阿东玩 AI

138-0000-0000 | adong@tsinghua.edu.cn | 清华大学 | 计算机科学与技术 · 硕士 | 1998-06



公开模板演示案例：经历与指标来自用户指定的 LLM-Resume-Template 公开模板，不代表真实候选人。清华大学计算机科学与技术专业硕士在读，预计 2026 年 6 月毕业。主攻大模型方向，在模型压缩、微调、Agent 决策优化和强化学习方面有系统研究与实践；熟悉 LLaMA 架构、RLHF、DPO、PPO、MCTS、DeepSpeed、PyTorch 和 Hugging Face Transformers。

教育背景

清华大学 | 计算机学院 | 计算机科学与技术 | 硕士 | 专业排名前 5% 2023.09—2026.06(预计)

清华大学 | 计算机系 | 计算机科学与技术 | 学士 2019.09—2023.06

专业技能

- 熟悉 Python 及 PyTorch 框架，掌握 Hugging Face Transformers 等大模型开发工具。
- 深入理解大模型微调技术，如 P-tuning、LoRA 等，并有丰富的提示工程实践经验。
- 熟悉 LLaMA 等主流模型架构，以及 RLHF、DPO、PPO、MCTS 等偏好对齐、强化学习和搜索增强方法。
- 掌握基于知识蒸馏的模型压缩方法，对大模型轻量化有深入研究。
- 熟悉 DeepSpeed 等分布式训练框架，具备大模型训练与部署性能优化能力。

科研经历

基于强化学习的自主决策 Agent 系统研究 2024.03—至今

- **问题背景**: 现有 LLM-based Agent 在复杂任务中存在决策路径冗余、缺乏长期规划能力等问题，难以在多步骤任务中做出最优决策序列。
- **研究内容**: 1) 提出基于 PPO 的 Agent 决策优化框架，将 LLM 作为策略网络，通过任务完成度、步骤效率等构建奖励函数，实现端到端的强化学习训练；在 ALFWorld、WebShop 等多个 Benchmark 上验证，对比 ReAct、Reflexion 等基线方法，任务成功率提升 12%~18%，平均决策步数减少 25%。
- 2) 设计分层强化学习架构，将复杂任务分解为高层规划和低层执行两个层次，高层 Agent 负责子目标分解，低层 Agent 负责具体动作执行；通过 Hindsight Experience Replay (HER) 提升样本利用效率，在稀疏奖励环境下收敛速度提升 40%。
- 3) 引入 Monte Carlo Tree Search (MCTS) 与 RL 结合的混合决策机制，在决策前进行多步模拟，优化探索-利用平衡；在长序列任务中，相比纯 RL 方法，样本效率提升 35%，最终性能提升 8%。
- **相关成果**: 第一作者论文已投稿 NeurIPS 2026，目前处于 Under Review 状态；开源代码在 GitHub 获得 500+ stars。

大模型知识蒸馏与模型压缩技术研究 2023.09—2024.02

- **问题背景**: 大规模语言模型部署成本高昂，在边缘设备和实时应用场景下难以满足推理延迟要求。
- **研究内容**: 1) 提出基于层级知识蒸馏的模型压缩方法，同时对齐 Teacher 和 Student 模型的注意力分布、隐层特征和输出概率分布，在保持 95% 性能的前提下，将 LLaMA-13B 压缩至 3B 参数量；2) 结合结构化剪枝和低秩分解技术，设计自适应剪枝策略，在 MMLU、GSM8K 等 Benchmark 上，相比标准蒸馏方法，平均性能提升 4.2%。
- **相关成果**: 第一作者论文发表于 NeurIPS 2024 (CCF-A)，论文被引用 15 次。

实习经历

字节跳动 | 人工智能实验室 | 大模型算法实习生 2024.06—2024.12

- **负责部分**: 参与豆包大模型的持续预训练和指令微调过程，负责领域数据构造与模型性能优化，参与 Agent 能力增强项目开发。
- **实习内容 1**: 根据代码生成和数学推理业务场景，挖掘模型在复杂逻辑推理上的能力短板，构造高质量 CoT 数据 5B tokens，进行持续预训练；在 HumanEval、MATH、GSM8K 等 benchmark 的消融实验中，稳定提升模型平均得分 5~8pp。
- **实习内容 2**: 针对模型在多轮对话中存在的指令遵循不足和上下文理解偏差问题，依据真实业务对话流程，收集和构造多轮对话数据集 50 万条，设计多任务学习目标，基于 LLaMA-3-8B 进行 SFT 和 DPO 训练，最终在内部 benchmark 上平均得分继续提升 3pp，用户满意度达到 87%。

- **实习内容 3**: 分析用户反馈数据，发现模型在工具调用场景下准确率不足（仅 72%），采用 ReAct 范式构造工具调用训练数据，基于 Qwen-14B 模型，融合 Function Calling 和 ReAct 数据进行混合训练，工具调用准确率提升至 89%；同时研究 Agent 规划能力改进，基于内部平台实现 Tree-of-Thought 推理框架，最终在复杂任务上成功率进一步提升 12%。

腾讯 | 微信 AI 团队 | 算法工程实习生 2023.06—2023.12

- **负责部分**: 参与智能客服大模型的训练与部署，负责模型推理性能优化和压缩加速。
- **实习内容 1**: 针对线上推理延迟高（P99 达到 3.2s）的问题，实施 KV Cache 优化、算子融合、批处理策略等多项优化措施，将推理延迟降低至 1.8s，QPS 提升 65%。
- **实习内容 2**: 研究模型量化技术，实现 INT8 量化和混合精度推理，在保持 98% 性能的前提下，显存占用降低 55%，使得单机可部署模型数量提升 2 倍，大幅降低部署成本。
- **实习内容 3**: 基于 vLLM 框架进行二次开发，实现动态批处理和连续批处理功能，优化调度策略；在高并发场景下（1000+ QPS），相比原生实现，吞吐量提升 80%，P95 延迟降低 40%。

项目经历

基于自我纠错机制的智能 RAG 检索系统 2024.01—2024.05

- **项目背景**: 为解决企业内部知识库检索效率低下的问题，设计并实现一套能够自主优化检索策略的 RAG 系统。
- **核心痛点**: 最初版本的 RAG 系统召回率仅有 65%，存在关键词匹配不准、语义理解偏差等问题，导致用户查询难以获取准确答案，几乎无法投入实际使用。
- **技术方案**: 1) 借鉴 ReAct 思想，设计迭代式检索策略——当检索结果置信度低于阈值时，Agent 自动触发反思机制，分析失败原因并重新生成优化后的查询；2) 实现多策略融合检索，包括向量检索、BM25 关键词检索和混合检索，根据查询类型自适应选择最优策略；3) 引入检索结果重排序模块，基于交叉编码器对召回文档进行精排，提升 Top-K 结果相关性。
- **实验验证**: 在企业内部知识库和 MS MARCO、Natural Questions 等公开数据集上进行了完整的消融实验，验证了各模块的有效性——迭代式检索使召回率提升 15%，多策略融合再提升 3%，重排序模块贡献 2%。
- **最终效果**: 系统召回率从初始的 65% 提升至 85%，平均检索轮次 1.3 次，用户满意度达到 82%；项目成果在组内进行分享，获得导师和团队成员认可，相关技术方案已应用于实际生产环境。

基于分层记忆架构的长程对话 Agent 系统 2023.07—2023.12

- **项目背景**: 现有 Agent 在长对话场景下，由于上下文窗口限制和记忆管理不当，导致信息遗忘严重、检索效率低下，难以维持长期交互。
- **架构设计**: 1) 工作记忆层维护最近 K 轮对话的完整上下文，直接输入 LLM；2) 短期记忆层对近期对话自动摘要，提取关键信息和决策依据；3) 长期记忆层基于重要性评分和时间衰减机制，筛选重要记忆永久保存。
- **关键技术**: 1) 实现基于 BM25 + 向量相似度的混合检索机制，快速定位相关历史记忆；2) 设计记忆压缩算法，将冗长对话通过 LLM 总结提炼；3) 引入记忆更新策略，当新信息与旧记忆冲突时自动触发记忆修正流程。
- **实验结果**: 在多轮对话任务上，信息有效保留率达到 92%（对比 Baseline 的 68%），记忆检索速度提升 3 倍（从平均 450ms 降至 150ms）；在 100 轮以上的超长对话中，任务完成率提升 28%。

个人荣誉

- NeurIPS 2024 第一作者论文 (CCF-A 类会议)，研究方向：大模型知识蒸馏与压缩。
- Kaggle LLM Science Exam 竞赛金牌 (Top 1%，128/2664)。
- 天池 TIANCHI LLM Fine-tuning Challenge 第一名。
- 清华大学 2024 年度学业优秀奖学金（一等）、优秀研究生；2023 年度国家奖学金。
- 第十四届中国大学生程序设计竞赛 (CCPC) 全国银奖；美国大学生数学建模竞赛 (MCM/ICM) Meritorious Winner。
- GitHub 开源项目 Agent-RL-Framework 获得 500+ stars。