大模型成为热门关键词:



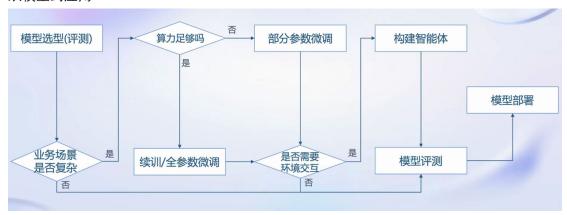
其原因归结为大模型是发展通用人工智能的重要途径:



大模型性能对比:

书生·浦语20B开源大模型性能 全面领先相近量级的开源模型(包括Llama-33B、 Llama-Llama2-能力 测试集 Llama-33B Llama2-13B以及国内主流的7B、13B开源模型) 以不足三分之一的参数量,达到Llama2-70B水平 69.75 47.73 59.55 综合考试 40.36 50.13 C-Eval (val) 31.83 41.4 59.01 58.8 37.47 AGI-Eval 22.03 30.93 37.37 33.53 33.92 40.02 BoolQ 78.75 82 42 67 87.46 84.43 86.61 87.74 TriviaQA 52.47 59.36 57.26 69.79 70.71 46.61 66.24 NaturalQuesti 20.17 24.85 16.32 25.15 30.89 33.41 34.16 31.59 29.85 68.78 14.17 34.73 43.74 CMRC 9.26 CSL 55 58.75 63.12 65.62 57.5 59.38 60 RACE (middle) 53.41 63.02 68.94 86.35 64.55 72.35 81.55 79.93 20.37 23.37 20.55 19.91 25.38 XSum 25.23 66.85 69.38 69.77 WinoGrande 64.64 64.01 67.32 69.38 ВВН 37.93 49.98 45.62 48.98 58.38 64.91 63.31 54.44 82.15 20.32 79.71 29.57 79.76 78.07 PIQA 14.02 18.9 17.07 25.61 17.68 18.9 26.22 HumanEval 编程 MBPP 20.6 26.8 30.8 28.4 39.6

从模型到应用:



在模型部署方面,主要考虑节省资源以及提升用户使用量。

Pre-trained



微调:





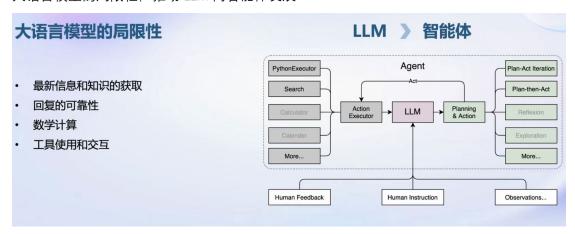
训练方案支持英伟达 20 系列以上显卡,显存 8G 可实现微调 70 亿参数模型。 评测平台架构:

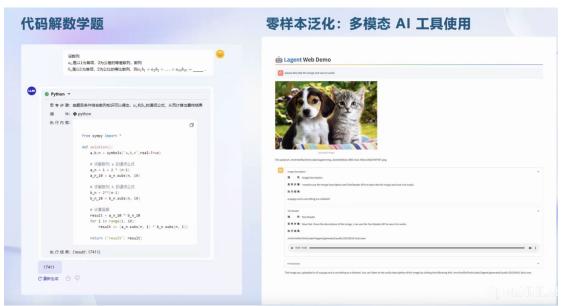


部署:



大语言模型的局限性,推动LLM 向智能体发展:





AgentLego(工具 collection):



总结:



开源大模型开放学习体系:数据-预训练-微调-部署-评测-应用