

宋佳桐

+86 16631625879 jt625879 songjiatong03@outlook.com

个人主页: mydocts.github.io/jiatong.github.io/



教育经历

香港中文大学 (深圳)

数据科学硕士

09/2025 - 06/2027

大连理工大学 (985)

GPA 90.5

数理基础科学学士

09/2021 - 06/2025

科研经历

Bookmark: 面向长链推理的测试时扩展框架

11/2025 - 12/2025

- 项目背景**: 在数学竞赛 (AIME) 等复杂推理场景下, 大模型面临上下文窗口限制和噪声累积问题 ("Lost in the Middle" 现象), 难以进行高质量长链推理。
- 核心创新**: 设计动态上下文压缩与检索机制: 将推理过程分解为多个子步骤 (<substep> 标签), 每步生成摘要并归档; 遇到回溯需求时 (<reflection>), 通过检索机制提取相关历史信息而非保留完整上下文。
- 项目难点**: 如何在极长推理链 (1.5w+ Tokens) 中通过 Bookmark 摘要保持核心逻辑不丢失, 以及在 GRPO 训练中解决通过奖励函数引导模型自主学习结构化标签与推理逻辑的对齐难题。
- 两阶段训练**: 第一阶段使用 SFT 让模型学习结构化标签生成; 第二阶段采用 GRPO 强化学习, 设计准确性奖励 ($w_{ans} = 1.0$) 与格式奖励 ($w_{fmt} = 0.1$) 优化策略。
- 项目成效**: 基于 Qwen3-8B 模型, 在 AIME 2024 达到 **53.33% Pass@1**, AIME 2025 达到 **41.67% Pass@1**, 平均推理 Token 数约 15k, 验证了长链深度推理的有效性。

项目经历

学术论文翻译垂直领域大模型微调

10/2025 - 12/2025

- 项目背景**: 针对学术论文翻译中专有名词不准、公式排版易碎以及 SOTA 模型 API 成本过高的痛点, 构建一款低成本、高并发且支持本地化部署的垂直领域翻译模型。
- 数据构造**: 分为 SFT 数据和 DPO 数据构造。通过 Teacher-Student 蒸馏利用 GPT-4o 生成高质量翻译对, 项目难点在于针对学术垂直领域特定的表达规范 (如公式保护、术语一致性) 设计最有效的 Prompt 模板。
- 项目难点**: 解决小参数量模型 (0.6B 级) 在偏好对齐过程中容易出现的“幻觉增强”与“灾难性遗忘”问题, 通过精细调节 Beta 系数与混合数据回填策略, 实现精度从 0.28 提升至 0.57 的显著跨越。
- 偏好对齐**: 选用 Qwen3-0.6B 为目标模型, 深入实践 DPO 与 GRPO 算法进行偏好对齐, 针对训练中的不稳定性、灾难性遗忘等问题进行了针对性调优。
- 实验结果**: BLEU:Base 模型 (0.2843) \rightarrow SFT LoRA(Rank 8, Alpha 32, **0.5092**) \rightarrow 全参数微调 (**0.5614**) \rightarrow DPO 训练 (**0.5710**); GRPO 训练表现稳定, 在 Step 100/200 分别达到 **0.5636/0.5670**。

实习经历

百融云创 | 推荐算法实习生

北京

07/2025 - 09/2025

- 业务背景**: 针对榕树贷款 App 提现环节, 精准匹配用户需求与可提现产品, 提升人均放款金额与业务增量价值。
- 开发实验**: 新增并落地 LTR 实验, 完善训练-发布-自动上线-校验闭环, 搭建天级效果分析; 人均放款金额提升 1.2%, 累计增收 10 万+;
- 特征迭代**: 基于 LGBM 与统计检验筛选高相关/高贡献特征, 精简特征集并提升模型稳定性与精度。

专业技能

- 计算机语言和技巧**: 熟悉 Python, SQL, Matlab, Latex, Linux, R
- 熟悉强化学习算法**: 熟悉 GRPO 及其变体算法 (包括 DAPO, DRPO, GFPO 等)
- 常用框架**: Pytorch, verl, TensorFlow, Scikit-learn, NumPy, Pandas, Matplotlib