9月 28号晚,联邦学习 FATE 开源社区第四期圆桌会完美落幕。作为首期面对社区用户全面开放的圆桌会,这期的报名获得空前热烈的支持。即将迈入 1.5 稳定版本的 FATE,也将继续探索更多的新功能。如作为构建联邦学习数据合作网络的基础设施——FATE Cloud,这也是本期圆桌会重点讨论项目。接下来,通过本文我们一起来回顾这期圆桌会吧。

以下为圆桌会上的精选问答

FATE cloud 具体什么时候开放?

十月下旬。

会议会分享 FedRec 的内容吗?

以前分享过,后面可以安排,欢迎找小助手要一下材料和回放。

组件任务重跑是咋个控制的呀; dag 执行流程全部落入存储,按计算轨迹重新计算吗?

训练的任务执行状态是落数据库的,如迭代到了 50 轮时组件 fail 了,重新启动这个组件即可。

新的 DSL 和旧的 dsl 差别,以及是否有算子时间、传输统计的样例?

旧的 DSL 在训练任务执行运行完成后生成预测 dsl,而新版本 dsl 需要制定预测组件。不限定生成次数,方便试错。旧版本里只能一次,因为是自动生成。

另外新的 DSL 有统一的风格,比如说配置 role_parameter\algorithm_parameter 旧版风格不一致,在新的 DSL 里面是 ok 的。

同时旧的 DSL 不支持预测阶段接入新组件(自动生成),而新的 DSL 是支持的。如极端例子,lr 模型预测的时候,可以增加新lr 训练模型和接入评估组件等。

算子时间方面,会通过算子操作级别来打印时间。

1.5 版本支持 tensorflow-gpu 吗?

TF 本身具备这样的能力,FATE 底层可以直接复用,但需要在接口层适配,测试充分后 FATE-1.5 正式版会支持该功能。

fate cloud 如何指定需要安装的主机(分开组件部署)?

可以对 k8s 集群的 node 节点打 label,通过在 nodeselector 里配置 label 信息,可以把组件部署到指定机器上。

组件重跑和组建多输出,支持上一个 job 的组件结果倒入下一个 job 作为输入吗?例如第一个 job 只有 psi,将其结果倒出,并作为下一个 job 的输入,下一个 job 直接是 secureboost 开头。

1.5 版本暂不支持, 但可以期待后续版本。

新旧 DSL 能兼容么?

兼容的,但可能提交任务的时候发现以前配错或多余的参数,在新版本里面会捕捉这些异常。我们通过 DSL_version 1 或 2 来区别,默认是 1,支持旧版本 DSL。如果跑在 spark 上需要用 DSL_version 2,因为涉及到一个数据转换组件。

停留在 1.5 可能会持续 2 年左右吗?

1.5 是一个长期维护的稳定版本,比如说未来即使更新到 2.0,我们还是会往 1.5 里面添加一些小的 feature,我们还是会持续推进新的版本迭代。

现在支持 CNN, RNN 训练非结构化数据么?

暂不支持,目前我们保留了数据转换接口,未来对图像或其他数据处理能力做了预备基础。目前来说版本开发内容较多,这块放在了次优先级。

hetero_nn 多分类预测结果只有一个,是什么原因呢(数据没有问题)?

测试用例有覆盖这个场景,但尚未发现此问题,欢迎在 github 上给我们提个 issus,把数据放上去一起研究,看看是否配置错误等其他方面原因。

Python 版 sdk 具备哪些功能?

fateflow cli v2 的功能都具备。

预热版本会释放全部功能么?

基本会释放属于 FATE 自身的所有新功能,但正式版 FATE 底层存储计算引擎 eggroll 新版本 2.2 会搭载发布,另外 preview 预热版大家使用过程中出现的问题也会在正式版进行修复。

手动 kill(cacel)的任务也可以重跑么?

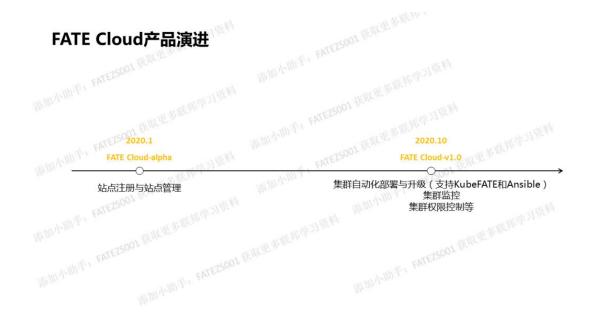
可以的, 基本所有失败任务都可以重跑的。

算子传输的带宽有统计么?如何统计与其他 party 联合执行训练的带宽?

目前是没有的,FATE 去统计一个分布式表 table,会带来性能损耗。但在其他底层如 spark、eggroll 去操作会好一点。另外其他第三方工具去统计流量出口也是可行的。

以下为 FATE Cloud 介绍, PPT 的部分内容:

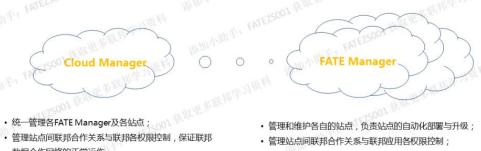
版面有限,有兴趣的朋友欢迎联系小助手(FATEZS001)获取完整内容。



FATE Cloud简介

FATE Cloud是构建和管理联邦数据合作网络的基础设施,为跨机构间、机构内部不同组织间提供了安全可靠、合规的数据 合作网络构建解决方案,构建工业级别的联邦学习云服务,实现多客户端(联邦站点)的云端管理。

FATE Cloud由负责联邦站点管理的云管理端Cloud Manager和站点客户端管理端FATE Manager组成



- 管理站点间联邦合作关系与联邦各权限控制,保证联邦 数据合作网络的正常运作;
- 管理和维护各自的站点,负责站点的自动化部署与升级;
- 管理站点间联邦合作关系与联邦应用各权限控制;

FATE Cloud功能概览



获取会议 PPT, 或对圆桌会还有别的疑问?欢迎联系 FATE 开源社区助手获得帮助。

原文链接: https://mp.weixin.qq.com/s/zkh1BcFNzjExinpst1j5Qg