

L1-书生大模型全链路开源体系

2024年9月30日 13:34



本视频介绍了书生葡语大模型的开源开放体系，包括技术发展、性能提升、模型架构、开源生态等。

LLM的各个步骤：

- 数据：提供丰富多样的开放数据
- 预训练：InternEvo支持大规模训练
- 微调：XTuner支持多种微调算法，适配多种开元生态，自动优化加速
- 部署：LM-Deploy提供搞笑的推理，可靠的量化，便捷服务，有状态的推理
- 评测：OpenCompass评测体系，国标认可，meta推荐的评测体系
- 应用：
 - 智能体Lagent框架，支持多种类型的智能体能力，多种大模型，简单易扩展的工具
 - HuixiangDou：RAG实践应用

视频要点：

- 开源开放体系涵盖数据收集、标注、训练、微调、评测、部署等全链路。
- 书生葡语大模型（英文名IN Turn LLM）多次开源，性能不断提升，达到国际先进水平。
- 模型具备强大的推理能力，上下文记忆达到百万级别，支持自主规划和搜索。
- 数据驱动，通过数据过滤、智能评估、指令生成等策略提升模型性能。
- 开源生态提供预训练框架、微调框架、部署工具、评测体系等，方便用户使用。
- 模型在推理能力、上下文记忆、自主规划等方面表现优异，可应

用于多种场景。

- 开源生态包含多种模态数据集，支持多种语言和任务类型。
- 实战营活动助力用户上手大模型开发，已有学员开发成功项目。

大模型的内容对于小白来说（比如本人），可能有点像天书，但好在有汪周谦老师，一点就透。