ICS 35.240.50
CCS L67

T/SIGA

团　　体　　标　　准

T/SIGA 003—2024

智能制造联邦学习应用指南

Guidance for federated learning application of intelligent manufacturing

2024-08-27 发布　　2024-08-28 实施

上海市图像图形学学会　发布

目　次

前　言

本文件依据 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由上海市图像图形学学会归口。

本文件起草单位：上海工程技术大学、东华大学、上海富数科技有限公司、公安部第三研究所、上海交通大学、上海保隆汽车科技股份有限公司、晋拓科技股份有限公司。

本文件主要起草人：赵晓丽、方志军、罗光圣、王国中、李高健、卞阳、杨天雅、李倩、姚晨、周军、黄军林、吴天磊、张东、韩乔。

智能制造联邦学习应用指南

1 范围

本文件给出了面向智能制造的联邦学习架构、分类、参与方角色、实现过程和软件系统等应用指南。

本文件适用于智能制造的联邦学习系统部署、开发和应用。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

下列术语和定义适用于本文件。

3.1

联邦学习 federated learning

一种分布式机器学习技术，其核心思想是通过在多个拥有本地数据的数据源之间进行分布式模型训练，在不需要交换本地个体或样本数据的前提下，仅通过交换模型参数或中间结果的方式，构建基于虚拟融合数据下的全局模型，从而实现数据隐私保护和数据共享计算的平衡，即“数据可用不可见”“数据不动模型动”的应用新模式。

[来源：IEEE P3652.1—2020，1.1，有修改]

3.2

联邦簇 federated cluster

以相同数据方案在资产数据上训练的模型的所有学习任务。

3.3

联邦学习任务 federated learning task

需要用算法训练的机器学习模型，简称任务。

3.4

联邦学习计划 federated learning program

对应一个联邦学习任务，表示联邦学习服务器和相关联邦学习客户端的联邦执行指令。

3.5

联邦学习软件系统 federated learning system

一套完整的、严密的用于实现联邦学习功能的软件。

3.6

元数据 metadata

数据集的字段信息，如字段名、字段类型、字段信息等。

3.7

局部模型 local model

各个客户端的模型。

3.8

全局模型 global model

各个客户端联邦聚合后的模型。

3.9

参与方 participant

一个或一组自然人或法人，指参与联邦学习互联互通任务的主体，包含发起方、数据方、算法方、计算方、结果方、协调方等多种角色。

3.10

算法组件 algorithm component

用于执行计算任务的一种可代替、可组合的部件，封装了算法功能的实现并提供一系列可用的接口。

3.11

同构 homogeneity

多个相同类型的样本参与完成一件事情。

3.12

异构 heterogeneity

多个不同类型的样本参与完成一件事情。

3.13

资产 assets

数据的来源，包括各种传感器、加工设备、机器人等。

3.14

客户端 end

位于联邦学习模型的最底层，能执行简单的计算来训练和评价模型，简称端。

3.15

诚实客户端 honest end

遵循协议规定、诚实地提供数据或执行计算任务的客户端。

3.16

半诚实客户端 semi—honest end

按照协议规定进行操作，但同时会尝试从其他方的输入或计算过程中获取更多信息的客户端。

3.17

恶意客户端 malicious end

试图挖掘其他参与方隐私信息的客户端。

3.18

边 edge

边缘设备，位于联邦学习模型的中间，常边缘设备比客户端有更强的计算能力，一般位于工厂的不同部门或不同的工厂。

3.19

云 cloud

中心服务器，位于联邦学习模型的最高层，具有更大的计算能力。

3.20

键值 key-value

根据关键字取值，具有极高的并发读写性能。

3.21

节点 node

互联互通网络的基本组成单元，对外提供交互接口。

4 缩略语

下列缩略语适用于本文件。

API（application programming interface）：应用程序接口

HMAC（hash—based message authentication code）：密钥相关的哈希运算消息认证码

HTTPS（hyper text Transfer protocol over secure socket layer）：超文本安全传输协议

RPC（remote procedure call）：远程过程调用

TCP（transmission control protocol）：传输控制协议

TLS（transport layer security）：传输层安全协议

UDP （user datagram protocol）：用户数据报协议

VPN（virtual private network）：虚拟专用网络

XML（extensible markup language）：可扩展标记语言

QPS（Queries per second）：每秒查询率

5 架构

5.1 智能制造联邦学习框架是分层联邦学习架构。第一层联邦学习的对象是针对相同设备产生的数据，这些数据具有相同的特征和模态，也就是同构数据；第二层联邦学习是针对不同部门或不同工厂边缘设备的数据，这些数据具有不同的特征和模态，也就是异构数据。

5.2 整个联邦学习模型如图 1 所示。模型基于工业互联网架构，由云边端组成，即客户端、边缘设备和可信第三方服务器。

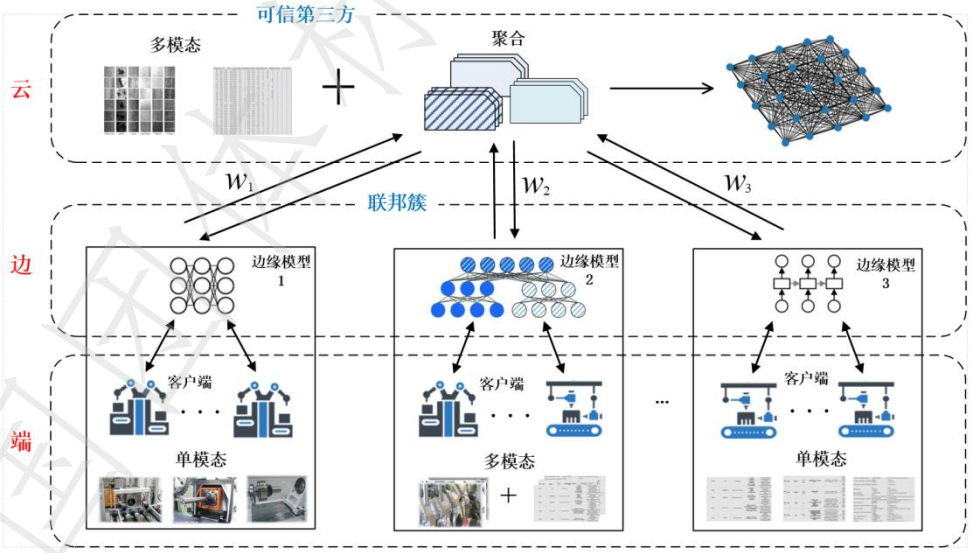

图 1 智能制造联邦学习架构

5.3 客户端（端）可以是设备、车间。位于分层联邦学习模型的最底层，能执行简单的计算来训练和评价模型。数据、模态同构的客户端组成联邦簇。簇内的客户端是同构的，可以减小训练的复杂度。

5.4 边缘设备（边）位于分层联邦学习模型的中间，是第一层联邦学习的聚合中心，不同的边缘设备间数据、模态异构。通常边缘设备比客户端计算能力更强，位于工厂的不同部门或不同的工厂。

5.5 可信第三方服务器（云）可以是功能强大的参与方，也可以不是参与方，其位于分层联邦学习模型的最高层，是第二层联邦学习的聚合中心，计算能力强，可以进行大模型训练，该层进行异构联邦学习。边缘设备同客户端相比，数量减少，降低了异构训练的难度。中心服务器进行大模型训练，泛化表

示能力强。

6 分类

6.1 横向联邦学习

客户端数据的样本特征重叠较多而样本重叠较少的情况下，取出双方用户特征相同而用户不完全相同的那部分数据进行训练。横向联邦学习宜用于拥有样本较少的数据方需要训练高质量模型的场景，以获得优于本地训练的效果。

6.2 纵向联邦学习

客户端数据的样本重叠较多而样本特征重叠较少的情况下，取出双方样本相同而用户特征不完全相同的那部分数据进行训练。纵向联邦学习宜用于拥有特征或标签不充分的数据方需要训练高质量模型的场景，以获得优于本地训练的效果。

6.3 联邦迁移学习

客户端数据的样本与样本特征重叠都较少的情况下，宜利用迁移学习来克服样本和标签不足的情况。联邦迁移学习主要用于参与联邦学习的客户端间样本和样本特征均重叠较少，且又需要互相协同获取高质量模型的场景，宜用于跨域联邦学习场景。

7 参与方角色

7.1 任务发起方

任务发起方是触发联邦学习任务的角色，并在任务执行前核实各方资源，其通常是计算方、调度方或结果使用方。

7.2 任务调度方

任务调度方负责联邦学习任务的分发和计算过程中各方资源和行为的协调。其根据联邦学习计划调度多个参与角色完成联邦学习任务，调度方宜获得可信第三方的认证，确保调度方可信性、安全性与正确性。调度方可由任务发起方兼任。

7.3 算法提供方

算法提供方根据联邦学习任务需求，提供联邦学习算法。其提供的算法以算法组件和算法描述的形式输出。算法组件包括计算逻辑、算法实现和算法参数；算法描述包括算法的版本号、接口说明、安全性说明、兼容性说明等；兼容性说明描述可以运行的计算环境、依赖的底层算子、适用的场景和兼容的版本信息等。

7.4 数据提供方

数据提供方提供联邦学习需要的元数据。对于联合建模和预测任务，提供模型所需的训练数据以及服务预测数据。

7.5 计算方

计算方为执行联邦学习任务提供计算能力。计算方自身算力不足时，可设立辅助计算方的角色进行辅助计算。计算方执行与其他参与方的协同计算，保证任务执行过程不会造成隐私数据泄露。

7.6 结果使用方

结果使用方是最终获得联邦学习任务计算结果。对于联合建模任务，获得最终的模型参数；对于联合预测任务，获得预测结果；对于联合分析任务，获得分析的结果。只有结果使用方才能获取计算的最终结果，其它非结果使用方不能获得结果，也不能通过中间计算过程，计算或推断出最终结果。

8 实现过程

8.1 总体流程

见附录 A。

8.2 任务创建

任务发起方利用调度方发起联邦学习任务。

8.3 任务分配

调度方将任务下发到每个数据提供方的计算节点，当某个节点有辅助计算节点时，在任务运行过程中会将某个计算逻辑下发到辅助计算节点。

8.4 数据准备

在每个工厂或生产线内部预处理本地数据，包括但不限于数据清洗、格式转换和特征提取，以准备好可用于模型训练的数据集。

8.5 联邦学习初始化

8.5.1 配置协议

定义通信协议和通信传输信息加密方式。

8.5.2 模型初始化

中央服务器（或协调服务器）初始化一个全局模型，并将其发送到各个工厂或生产线的本地设备。

8.6 本地模型训练

各个工厂或生产线接收到全局模型后，使用本地数据进行模型训练。本地数据不会离开各自的工厂或生产线，每个本地节点根据其数据执行多轮梯度下降或其他优化算法，以更新模型参数。

8.7 模型间通信

8.7.1 模型参数上传

各个工厂或生产线将本地训练得到的模型参数(梯度或权重更新)或知识加密后上传到中央服务器。

8.7.2 隐私保护

在上传过程中，使用差分隐私、同态加密等技术保护数据隐私，确保上传的参数不泄露敏感信息。

8.8 全局模型聚合

8.8.1 参数聚合

中央服务器接收来自各个工厂或生产线的加密模型参数，使用联邦平均或其他聚合算法更新全局模型。

8.8.2 更新全局模型

将聚合后的全局模型参数更新到全局模型中，并准备分发给各个本地节点。

8.9 模型分发与迭代

8.9.1 模型分发

中央服务器将更新后的全局模型分发给各个工厂或生产线，开始下一轮的本地模型训练。

8.9.2 迭代训练

上述本地训练、模型更新和全局聚合过程重复进行多轮，直至全局模型收敛或达到预设的训练轮次。

8.10 模型验证与部署

8.10.1 模型验证

在训练过程中和结束后，验证全局模型的性能，保证其在各个工厂或生产线的生产数据上都能取得良好的效果。

8.10.2 模型部署

将最终的全局模型部署到各个工厂或生产线，应用于实际生产过程中的预测、优化和决策等工作。

9 软件系统

9.1 框架

9.1.1 软件系统包括但不限于接入层模块、基础技术模块、计算模块、管理模块，其框架参见图 2。

9.1.2 接入层是参与方接入系统需要的各种协议以及连接器，实现网络互联。

9.1.3 基础技术模块为上层联邦学习计算模块提供系统通用的、维持系统正常运作的支撑性技术服务。

9.1.4 计算模块是实际完成联邦学习计算和交互的模块，包括但不限于联邦数据探查、联邦建模、联邦推理。

9.1.5 管理模块是为联邦学习系统提供统一管理能力，包括但不限于项目管理、模型管理、用户管理、运维管理、结果管理、参与方管理、算法管理、数据管理。

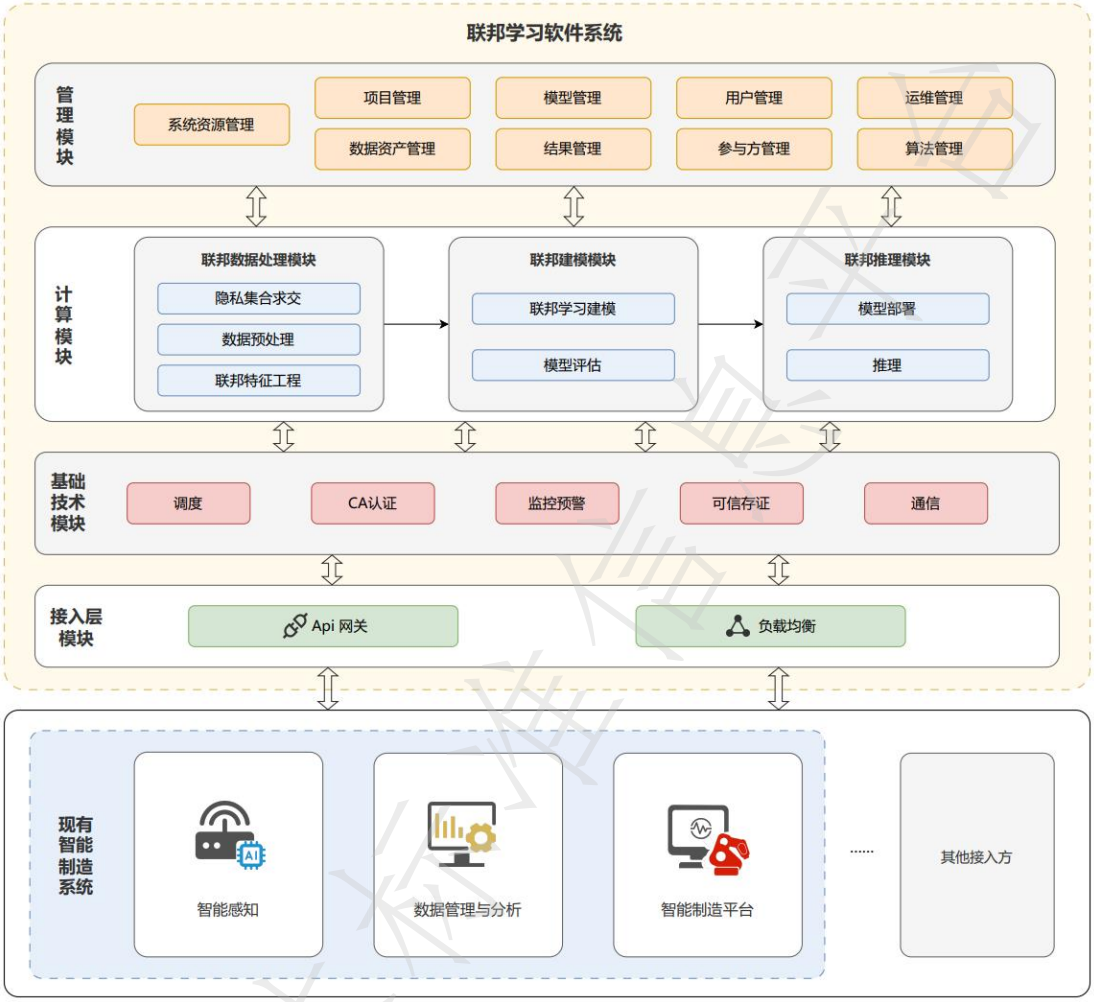

图 2 联邦学习软件系统

9.2 和现有智能制造系统对接

9.2.1 现有智能制造系统相对成熟，联邦学习系统需要和现有的制造系统做好对接，以发挥联邦学习系统的作用。系统间的对接如图 2 下方所示，需在数据类型和传输协议上做好协同。

9.2.2 联邦学习系统宜支持不同类型的数据提供方，涵盖目前智能制造系统通用的数据类型，包括但不限于数据库和文件。数据库类型包括但不限于 MySQL、Oracle、Hive、HBase 等；文件类型包括但不限于 txt、csv、xml、key-value。

9.2.3 数据交互方式宜开发对应系统的对外接口来实现，开发的接口宜基于 webAPI 方式，以避免原系统直接与目标系统的数据层进行交互。既保证系统的独立性和安全性，又实现不同系统间的数据交互。

9.2.4 系统间的传输宜采用 https 方式，保证信息传输的安全性，宜采用成熟的加密方式。

9.3 功能建议

9.3.1 身份认证

身份认证宜满足以下要求：

a) 对隐私计算过程中的关键环节进行身份认证，保证操作方身份及行为的合法性和抗抵赖性；

b) 使用HMAC、数字签名、动态口令等技术对隐私计算系统中各参与方进行身份认证与管理；

c）各参与方之间通信时进行身份认证，保证身份合法性；

d）任务发起方对任务请求信息进行签名，任务调度方对身份进行审核，并向任务涉及的相关方请求任务审核。当身份和任务审核都通过时才允许任务发起方创建任务；

e）结果方请求计算结果时，任务调度方进行身份认证和权限鉴别；

f）采用两种或两种以上组合的认证方式实现身份认证，且其中至少一种认证方式宜使用密码技术来实现，身份认证方式包括但不限于口令、证书、令牌、短信验证码、邮箱验证码、生物识别；

g）身份认证信息具有复杂度要求并定期更换；

h）采用密码技术保证身份认证信息在传输过程中的完整性和保密性。

9.3.2 授权访问

授权访问宜满足以下要求：

a）系统能对各参与方进行相应的权限设置和访问控制，避免出现信息泄露或操作风险；

b）按照最小安全访问原则设置访问控制权限；

c）支持基于用户账户和权限分配的细粒度访问控制，支持仅授权用户才能访问特定资源；

d）对数据提供方的数据使用进行控制，防止数据在数据提供方非授权的情况下被使用；

e）任务调度方宜对未被授权的计算请求协调发起数据使用授权申请，申请内容包含授权主体、授权条件、授权内容、数据使用范围等。数据提供方同意后向使用方发送授权，用于后续计算时的权限认证；

f）数据提供方对每个任务请求验证其数据使用授权的合法性，包括但不限于授权是否有效、数据使用范围和使用期限是否合理，进而保证访问者是否具备操作该对象的权限；

g）保证数据是在数据方授权的情况下被使用；

h）支持根据用户身份、行为、环境，以及安全监测情况等综合因素，动态调整用户的访问权限；

i）数据提供方具备取消数据使用授权的权限;

j）采用密码技术保证访问控制信息的完整性；

k）采用密码技术保证重要信息资源安全标记的完整性。

9.3.3 存证审计

存证审计宜满足以下要求：

a）具备对各参与方的用户操作日志和结果存证的审计能力，对于违背约定的数据提供方、计算方和结果使用方能通过存证、审计等方法进行发现、追踪；

b）对隐私计算全过程中各参与方的操作行为、模型参数、状态信息、计算过程、数据的访问记录等信息进行存证；

c）对数据提供方和结果使用方的计算结果和信息进行存证和记录，保证信息安全性与结果可追溯性；

d）采用密码技术保证日志及存证信息的完整性，并提供查询功能，能够对重要及异常操作进行回溯和审计；

e）将区块链技术用于辅助存证审计，实现数据全节点可见，链上数据可追溯、不可篡改；

f）使用核准的密码技术对存证内容和审计记录进行保护，防止非授权的篡改和窃取。

9.3.4 联邦计算

联邦计算宜满足以下要求：

a）数据计算既能处理结构化数据也能处理非结构化数据；

b）数据提供方将传输的信息通过密码变换转换为密态数据，提供给计算方；

c）具备数据格式转换、精度处理等功能；

d）提供多种联邦计算能力，包括但不限于数据清洗、特征选择等特征工程，常用的联邦监督学习算法，常用的联邦非监督学习算法，联邦深度学习算法等；

e）支持联邦机器学习模型的多种评价指标的计算，包括但不限于损失（loss）、曲线下面积（AUC, area under curve）值、准确率（accuracy）、F1 得分（F1 score）。评价指标的具体计算公式可参见附录 B；

f）支持联邦模型在线、离线推理和批量推理；

g）支持联邦算法在线编写、修改、调试、提交等。

9.3.5 用户加入和退出机制

用户加入和退出机制宜满足以下要求：

a）联邦学习客户端包括诚实、半诚实和恶意客户端，在用户加入时具有判别机制，拒绝恶意客户端加入；

b）对于诚实和半诚实客户端，系统能够计算联邦客户端的贡献度，并根据贡献度给予相应的激励，提高联邦参与者的积极性；

c）对于一段时间内没有贡献或拖慢整个联邦学习系统的客户端，给出预警、限期整改和强制退出等不同类别的警戒措施。

9.4 非功能建议

9.4.1 系统安全性

9.4.1.1 数据安全性

数据安全性宜确保数据全生命周期的安全性：

a）在数据进入联邦学习系统前对法律法规禁止、未经个人信息主体授权，与当前业务场景无关、无法利用的数据等进行数据去除；

b）对法律法规禁止的输入数据进行匿名化处理，使个人信息主体无法被识别或关联；

c）具备接收数据访问权限申请、对权限进行授权和管理的功能，保证流入数据在数据提供方授权后被访问，过程数据和结果数据的访问权限须符合参与方约定；

d）对原始数据、中间数据及结果数据进行分层存储，保证原始数据与结果数据之间的独立性；

e）对数据的摘要等数据、数据关键操作、操作主体进行存证，保证具有追溯能力。存证数据本身需要进行脱敏、加密等处理，对存证数据约定的存储时限符合法律法规的要求；

f）各参与方遵循最小化原则存储过程数据，在保证可追溯性的前提下，在数据隐私计算完成后采用覆写法等方式删除非必要的过程数据；

g）保证敏感信息已被消除，无法通过联邦学习处理后的数据进行推断、重建、还原。保证在采用所有合理范围内可能使用的重识别方法时，处理后的数据在不结合额外信息的情况下无法重标识到个人。

9.4.1.2 计算安全性

计算安全性宜确保在计算过程中安全性：

a) 确保各参与方的原始数据不出本地，隐私数据不可被逆推、窃取或篡改；

b）保证计算结果只被约定的结果使用方获知，不被其他参与方知晓，保障结果隐私性；

c）采用同态加密、多方安全计算、差分隐私、可信硬件等安全技术保护中间数据，防止从中间数据逆推出原始数据及隐私参数。所使用的密码方案具有可证明安全性，且能够较全面地抵抗已知攻击；

d）在有第三方的架构中，保证第三方未经授权不可获知或解密出模型参数明文、可逆推出模型参数或用户隐私数据的中间参数明文；

e）采用杂凑函数、分组加密、公钥加密、数字签名等国家密码管理部门批准的商用密码算法标准。采用密钥长度及密钥管理方式等符合国家密码管理部门与行业主管部门要求。

9.4.1.3 传输安全性

传输安全性宜保证传输通道和通信网络的安全性：

a）使用SSL3.2/TLS1.1以上的安全通信机制，具体版本和加密套件由各参与方协商确定；

b）传输过程中使用的签名算法、杂凑算法、密钥协商算法符合国家密码主管部门的要求，采用国家商用密码算法：SM2、SM3、SM4等；

c）采用HTTPS等安全传输协议，若通过运营商进行数据传输，采用VPN或者专线等技术保证传输通道的安全性；

d）各参与方采取适当的技术和管理措施保证传输过程安全。可采用手段包括但不限于备份存储、二次加密、一次一密、前置机技术、多租户隔离；

e）一般公网、内网场景中使用TCP协议进行传输。对于传输效率有更高要求的，使用UDP协议进行传输，保证基于UDP协议可以在复杂网络环境之中稳定运行。

9.4.2 系统性能

系统性能宜保证计算满足一定的性能：

a）在安全合规的前提下，支持完成海量数据的处理、计算；

注：海量数据指支持十亿级数据的安全求交，百亿级数据的匿踪查询，亿级数据的多方安全计算，万级数据、百级特征的联邦学习建模，亿级数据的联邦学习在线推理。

b）模型推理的耗时、吞吐量等分别满足实时和非实时业务场景要求；

c）对数据的计算和通信耗时符合各参与方预期，同时支持并行计算和分布式计算。

9.4.3 易用性

易用性宜满足以下要求：

a）联邦学习平台能对平台运行时使用的资源情况进行监控告警，如内存、中央处理器（CPU）使用率、网络等；

b）能对任务运行过程中的详细日志能进行查看和下载；

c）能对平台的网络通信量进行观测，如出入包的大小等；

d）能对平台的负载进行监控，如服务调用QPS；

e）能对计算流程进行监控，如模型指标等；

f）联邦学习平台需要具备完善的部署和操作说明文档，方便用户理解、部署、接入和使用。

g）能支持可视化拖拉拽操作，包括数据管理、任务管理、模型管理等交互操作界面，降低用户学习和使用的成本；

h）能提供算法开发框架，方便用户可以基于框架开发满足安全约束条件的自定义算法开发；

i）能提供良好的应用设计，减少升级时对用户的打扰，最小化用户升级成本；

j）能支持容器化部署，方便用户快速部署；

k）在保持系统基础设施稳定和改动量最小的情况下可灵活实现集群资源的弹性扩展，如：资源层面的弹性增减，如CPU、存储设备、输入/输出（I/O）设备等；系统层面，调度系统、计算引擎、数据服务均可弹性扩展；

l）具备鲁棒性，如断点续跑等能力，在模型训练执行意外失败时，可基于断点，继续运行；

m）具备主备、多活等高可用机制，单节点出现故障时可自动隔离及切换，对外服务不存在单点。

附　录 A
（资料性）
智能制造联邦学习实现过程流程图

图 A.1 给出了一种智能制造联邦学习实现过程流程图。

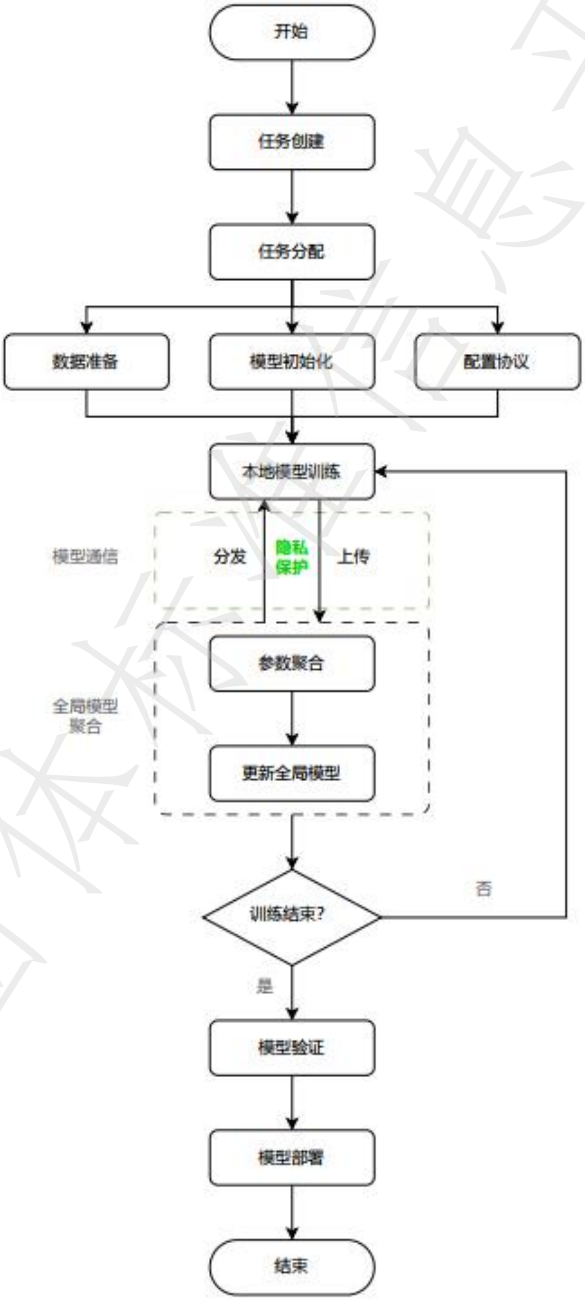

图 A.1

附　录 B
（资料性）
评价指标计算公式

B.1　损失（loss）

交叉熵损失，可按式（B.1）计算。

$$L = -[y \log \hat{y} + (1-y)\log(1-\hat{y})] \quad \text{（B.1）}$$

式中：

L — 交叉熵损失；

y— 真实标签的独热编码；

$\hat{y}$ — 模型预测的概率分布。

B.2　曲线下面积（AUC, Area Under Curve）

可按式（B.2）计算。

$$AUC = \frac{1}{2}\sum_{i=1}^{m-1}(x_{i+1} - x_i)\cdot(y_i + y_{i+1}) \quad \text{（B.2）}$$

式中：

AUC — 曲线下面积；

m — 样本点总数；

x_i — 样本点的假正例率；

y_i — 样本点的真正利率。

B.3　准确率（accuracy）

可按式（B.3）计算。

$$acc(f;D) = \frac{1}{m}\sum_{i=1}^{m}\mathrm{II}[f(x_i) = y_i] \quad \text{（B.3）}$$

式中：

$acc（f;D）$ — 准确率；

$II（\cdot）$ — 指示函数；

x_i — 样本；

y_i — 样本的真实标签；

$f(\cdot)$ — 学习器；

D — 数据集。

B.4 F1 得分（F1 score）

可先按式（B.4）和（B.5）计算查准率和召回率，再按式（B.6）计算 F1 得分。

$$precision = \frac{TP}{TP + FP} \quad \text{（B.4）}$$

式中：

precision — 查准率；

TP — 真正例；

FP — 假正例。

$$recall = \frac{TP}{TP + FN} \quad \text{（B.5）}$$

式中：

recall — 召回率；

TP — 真正例；

FN — 假反例。

$$F1 = \frac{2 \cdot precision \cdot recall}{precision + recall} = \frac{2 \cdot TP}{m + TP - TN} \quad \text{（B.6）}$$

式中：

*F*1 — F1 得分；

precision — 查准率；

recall — 召回率；

TP — 真正例；

TN — 真反例。

参 考 文 献

[1] YD/T 4691—2024 隐私计算 联邦学习产品安全要求和测试方法

[2] IEEE P3652.1—2020 IEEE Guide for Architectural Framework and Application of Federated Machine Learning