计算机学院毕业论文（设计）开题答辩记录表

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 学生姓名 | | 蔡灯 | | 学 号 | | | U201614739 | | | 专 | 业 | 计算机科学与技术 | | 指导教师姓名 | 李瑞轩 |
| 指导教师职称 | | | 教授 | |  | 论文(设计)题目 | | | 不同通信模式下分布式深度学习模型训练方法效率评估 | | | | | | |
| 答辩组成员 | 姓名 | | | | 职称 | | | 开题答辩日期 | | | | | 3月9日 | | |
| 李玉华 | | | | 副教授 | | | 记录人 | | | | | 朱浩 | | |
| 辜希武 | | | | 讲师 | | |  | | | | |  | | |
|  | | | |  | | |  | | | | |  | | |
| 答辩记录内容  一、论文设计要点   1. 使用卷积神经网络在不同数据集上训练； 2. 使用经典优化算法SGD的同步，异步，限制异步频率的部分同步，补偿梯度延迟的异步变体； 3. 使用parameter-server通信架构；   二、提问与回答   1. 讲讲你所使用的优化算法的优缺点？   答：使用同步SGD训练模型，模型准确率最高，而且也稳定，在理论上和单机上训练模型等价，但是训练时间长；使用异步SGD,效果恰恰相反，模型准确率低，不稳定容易震荡，但是训练时间最短；限制异步频率的部分同步SGD，效果则是这两种算法的折中；补偿梯度延迟的异步SGD，是在异步的基础上降低了梯度延迟的误差，使得模型准确率获得提高的同时还没有增加多少计算力和存储代价。   1. 介绍一下你所使用通信架构parameter-server?   答：Parameter-server架构（PS架构）是分布式深度学习使用最频繁的通信架构。在PS架构中，分布式集群将节点分为两类：计算节点，参数服务节点。计算节点训练本地所分配到的数据集子集，并计算出相应的本地局部模型参数梯度，并发送给参数服务节点；参数服务节点根据从各个节点收集到的局部模型参数梯度，更新全局模型参数，同时将全局模型参数广播发送给各个节点。  三、存在的问题与建议  答：目前最主要的问题实验的环境存在诸多限制，搭建合适的分布式环境，使用什么样的模型训练。结合现有的资源，只能在本人的8G笔记本上使用虚拟机搭建几台虚拟节点，每个节点的内存不超过2G，这使得在复杂的数据集上可能无法选择高精度高复杂的模型。 | | | | | | | | | | | | | | | |