学习周报

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 姓名 | | 郝丹萌 | 时间范围 | 2024.11.11-2024.11.17 |
| 周次 | | 第五周 | 研究方向 | 大模型高效参数微调 |
| 本周完成工作 | 1. 看神经网络和深度学习等基础课程。 2. 阅读nanoGPT的源代码，并在少量的数据上训练和微调。 3. 论文阅读《Open Large-scale Language Models》。 | | | |
| 本周  问题汇报 | 1. 不理解nanoGPT代码中权重共享，共享权重训练后效果不会变差么。 2. checkpoint和预训练权重的区别和联系。 3. 为什么前馈网络使用了两层线性变换，而不是一层。 | | | |
| 下周工作计划 | 1. 继续听基础课。 2. 阅读Baichuan-7B的源代码。  3、看QLora相关论文、代码，进行微调。 | | | |

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)

