

Llama3 技术报告

advanceflow

日期: April 19, 2024

摘 要

1 摘要

2 预训练

3 监督式微调 (SFT)

4 拒绝采样

5 近端策略优化 (PPO)

6 直接策略优化 (DPO)

7 微调

8 实验

9 结论

参考文献