建事实性偏好对齐目标, 实现了无人工标注下的模型域适应自主对齐。PCL (ACL Findings' 25) 从自动生成对话中学习人设偏好。额外的, 在 Agent RL 方面, 于 **WebExplorer (Tech. Report)** 中针对长程复杂任务设计了一套自动化数据合成算法, **SmartRAG (ICLR'25)** 通过 Reward Shaping 和 RL 在模型探索中实现了成本与准确率的平衡。

学术论文

The First Few Tokens Are All You Need: An Efficient and Effective Unsupervised Prefix Fine-Tuning Method for Reasoning Models

作者: Ke Ji, Jiahao Xu, Tian Liang, Qiuwei Liu, Zhiwei He, Xingyu Chen, Xiaoyuan Liu, Zhijie Wang, Junying Chen, Benyou Wang, Zhaopeng Tu, Haitao Mi, Dong Yu *Neurips2025* 一作

背景: 当前基于 RFT 范式下的 Self-training 方法数据生产成本高, 可以通过采样快速得到正确路径的简单题对模型增益小, 而对模型有增益的难题所需的采样时间往往巨大, 如何使模型在其准确率较低的场景中进一步增强是一个主要挑战。

- 本文提出并通过定量实验证前缀一致性 (Prefix Self-Consistency) 这一现象, 从贝叶斯角度建模该问题, 并受此启发提出一种无监督的前序步骤微调方法 (UPFT)。
- 无监督设置下 UPFT 效率显著优于传统全 token 微调, 同时仅需单次采样即可取得与传统方法 16 次采样有竞争力的性能, 并将训练成本降低约 75%。UPFT 协议简单但有效, 所需要的只是模型本身和问题集合, 不需要基于标准答案进行大量拒绝采样。

Unlocking LLMs' Self-Improvement Capacity with Autonomous Learning for Domain Adaptation

ACL2025 Findings 一作

作者: Ke Ji, Junying Chen, Anningzhe Gao, Wenyu Xie, Xiang Wan, Benyou Wang

背景: 大模型在进行 Domain Adaptation 时, 简单的持续预训练 (Continue Pretraining) 效率较低, 而基于指令的监督微调 (SFT) 不仅依赖昂贵的人工标注, 现有的自动化数据合成策略会引入较多的事实性错误噪音, 容易导致模型产生幻觉或过度自信。

- 提出了一种基于知识回溯的自主学习框架, **强化模型的事实性偏好**, 提升模型在目标领域回答的准确性。利用开卷回答 (Fact-based Answer) 作为事实基准, 闭卷生成作为潜在幻觉负样本 (Hallucinated Answer), 自动构建偏好对并基于 DPO 优化模型事实性偏好。
- 本方法无需任何人工或强模型 (GPT-4) 标注, 仅依赖模型本身和无标注语料即可实现迭代。在 3 个领域 (医学、法律、通用常识) 的 9 个数据集上, 验证了在文档规模从 1K 扩增到 1M 过程中, 该方法能显著提升知识准确性。

WebExplorer: Explore and Evolve for Training Long-Horizon Web Agents

Technical Report

作者: Junteng Liu, Yunji Li, Chi Zhang, Jingyang Li, Aili Chen, Ke Ji, Weiyu Cheng, Zijia Wu, Chengyu Du, Qidi Xu, Jiayuan Song, Zhengmao Zhu, Wenhui Chen, Pengyu Zhao, Junxian He

背景: 现有 Web Agent 在处理长程复杂任务时表现不佳, 主要受限于高质量、高难度训练数据的匮乏。传统的基于规则或图谱的数据生成方法难以模拟真实且复杂的网络探索行为。

- 提出了一种基于模型的“探索与演化”数据合成框架, 利用 LLM 自主探索模拟真实信息空间, 并设计“由长变短”的 query 进化策略, 自动生成需多步推理的高难度问题。
- 构建了 WebExplorer-8B 模型, 通过监督微调 (SFT) 与强化学习 (RL) 结合的训练范式, 支持 128K 上下文及 100 轮工具调用。该 8B 模型在 BrowseComp 等高难度评估基准上取得了与现有方法 72B 参数量模型有竞争力的性能。

Hierarchical Verbalizer for Few-shot Hierarchical Text Classification

ACL2023 一作

作者: Ke Ji, Yixin Lian, Jingsheng Gao, Baoyuan Wang

背景: 目前预训练语言模型在面临结构化任务时性能不佳, 现有工作往往依赖于大量有标注数据, 同时目前层次化结构任务的评测指标缺乏对于层次结构一致性的考虑。

- 本文首次强调了结构化分类任务的数据稀缺问题, 定义了少样本实验设置, 后续在此领域的工作均使用本文提出的少样本数据切分方法, 为研究低资源下层次文本分类提供了一个有价值的实验设置基础。
- 提出了一个多层次 Verbalizer 方法 HierVerb, 通过统一上下游任务范式, 高效迁移 LLM 中的先验知识适应到结构化任务, HierVerb 相比较于先前方法平均有 micro 57.58% 以及 macro-F1 74.79% 的提升。

Domain-Hierarchy Adaptation via Chain of Iterative Reasoning for Few-shot Hierarchical Text Classification

IJCAI2024 一作

作者: **Ke Ji**, Peng Wang, Wenjun Ke, Guozheng Li, Jiajun Liu, Jingsheng Gao, Ziyu Shang

背景: 预训练语言模型中固有的非结构化先验知识使得其在进行结构化多分类任务中表现不佳, 现有范式与预训练模型的无结构先验知识背道而驰, 忽视了建模层次结构任务的重要性。

- 本文首次将层次文本分类任务建模成层次结构迭代推理问题, 相比以往方法在层次一致性指标上提升 5%。

SmartRAG: Jointly Learn RAG-Related Tasks From the Environment Feedback

ICLR2025 三作

作者: Jingsheng Gao, Linxu Li, **Ke Ji**, Weiyuan Li, Yuzhuo Fu, Bin Dai

背景: 现有 RAG 系统往往独立训练检索器与生成器, 导致模块间缺乏协同, 模型难以根据问题难易动态调整检索策略, 造成计算资源浪费和推理效率低下。

- 提出了一种基于环境反馈的端到端联合优化框架, 利用 PPO 将“何时检索”、“检索什么”与“如何回答”三个子任务统一建模。
- 设计了权衡准确率与检索成本的复合奖励函数, 首次将多步决策 RL 思路应用于 RAG 全流程。实验表明, 该方法能以显著降低的检索频次在多个开放域问答基准上取得优于 Self-RAG 的性能。

Enhancing Persona Consistency for LLMs' Role-Playing using Persona-Aware Contrastive Learning

ACL2025 Findings 一作

作者: **Ke Ji**, Yixin Lian, Linxu Li, Jingsheng Gao, Weiyuan Li, Bin Dai

背景: LLM 在角色扮演任务中常出现角色一致性差、情感表达缺失的问题, 现有方法依赖昂贵的人工标注, 且容易导致模型通用知识遗忘。

- 提出了一种无需人工标注的自进化对齐框架 (PCL), 创新性地设计了“角色链”(Chain of Persona) 机制, 引导模型通过自我提问与反思生成符合人设的高质量回复。

RAG-Instruct: Boosting LLMs with Diverse Retrieval-Augmented Instructions

EMNLP2025 三作

作者: Wanlong Liu, Junying Chen, **Ke Ji**, Li Zhou, Wenyu Chen, Benyou Wang

背景: 现有 RAG 微调方法大多基于特定的 QA 数据集, 不仅任务类型单一, 且在面对检索无用或部分有用的复杂检索场景时, 模型的鲁棒性有限。

- 提出了一种通用的 RAG 指令数据合成框架, 引入“指令模拟”技术, 利用现有高质量指令集引导生成过程, 确保合成数据的任务多样性与高质量。
- 定义了五种 RAG 查询范式(从无用文档到多文档回答), 覆盖现实世界复杂的检索场景。基于该方法构建的数据集使 LLM 在 11 项不同任务上展现出优于 Self-RAG 等基线的零样本性能和鲁棒性。

实习经历

MiniMax 大模型研究实习生

2025.6-至今

- Interleaved Agent: 负责 HLE 榜单, 主要提升模型解决复杂 STEM 题的能力, 此类题涉及复杂推理与数值计算, 以及冷门专业知识。设计实验定义观测指标来迭代用于提升此能力的数据, 消融和实验 Interleaved agent 冷启动和 RL 中的各种变量。
- CoT Synthesis: 探索 COT 数据对性能的影响, verifiable 场景采用 1Q 多 A 训练 + 迭代数据过滤, non-verifiable 场景采用多源 response 融合来合成高质量 thinking。

腾讯科技(深圳)有限公司 AI Lab 语言智能研究实习生

2024.10-2025.5

- 主要工作内容: 关注大模型推理能力的提升, 研究推理场景面临的复杂模式域适应以及高成本采样等挑战, 探索高效 self-training 算法。
- 发表论文 [UPFT](#), 一种无需拒绝采样的高效 Self-training 算法。相比较于传统 RFT 范式的算法, UPFT 在采样成本仅为 1/16 和训练时间缩短 75% 的成本下, 取得了有竞争力的性能, 文章在 [Twitter](#) 和 [量子位](#) 上被广泛转载。

北京红棉小冰股份有限公司(原微软小冰) 自然语言处理研究实习生

2022.1-2023.2

- 调研 prompt-based tuning 和大模型训练推理框架, 参与负责超级小冰大模型的数据收集以及模型训练。
- 构建对话系统以及 NLU 服务, 同时发表一作论文 [ACL2023](#), 在低资源结构化场景下 (1-shot) 取得了 74.79% macro-F1 的提升。

项目经历

HuatuoGPT-II, One-stage Training for Medical Adaption of LLMs

COLM2024 三作

作者: Junying Chen, Xidong Wang, **Ke Ji**, Anningzhe Gao, Feng Jiang, Shunian Chen, Hongbo Zhang, Song Dingjie, Wenyu Xie, Chuyi Kong, Jianquan Li, Xiang Wan, Haizhou Li, Benyou Wang

背景: 当前大模型在进行 domain-specific adaptation 时, 面临复杂的数据混合协议以及两阶段域适应引入的灾难性遗忘两个主要挑战。

- 提出一种有效的一阶段域适应方法, 该方法基于统一的指令数据格式, 将 Continue Pretrain 和 SFT 两阶段融合, 并设计了一种基于重要性排序的数据采样策略。
- HuatuoGPT-II 在多个医疗资格和专业评测中取得了优异成绩, 并通过 2023 年的国家执业药师考试, 同时大量的消融实验证了本文方法的有效性。
- 项目 Repo: [Github: 400 star](#); 模型权重: [HuatuoGPT-II](#); 开源数据: [持续预训练语料](#) 和 [指令微调数据集](#)。

HuatuoGPT-o1, Towards Medical Complex Reasoning with LLMs

ACL2025 Findings 三作

作者: Junying Chen, Zhenyang Cai, **Ke Ji**, Xidong Wang, Wanlong Liu, Rongsheng Wang, Benyou Wang

背景: 医疗场景下对大模型输出内容的可靠性要求苛刻, 现有工作更多集中在如何“记忆”医疗知识, 同时对高质量标注数据需求较大, 而对于如何增强医疗模型的推理能力关注较少。

- 首次将复杂推理应用在医疗领域来增强模型回答的可靠性, 提出 HuatuoGPT-o1 模型。对于医疗领域的推理过程复杂且步骤不明确导致的正确性验证困难这一问题, 我们利用具有挑战性的闭卷考试题目构建可验证医学问题, 然后基于可验证医学问题进行强化学习, 最后通过实验证了本文所提框架的有效性。
- 项目 Repo: [Github: 1.3k star](#); 模型权重: [HuatuoGPT-o1](#); 开源数据: [高质量可验证医疗问答数据集](#)。

所获主要荣誉和学术服务

- 主要荣誉

- AIMO2 金牌 (HUAWEI-CUHK SZ)
- 东南大学优秀毕业生、东南大学优秀硕士学位论文
- 国家奖学金 (1/130)、研究生优秀奖学金、三好学生标兵 (1/130)
- 一等奖学金、优秀毕业生

2025 年
2024 年
2023 年
本科

- 审稿人

- ICLR (2025-); ACL (2024-); EMNLP (2024-); CVPR (2026-); AAAI (2024-)

- 教学助理

- 2025.02 - 2025.06, [CSC6052 Natural Language Processing](#), 香港中文大学 (深圳)
- 2024.09 - 2025.01, [CSC6203 Large Language Models](#), 香港中文大学 (深圳)

技能

语言技能: 普通话 (母语), 英语 (TOEFL-95)

常用编程技能与框架: pytorch; linux; vllm; verl; ms-swift; huggingface; deepspeed