学习周报

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 姓名 | | 郝丹萌 | 时间范围 | 2024.11.18-2024.11.24 |
| 周次 | | 第六周 | 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. 学习神经网络和深度学习的基础课程，重点学习了神经网络优化算法、损失函数的选择以及训练过程中的常见问题与解决策略。通过对优化算法的学习，加深了对大模型训练细节的理解。   3、学习并阅读了 QLora 的相关论文及其源代码，理解了其在大模型微调中的优势及实现方法。基于对 QLora 原理的理解，在少量数据上进行微调实验，初步观察到训练过程中模型效果的提升。 | | | |
| 本周  问题汇报 | 1. 在微调 QLora 时，尚不清楚如何选择合适的超参数来提高小数据集上的训练效果，尚需进一步实验。 2. 在对比 QLora 微调的训练效果时，发现训练过程中的损失函数下降较为平缓，是否是超参数设置问题，需要调整并重新实验验证。 | | | |
| 下周工作计划 | 1. 继续听基础课。 2. 继续深入学习基础课程，特别是与大模型优化内容相关的课程，进一步完善对模型优化技术的理解。 3. 阅读AdaLoRA论文和源代码，研究 AdaLoRA 的原理和实现，特别关注 AdaLoRA 在大语言模型微调中的应用，理解其如何利用自适应学习率和参数选择来优化训练过程。 4. 在少量数据集上实现 AdaLoRA 微调。 5. 研究 AdaLoRA 的超参数设置，特别是如何在不同任务中调整超参数，以提高微调效果。 | | | |

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)

