学习周报

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 姓名 | | 郝丹萌 | 时间范围 | 2024.11.25-2024.12.01 |
| 周次 | | 第七周 | 研究方向 | 大模型高效参数微调 |
| 本周完成工作 | 1. 看与模型微调相关的课程。 2. 继续学习了与大模型优化内容相关的课程，有关大规模预训练模型的优化技术，如梯度裁剪、学习率调度和分布式训练策略。 3. 阅读AdaLoRA 的论文与源码。 4. 阅读并理解了 AdaLoRA 的相关论文及其实现代码。 5. 在少量数据集上实现了 AdaLoRA 微调。 6. 学习学习率、批次大小、训练周期等超参数对训练效果的影响。 | | | |
| 本周  问题汇报 | 1. 在实现 AdaLoRA 微调时，部分超参数的选择仍存在不确定性，特别是在小数据集上的表现不如预期。需要进一步探索适用于该任务的超参数配置。 2. 在与 QLora 的对比实验中，虽然 AdaLoRA 在某些指标上略有优势，但收敛速度和稳定性仍有提升空间，超参数的调整需要更加细致，以达到更好的平衡。 3. 对 AdaLoRA 的一些实现细节还不完全理解，计划通过进一步分析代码和论文，解决这些问题。 | | | |
| 下周工作计划 | 1、看模型微调的课程。  2、学习prefix tuning和quantization相关知识。 | | | |

学习内容笔记

机器学习相关笔记、机器学习相关代码练习：[Meng592/LearningNotes at master](https://github.com/Meng592/LearningNotes/tree/master)

