大语言模型高效训练方法研究 预训练成本高昂 目标多样模型庞大 研究背景 亟需精准扩展定律 导致微调消耗资源大 微调低资源 研究目标 预训练可预测 提示微调数据 超参稳定性弱 预测开销大 少参微调模块 研究挑战 性能涌现性强 干预程度小 利用效率低下 配置寻优困难 运用能力激活 实现模块自动 挖掘超参数 构建可复用 寻优与轻便框架 研究方法 与知识提示微调 扩展规律与 扩展定律与 性能扩展规律 高效退火迭代 降低数据需求 降低存算需求