(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 114330759 B
(45)授权公告日 2022.08.02

(21)申请号 202210217697.2
(22)申请日 2022.03.08
(65)同一申请的已公布的文献号
申请公布号 CN 114330759 A
(43)申请公布日 2022.04.12
(73)专利权人 富算科技(上海)有限公司
地址 200135 上海市浦东新区中国(上海)自由贸易试验区浦东大道1200号2层A区
(72)发明人 尤志强 卞阳
(74)专利代理机构 北京超凡宏宇专利代理事务所(特殊普通合伙) 11463
专利代理师 蒋姗
(51)Int.Cl.
G06N 20/20(2019.01)
G06F 21/60(2013.01)
(56)对比文件
CN 111242316 A,2020.06.05
CN 112906912 A,2021.06.04
CN 111340247 A,2020.06.26
CN 112818374 A,2021.05.18
CN 112383396 A,2021.02.19
CN 111553470 A,2020.08.18
CN 113194126 A,2021.07.30
CN 113821313 A,2021.12.21
CN 114091103 A,2022.02.25
CN 113570069 A,2021.10.29
CN 113779608 A,2021.12.10
CN 113689003 A,2021.11.23
CN 112529101 A,2021.03.19
CN 113011603 A,2021.06.22
CN 113837399 A,2021.12.24
US 2021073678 A1,2021.03.11
Runhua Xu等.FedV: Privacy-Preserving Federated Learning over Vertically Partitioned Data.《arXiv》.2021,摘要、第1-6节.
Kuihe Yang等.Model Optimization Method Based on Vertical Federated Learning.《2021 IEEE International Symposium on Circuits and Systems (ISCAS)》.2021,1-5.
(续)
审查员 李华

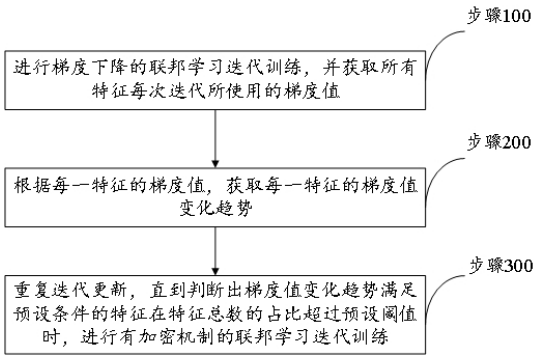
权利要求书3页 说明书13页 附图8页

(54)发明名称

一种纵向联邦学习模型的训练方法及系统

(57)摘要

本申请提供一种纵向联邦学习模型的训练方法及系统,在纵向联邦学习场景,将模型的训练流程拆分为前后两个阶段,前一训练阶段,模型信息量少且不稳定,在梯度中间值采取明文的方式由带有标签数据的发起方通信给数据参与方进行模型的学习,该过程几乎不泄露有效信息,因此,前一训练阶段进行不加密的联邦学习迭代训练,重复迭代更新,直到判断出梯度值变化趋势满足预设调节的特征在特征总数的占比超过预设阈值时,开始后一训练阶段进行有加密机制的联邦学习迭代训练。通过对模型训练过程的差异化处理,在保护数据安全的前提下,能够加快联邦学习算法的运行速度,明显提升算法性能。



CN 114330759 B

[转续页]

[接上页]

(56)对比文件

夏家骏等.基于秘密共享与同态加密的纵向联邦学习.《信息通信技术与政策》.2021,(第6期),19-26.

1.一种纵向联邦学习模型的训练方法,其特征在于,包括:

进行梯度下降的无加密机制的联邦学习迭代训练,并获取所有特征每次迭代所使用的梯度值;

根据每一特征的梯度值,获取每一特征的梯度值变化趋势;以及

重复迭代更新,直到判断出梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值时,进行有加密机制的联邦学习迭代训练;

其中,在开始训练时,梯度值变化趋势为梯度夹角值tan(angle)变大;所述梯度夹角值tan(angle)为:

$\tan(angle) = |(k_i - k_{i-1})/(1+k_i \times k_{i-1})|$

其中,k_i为第i次迭代获取的梯度值,k_{i-1}为第i-1次迭代获取的梯度值;

所述预设条件为所述梯度夹角值tan(angle)开始变小。

2.如权利要求1所述的方法,其特征在于,所述联邦学习迭代训练包括:

数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积;

对每一样本特征,将数据参与方和模型发起方的内积相加,得到总内积值,使用sigmoid函数对总内积值进行转换得到预测值;

对每一样本特征,将预测值与对应的模型发起方的真实标签做差,得到梯度中间值;以及

根据梯度中间值分别计算数据参与方和模型发起方的样本特征的梯度值,并更新样本特征的特征权重。

3.如权利要求2所述的方法,其特征在于,所述梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值,包括:

模型发起方和数据参与方的所有样本特征中梯度值变化趋势满足预设条件的数量在特征总数的占比超过预设阈值;其中,所述特征总数为模型发起方和数据参与方的样本特征数之和。

4.如权利要求1所述的方法,其特征在于,所述有加密机制的联邦学习迭代训练,包括:

数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积;

对每一样本特征,将数据参与方和模型发起方的内积相加,得到总内积值,使用sigmoid函数对总内积值进行转换得到预测值;

对每一样本特征,将预测值与对应的模型发起方的真实标签做差,得到梯度中间值;

由模型发起方,根据梯度中间值计算模型发起方的梯度值,利用模型发起方的梯度值更新样本特征的特征权重,并且,对梯度中间值进行加密得到加密梯度中间值,并发送至数据参与方;

由数据参与方,根据加密梯度中间值计算数据参与方的梯度并增加掩码,得到加密掩码梯度并发送至模型发起方;

由模型发起方,对加密掩码梯度利用私钥解密,得到掩码梯度并发送至数据参与方;以及

由数据参与方,对掩码梯度去除掩码,得到数据参与方的梯度值,并利用数据参与方的梯度值更新特征权重。

5.如权利要求4所述的方法,其特征在于,所述有加密机制的联邦学习迭代训练采用分

批训练的方式；

所述根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度并发送至模型发起方，包括：根据加密梯度中间值计算数据参与方的梯度后，对各对应样本特征的加密梯度值做聚合操作，并增加掩码，得到加密掩码聚合梯度并发送至模型发起方；

利用梯度值更新特征权重，包括：计算每一样本特征的梯度值均值，利用梯度值均值更新特征权重。

6.如权利要求4所述的方法，其特征在于，所述加密的方式包括半同态加密、全同态加密或mpc秘密共享。

7.如权利要求4所述的方法，其特征在于，还包括：

模型发起方基于真实标签和预测值计算模型的损失值，根据损失值判断模型是否收敛：

若收敛，则确定模型训练完成；

若不收敛，则继续迭代更新。

8.一种纵向联邦学习模型的训练方法，其特征在于，应用于数据参与方，包括：

计算数据参与方的每一样本特征的特征值与特征权重的内积，向模型发起方发送数据参与方的内积；

接收模型发起方发送的梯度中间值；

根据梯度中间值计算数据参与方的梯度值，并更新样本特征的特征权重；以及，数据参与方判断每一样本特征的梯度值变化趋势是否满足预设条件，向模型发起方发送数据参与方的梯度值变化趋势满足预设条件的梯度值数量；

重复迭代更新，直到接收模型发起方发送的用于进行有加密机制的联邦学习迭代训练的指令；

计算数据参与方的每一样本特征的特征值与特征权重的内积，向模型发起方发送数据参与方的内积；

接收模型发起方发送的加密梯度中间值；

根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度，并向模型发起方发送加密掩码梯度；

接收模型发起方发送的掩码梯度；

对掩码梯度去除掩码，得到数据参与方的梯度值，并利用数据参与方的梯度值更新特征权重；

其中，在开始训练时，梯度值变化趋势为梯度夹角值tan(angle)变大；所述梯度夹角值tan(angle)为：

$$\tan(angle) = \left| (k_i - k_{i-1}) / (1 + k_i \times k_{i-1}) \right|$$

其中，k_i为第i次迭代获取的梯度值，k_{i-1}为第i-1次迭代获取的梯度值；

所述预设条件为所述梯度夹角值tan(angle)开始变小。

9.一种纵向联邦学习模型的训练方法，其特征在于，应用于模型发起方，包括：

计算模型发起方每一样本特征的特征值与特征权重的内积；

接收数据参与方的样本特征的特征值与特征权重的内积；

对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用

sigmoid函数对总内积值进行转换得到预测值；

对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值，向数据参与方发送梯度中间值；

根据梯度中间值计算模型发起方的样本特征的梯度值，并更新样本特征的特征权重；

模型发起方接收数据参与方发送的数据参与方的梯度值变化趋势满足预设条件的梯度值数量；

重复迭代上述过程，直到判断出模型发起方和数据参与方的所有样本特征中梯度值变化趋势满足预设条件的数量在特征总数的占比超过预设阈值时，向数据参与方发送用于进行有加密机制的联邦学习迭代训练的指令；其中，所述特征总数为模型发起方和数据参与方的样本特征数之和；

计算模型发起方的每一样本特征的特征值与特征权重的内积；

接收数据参与方的内积；

对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；

对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值；

根据梯度中间值计算模型发起方的梯度值，利用模型发起方的梯度值更新样本特征的特征权重，并且，对梯度中间值进行加密得到加密梯度中间值，向数据参与方发送加密梯度中间值；

接收数据参与方发送的加密掩码梯度；

对加密掩码梯度利用私钥解密，得到掩码梯度，向数据参与方发送掩码梯度；

其中，在开始训练时，梯度值变化趋势为梯度夹角值tan(angle)变大；所述梯度夹角值tan(angle)为：

$$\tan(angle) = \left| (k_i - k_{i-1}) / (1 + k_i \times k_{i-1}) \right|$$

其中，k_i为第i次迭代获取的梯度值，k_{i-1}为第i-1次迭代获取的梯度值；

所述预设条件为所述梯度夹角值tan(angle)开始变小。

10.一种纵向联邦学习模型的训练系统，其特征在于，包括：

联邦学习模块，用于进行梯度下降的无加密机制的联邦学习迭代训练，并获取所有特征每次迭代所使用的梯度值；

趋势获取模块，用于根据每一特征的梯度值，获取每一特征的梯度值变化趋势；

加密学习模块，用于判断出梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值时，进行有加密机制的联邦学习迭代训练；

其中，在开始训练时，梯度值变化趋势为梯度夹角值tan(angle)变大；所述梯度夹角值tan(angle)为：

$$\tan(angle) = \left| (k_i - k_{i-1}) / (1 + k_i \times k_{i-1}) \right|$$

其中，k_i为第i次迭代获取的梯度值，k_{i-1}为第i-1次迭代获取的梯度值；

所述预设条件为所述梯度夹角值tan(angle)开始变小。

一种纵向联邦学习模型的训练方法及系统

技术领域

[0001] 本申请涉及机器学习技术领域，具体而言，涉及一种纵向联邦学习模型的训练方法及系统。

背景技术

[0002] 联邦学习最早由谷歌提出，主要是为了应对用户数据安全保护和数据隐私监管所需。其作为一种多方安全计算的实现技术，能够实现在原始数据不出门的前提下，让数据的价值进行流动，逐步被应用于金融风控、个性化推荐等领域。联邦学习根据不同的业务使用场景，主要包括纵向联邦学习、横向联邦学习以及联邦迁移算法三种类型。目前联邦学习已经可以支持多种机器学习算法。

[0003] 联邦学习提供了数据使用的安全性保证，能够规避数据安全监管风险，当下的联邦学习在工程实现上，普遍对安全性给予更多的关注，为了保护用户数据安全，使用加密算法或者多方安全计算秘密共享等方式，实现数据的隐私计算，但是，现有技术对大数据量进行加解密涉及大量的计算操作，或者采用秘密共享又会使得数据通信量倍数扩大，将导致联邦学习算法的运行速度较慢，算法性能较低。现有的联邦学习模型与明文数据集中式的模型训练速度相比，相差数倍甚至数十倍，并且随着数据量的扩大，性能差距愈发明显。而现实业务中，企业与企业之间，用户与企业服务之间都非常强调效率，如果性能损失很大，一个任务执行效率非常慢，会影响联邦学习在实际业务中落地，用户也无法接受企业所提供的服务，造成业务受损，用户流失等负面结果，进而对企业的正常发展产生不利影响。

发明内容

[0004] 本申请实施例的目的在于提供一种纵向联邦学习模型的训练方法及系统，用以解决现有技术对大数据量进行加解密涉及大量的计算操作，或者采用秘密共享又会使得数据通信量倍数扩大，将导致联邦学习算法的运行速度较慢，算法性能较低的问题。

[0005] 本申请实施例提供的一种纵向联邦学习模型的训练方法，包括：

[0006] 进行梯度下降的联邦学习迭代训练，并获取所有特征每次迭代所使用的梯度值；

[0007] 根据每一特征的梯度值，获取每一特征的梯度值变化趋势；以及

[0008] 重复迭代更新，直到判断出梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值时，进行有加密机制的联邦学习迭代训练。

[0009] 上述技术方案中，在纵向联邦学习场景，将模型的训练流程拆分为前后两个阶段，前一训练阶段，模型信息量少且不稳定，在梯度中间值采取明文的方式由带有标签数据的发起方通信给数据参与方进行模型的学习，该过程几乎不泄露有效信息，因此，前一训练阶段进行不加密的联邦学习迭代训练，重复迭代更新，直到判断出梯度值变化趋势满足预设调节的特征在特征总数的占比超过预设阈值时，开始后一训练阶段进行有加密机制的联邦学习迭代训练。通过对模型训练过程的差异化处理，在保护数据安全的前提下，能够加快联邦学习算法的运行速度，明显提升算法性能。

[0010] 在一些可选的实施方式中，在开始训练时，梯度值变化趋势为梯度夹角值tan(angle)变大；梯度夹角值tan(angle)为：

[0011] $\tan(angle) = |(k_i - k_{i-1})/(1+k_i \times k_{i-1})|$

[0012] 其中，k_i为第i次迭代获取的梯度值，k_{i-1}为第i-1次迭代获取的梯度值。

[0013] 上述技术方案中，为了将训练阶段合理拆分为前后两个阶段，引入了梯度夹角值的概念，例如：对一个特征的特征值来说，其特征权重所对应的梯度值即为导数，同时也是斜率，可以看做一条直线，相邻两次迭代所用到的两个梯度值可以看作是两条直线，这两条直线的夹角的变化反映了梯度值变化趋势，因此，梯度值变化趋势可以通过梯度夹角值来进行量化表示。

[0014] 在一些可选的实施方式中，预设条件为梯度夹角值tan(angle)开始变小。

[0015] 上述技术方案中，由于模型训练是使用梯度下降优化算法进行迭代的，在模型训练的起始阶段，梯度值变化较大，慢慢的逐步变缓，当快达到局部最优点是，梯度变缓将变得非常小，所以整个学习过程中连续两次梯度方向直线夹角先增大后减小，在单个特征场景下，将特征对应的梯度夹角值是否开始减小作为判断是否进入有加密机制的训练阶段，具备合理性。推广到多个特征的场景下，梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值时，进行有加密机制的联邦学习迭代训练，也具备合理性，其中预设阈值例如1/2、1/3等。在一些可选的实施方式中，联邦学习迭代训练包括：

[0016] 数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积；

[0017] 对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；

[0018] 对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值；以及

[0019] 根据梯度中间值分别计算数据参与方和模型发起方的样本特征的梯度值，并更新样本特征的特征权重。

[0020] 上述技术方案中，在梯度中间值采取明文的方式由模型发起方通信给数据参与方进行模型的学习，该训练过程不采用加密方式通信，具有运行速度快的优点，并且由于是前一训练阶段，模型信息量少且不稳定，几乎不泄露有效信息。

[0021] 在一些可选的实施方式中，梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值，包括：

[0022] 模型发起方和数据参与方的所有样本特征中梯度值变化趋势满足预设条件的数量在特征总数的占比超过预设阈值；其中，特征总数为模型发起方和数据参与方的样本特征数之和。

[0023] 上述技术方案中，综合统计模型发起方和数据参与方的梯度值变化趋势的情况，以评估是否进入后一阶段的有加密机制的联邦学习迭代过程，相较于只统计一方的梯度值变化趋势更具有合理性，更能准确评估联邦学习整体的训练情况。

[0024] 在一些可选的实施方式中，有加密机制的联邦学习迭代训练，包括：

[0025] 数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积；

[0026] 对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；

[0027] 对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值；

[0028] 由模型发起方，根据梯度中间值计算模型发起方的梯度值，利用模型发起方的梯度值更新样本特征的特征权重，并且，对梯度中间值进行加密得到加密梯度中间值，并发送至数据参与方；

[0029] 由数据参与方，根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度并发送至模型发起方；

[0030] 由模型发起方，对加密掩码梯度利用私钥解密，得到掩码梯度并发送至数据参与方；以及

[0031] 由数据参与方，对掩码梯度去除掩码，得到数据参与方的梯度值，并利用数据参与方的梯度值更新特征权重。

[0032] 上述技术方案中，随着迭代次数的增加，特征权重逐步找到全局最优或者局部最优点，预测值与真实值直接的误差越来越小，此时梯度信息能够提供的有效信息较多，如果泄露梯度信息，将会产生数据信息泄露的安全隐患，因此，在后一阶段的联邦学习过程中，模型发起方和数据参与方采用加密方式通信，以保护双方隐私数据。

[0033] 在一些可选的实施方式中，有加密机制的联邦学习迭代训练采用分批训练的方式；

[0034] 根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度并发送至模型发起方，包括：根据加密梯度中间值计算数据参与方的梯度后，对各对应样本特征的加密梯度值做聚合操作，并增加掩码，得到加密掩码聚合梯度并发送至模型发起方；

[0035] 利用梯度值更新特征权重，包括：计算每一样本特征的梯度值均值，利用梯度值均值更新特征权重。

[0036] 上述技术方案中，有加密机制的联邦学习迭代训练采用分批训练的方式，分批训练做聚合能够提高安全性，适用于一些对安全要求更高的场景下。

[0037] 在一些可选的实施方式中，加密的方式包括半同态加密、全同态加密或mpc秘密共享。

[0038] 上述技术方案中，模型发起方和数据参与方之间使用半同态加密、全同态加密或者mpc秘密共享等方式传输梯度信息，实现双方特征权重的隐私计算。

[0039] 在一些可选的实施方式中，还包括：

[0040] 模型发起方基于真实标签和预测值计算模型的损失值，根据损失值判断模型是否收敛：

[0041] 若收敛，则确定模型训练完成；

[0042] 若不收敛，则继续迭代更新。

[0043] 本申请实施例提供的一种纵向联邦学习模型的训练方法，应用于数据参与方，包括：

[0044] 计算数据参与方的每一样本特征的特征值与特征权重的内积，向模型发起方发送数据参与方的内积；

[0045] 接收模型发起方发送的梯度中间值；

[0046] 根据梯度中间值计算数据参与方的梯度值，并更新样本特征的特征权重；以及，数

据参与方判断每一样本特征的梯度值变化趋势是否满足预设条件，向模型发起方发送数据参与方的梯度值变化趋势满足预设条件的梯度值数量；

[0047] 重复迭代更新，直到接收模型发起方发送的用于进行有加密机制的联邦学习迭代训练的指令；

[0048] 计算数据参与方的每一样本特征的特征值与特征权重的内积，向模型发起方发送数据参与方的内积；

[0049] 接收模型发起方发送的加密梯度中间值；

[0050] 根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度，并向模型发起方发送加密掩码梯度；

[0051] 接收模型发起方发送的掩码梯度；

[0052] 对掩码梯度去除掩码，得到数据参与方的梯度值，并利用数据参与方的梯度值更新特征权重。

[0053] 上述技术方案中，数据参与方，在前一阶段的训练中，在梯度中间值采取明文的方式接收模型发起方发送的梯度中间值，该训练过程不采用加密方式通信，具有运行速度快的优点，并且由于是前一训练阶段，模型信息量少且不稳定，几乎不泄露有效信息。随着迭代次数的增加，特征权重逐步找到全局最优或者局部最优点，预测值与真实值直接的误差越来越小，此时梯度信息能够提供的有效信息较多，如果泄露梯度信息，将会产生数据信息泄露的安全隐患，因此，数据参与方，在后一阶段的联邦学习过程中，接收模型发起方发送的加密梯度中间值，以保护双方隐私数据。

[0054] 本申请实施例提供的一种纵向联邦学习模型的训练方法，应用于模型发起方，包括：

[0055] 计算模型发起方每一样本特征的特征值与特征权重的内积；

[0056] 接收数据参与方的样本特征的特征值与特征权重的内积；

[0057] 对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；

[0058] 对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值，向数据参与方发送梯度中间值；

[0059] 根据梯度中间值计算模型发起方的样本特征的梯度值，并更新样本特征的特征权重；

[0060] 模型发起方接收数据参与方发送的数据参与方的梯度值变化趋势满足预设条件的梯度值数量；

[0061] 重复迭代上述过程，直到判断出模型发起方和数据参与方的所有样本特征中梯度值变化趋势满足预设条件的数量在特征总数的占比超过预设阈值时，向数据参与方发送用于进行有加密机制的联邦学习迭代训练的指令；其中，特征总数为模型发起方和数据参与方的样本特征数之和；

[0062] 计算模型发起方的每一样本特征的特征值与特征权重的内积；

[0063] 接收数据参与方的内积；

[0064] 对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；

[0065] 对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值；

[0066] 根据梯度中间值计算模型发起方的梯度值，利用模型发起方的梯度值更新样本特征的特征权重，并且，对梯度中间值进行加密得到加密梯度中间值，向数据参与方发送加密梯度中间值；

[0067] 接收数据参与方发送的加密掩码梯度；

[0068] 对加密掩码梯度利用私钥解密，得到掩码梯度，向数据参与方发送掩码梯度。

[0069] 上述技术方案中，模型发起方，在前一阶段的训练中，在梯度中间值采取明文的方式发送的梯度中间值至数据参与方，该训练过程不采用加密方式通信，具有运行速度快的优点，并且由于是前一训练阶段，模型信息量少且不稳定，几乎不泄露有效信息。随着迭代次数的增加，特征权重逐步找到全局最优或者局部最优点，预测值与真实值直接的误差越来越小，此时梯度信息能够提供的有效信息较多，如果泄露梯度信息，将会产生数据信息泄露的安全隐患，因此，模型发起方，在后一阶段的联邦学习过程中，发送加密梯度中间值至数据参与方，以保护双方隐私数据。

[0070] 本申请实施例提供的一种纵向联邦学习模型的训练系统，包括：

[0071] 联邦学习模块，用于进行梯度下降的联邦学习迭代训练，并获取所有特征每次迭代所使用的梯度值；

[0072] 趋势获取模块，用于根据每一特征的梯度值，获取每一特征的梯度值变化趋势；

[0073] 加密学习模块，用于判断出梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值时，进行有加密机制的联邦学习迭代训练。

[0074] 上述技术方案中，在纵向联邦学习场景，将模型的训练流程拆分为前后两个阶段，利用联邦学习模块进行前一训练阶段，该训练阶段模型信息量少且不稳定，在梯度中间值采取明文的方式由带有标签数据的发起方通信给数据参与方进行模型的学习，该过程几乎不泄露有效信息，因此，前一训练阶段进行不加密的联邦学习迭代训练，重复迭代更新，直到判断出趋势获取模块所得到的梯度值变化趋势满足预设调节的特征在特征总数的占比超过预设阈值时，利用加密学习模块开始后一训练阶段进行有加密机制的联邦学习迭代训练。通过对模型训练过程的差异化处理，在保护数据安全的前提下，能够加快联邦学习算法的运行速度，明显提升算法性能。

附图说明

[0075] 为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例中所需要使用的附图作简单地介绍，应当理解，以下附图仅示出了本申请的某些实施例，因此不应被看作是对范围的限定，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他相关的附图。

[0076] 图1为梯度下降优化算法迭代过程的示意图；

[0077] 图2为模型迭代过程中梯度的变化示意图；

[0078] 图3为两阶段联邦学习模型的训练方法与迭代次数的关系示意图；

[0079] 图4为梯度夹角值的示意图；

[0080] 图5为本申请实施例提供的一种纵向联邦学习模型的训练方法步骤流程图；

[0081] 图6为本申请实施例提供的联邦学习迭代训练工作流程图；

[0082] 图7为本申请实施例提供的有加密机制的联邦学习迭代训练工作流程图；

[0083] 图8为本申请实施例提供的一种纵向联邦学习模型的训练方法中数据参与方和模型发起方的工作流程图；

[0084] 图9为本申请另一实施例提供的一种纵向联邦学习模型的训练方法中数据参与方和模型发起方的工作流程图；

[0085] 图10为本申请实施例提供的一种纵向联邦学习模型的训练系统功能模块图。

[0086] 图标：1-联邦学习模块，2-趋势获取模块，3-加密学习模块。

具体实施方式

[0087] 下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行描述。

[0088] 本申请实施例以纵向联邦学习场景下逻辑回归算法的训练流程进行原理阐述，说明本申请的一个或多个实施例中将联邦学习训练阶段分为前后两个阶段的差异化训练的合理性。需明确的是，除了本实施例的逻辑回归算法场景，其他使用梯度下降的算法的场景下都可以应用本申请实施例的一种纵向联邦学习模型的训练方法及系统。

[0089] 使用梯度下降优化算法迭代的过程如图1所示，在两次模型训练过程中，A点、C点代表初始状态参数，B点、D点分别为A点、C点对应的局部最优点，J为损失函数，从该几何图中优化路径的陡峭程度变化可以看出，模型训练的起始阶段，即，从A点出发的与A点相近的点，从C点出发与C点相近的点，梯度值变化比较大，慢慢的逐步变缓，当快到达局部最优点B或D时，梯度变化将变得非常小。这个过程对应于参数更新中，参数逐步收敛。由于梯度变缓，使得参数的更新变化幅度也降低，当达到收敛阈值，则停止更新，模型学习到有效的参数。

[0090] 图2以更直观的形式，展现了模型迭代过程中，梯度的变化情况。如图2所示，迭代初始阶段，损失函数J(θ)与对应参数θ的导数，即梯度，变化非常剧烈。这是因为初始阶段的参数为随机点，也就是初始化生成的随机数作为参数的起始点，其与真实最优点的差距比较大，因此梯度变化较明显。随着迭代轮数增多，参数逐步找到全局或者局部最优点，也就是此时的预测值与真实值之间的误差越来越小，此时梯度信息能够提供的信息比初始阶段要更多，如果在迭代后期梯度信息泄露，将会产生数据信息的泄露。然而在起始阶段，即使梯度信息被暴露，也不会产生明显的影响。因为此时的特征权重依然具有较大随机性，只是处于参数调整为最优值的前期。

[0091] 因此，申请人在图3中引入两阶段联邦学习模型的训练方法。可以看到在第一阶段，也就是在迭代次数iter=0，iter=1，iter=2三轮迭代中，梯度的变化非常明显，整个梯度绝对值由初始状态的最大值急剧衰减，当然在实际过程中，这个过程往往伴随这震荡衰减，但梯度绝对值总体是趋向减小的方向。然后iter=2是一个拐点，在第二阶段在iter=3，iter=4梯度绝对值已经基本处于微调，逐步靠近最优点，梯度值也趋向于0(如iter=5所示)，说明在第二阶段，模型已经基本学到了相当多的信息。据以上所述，由于两阶段的模型信息量差异，因此可以对两个不同阶段采取不同的数据安全保护手段。在本发明中，第一阶段因为模型信息量少且不稳定，在梯度中间值采取明文的方式由带有标签数据的发起方通信给数据参与方进行模型的学习，此过程几乎不泄露有效信息。第二阶段由于模型已经具

备足够的信息量，且逐步趋近于模型全局最优解或者局部最优解，因此对于梯度信息需要采取半同态加密手段或者多方安全计算秘密共享技术来进行数据信息保护。

[0092] 为了确定两个阶段的划分界限，在此引入夹角(两条直线形成的小于90度的角)概念。假如模型中只有一个特征变量x，对于x的特征权重w来说，梯度即为导数，同时也是斜率。初始阶段连续两次梯度值变化很大，这两次梯度方向的直线所形成的夹角相对较小。随着迭代进行，连续两次梯度值变化逐步变小，梯度方向的直线所形成的夹角逐步变大，当达到一个临界点，连续两次梯度值变化更小，梯度方向的直线形成的夹角又开始变小。所以整个过程中连续两次梯度方向直线夹角先增大后减小，从图3中的5轮迭代，也可以看出同样的变化趋势。

[0093] 请参照图4，图4为梯度夹角值的示意图，申请人发现可以采用梯度夹角值tan(angle)来表征这个迭代过程，tan(angle) =| (k2-k1)/(1+k1×k2)|。

[0094] 下面举例说明：假如连续5次梯度值为-5，-3，-2，-1，-0.8，套用假设的梯度夹角值，可以看到：第一个梯度夹角值：|(-5+3)/(1+15)|=1/8；第二个梯度夹角值：|(-3+2)/(1+6)|=1/7；第三个梯度夹角值：|(-2+1)/(1+2)|=1/3；第四个梯度夹角值：|(-1+0.8)/(1+0.8)|=1/9。从这四次计算来看，梯度夹角值确实存在先变大后变小的过程。那么推广到k个x特征场景，申请人定义当存在1/2个特征的特征权重对应的连续梯度方向直线对应的夹角发生由大变小的情形时，该时刻即为临界点，也就是第二阶段开始的拐点。如果在某些场景下，需要更严格限制，则可以设置当1/3特征发生以上夹角变化情况即为第二阶段的拐点。

[0095] 基于上述的原理，下面详细阐述本申请实施例提供的一种纵向联邦学习模型的训练方法。

[0096] 请参照图5，图5为本申请实施例提供的一种纵向联邦学习模型的训练方法步骤流程图，包括：

[0097] 步骤101、进行梯度下降的联邦学习迭代训练，并获取所有特征每次迭代所使用的梯度值；

[0098] 步骤102、根据每一特征的梯度值，获取每一特征的梯度值变化趋势；

[0099] 其中，梯度值变化趋势可以通过梯度夹角值或其他能够反映梯度值变化的参数来反映。

[0100] 步骤103、重复迭代更新，直到判断出梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值时，进行有加密机制的联邦学习迭代训练。

[0101] 本申请实施例中，在纵向联邦学习场景，将模型的训练流程拆分为前后两个阶段，前一训练阶段，模型信息量少且不稳定，在梯度中间值采取明文的方式由带有标签数据的发起方通信给数据参与方进行模型的学习，该过程几乎不泄露有效信息，因此，前一训练阶段进行不加密的联邦学习迭代训练，重复迭代更新，直到判断出梯度值变化趋势满足预设调节的特征在特征总数的占比超过预设阈值时，开始后一训练阶段进行有加密机制的联邦学习迭代训练。通过对模型训练过程的差异化处理，在保护数据安全的前提下，能够加快联邦学习算法的运行速度，明显提升算法性能。

[0102] 在一些可选的实施方式中，在开始训练时，梯度值变化趋势为梯度夹角值tan(angle)变大；梯度夹角值tan(angle)为：

[0103] $\tan(angle) = \left| (k_i - k_{i-1}) / (1 + k_i \times k_{i-1}) \right|$

[0104] 其中，k_i为第i次迭代获取的梯度值，k_{i-1}为第i-1次迭代获取的梯度值。

[0105] 本申请实施例中，为了将训练阶段合理拆分为前后两个阶段，引入了梯度夹角值的概念，例如：对一个特征的特征值来说，其特征权重所对应的梯度值即为导数，同时也是斜率，可以看做一条直线，相邻两次迭代所用到的两个梯度值可以看作是两条直线，这两条直线的夹角的变化反映了梯度值变化趋势，因此，梯度值变化趋势可以通过梯度夹角值来进行量化表示。

[0106] 在一些可选的实施方式中，预设条件为梯度夹角值tan(angle)开始变小。

[0107] 这里的梯度夹角值开始变小可以是梯度夹角值开始变小的该次迭代，也可以是梯度夹角值开始变小的下次迭代或下下次迭代等，可根据实际需求进行设置。

[0108] 本申请实施例中，由于模型训练是使用梯度下降优化算法进行迭代的，在模型训练的起始阶段，梯度值变化较大，慢慢的逐步变缓，当快达到局部最优点是，梯度变缓将变得非常小，所以整个学习过程中连续两次梯度方向直线夹角先增大后减小，在单个特征场景下，将特征对应的梯度夹角值是否开始减小作为判断是否进入有加密机制的训练阶段，具备合理性。推广到多个特征的场景下，梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值时，进行有加密机制的联邦学习迭代训练，也具备合理性，其中预设阈值例如1/2、1/3等，预设阈值的设置可以根据对安全性要求的强度进行调节，要求越高则阈值设置越小。

[0109] 联邦学习的参与方包括模型发起方和数据参与方，且满足以下条件：

[0110] (1)模型发起方，需要具有标签数据及部分特征数据。

[0111] (2)数据参与方仅具有部分特征数据。

[0112] (3)各参与方所持有的特征数据数量都分别需要大于3个，不允许存在单特征或者无特征。之所以控制数量，是为了避免信息泄漏风险。

[0113] (4)当然标签数据不一定在模型发起方，也可以在数据参与方，那么相应的流程需要做一定的调整，具有标签数据的一方需要计算梯度中间值信息以及加解密动作。本申请的一个或多个实施例中均以标签数据在模型发起方为例进行阐述。

[0114] (5)特征中的值不允许全0，或者全1。

[0115] 请参照图6，图6为本申请实施例提供的联邦学习迭代训练工作流程图，包括：

[0116] 步骤101、数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积；

[0117] 步骤102、对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；

[0118] 其中，由于针对的场景为纵向联邦学习，所以数据参与方和模型发起方两方的样本特征是经过安全求交对齐的，样本特征数也是一致的，也就是说每一特征在数据参与方和模型发起方分别对应有样本特征，因此，可以分别按照每个样本将数据参与方和模型发起方的内积相加，得到总内积值。sigmoid的函数形式为：$y=1/(1+e^{-z})$；总内积值$z=W\times X_i$，$W\times X\_i=W\_a\times X\_ai+W\_b\times X\_bi$，W_a为模型发起方的特征权重，W_b为数据参与方的特征权重，X_ai为模型发起方的特征值，X_bi为数据参与方的特征值。

[0119] 步骤103、对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值；

[0120] 在两方场景的训练流程中，模型发起方持有特征和标签信息，而数据参与方仅持有特征，没有标签信息，其也是模型发起方和数据参与方的区别之一。

[0121] 步骤104、根据梯度中间值分别计算数据参与方和模型发起方的样本特征的梯度值，并更新样本特征的特征权重。

[0122] 本申请实施例中，在梯度中间值采取明文的方式由模型发起方通信给数据参与方进行模型的学习，该训练过程不采用加密方式通信，具有运行速度快的优点，并且由于是前一训练阶段，模型信息量少且不稳定，几乎不泄露有效信息。

[0123] 在一些可选的实施方式中，梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值，包括：模型发起方和数据参与方的所有样本特征中梯度值变化趋势满足预设条件的数量在特征总数的占比超过预设阈值；其中，特征总数为模型发起方和数据参与方的样本特征数之和。

[0124] 本申请实施例中，综合统计模型发起方和数据参与方的梯度值变化趋势的情况，以评估是否进入后一阶段的有加密机制的联邦学习迭代过程，相较于只统计一方的梯度值变化趋势更具有合理性，更能准确评估联邦学习整体的训练情况。

[0125] 请参照图7，图7为本申请实施例提供的有加密机制的联邦学习迭代训练工作流程图，包括：

[0126] 步骤301、数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积；

[0127] 步骤302、对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；

[0128] 同样的，由于针对的场景为纵向联邦学习，所以数据参与方和模型发起方两方的样本特征是经过安全求交对齐的，样本特征数也是一致的，也就是说每一特征在数据参与方和模型发起方分别对应有样本特征，因此，可以分别按照每个样本将数据参与方和模型发起方的内积相加，得到总内积值。sigmoid的函数形式为：$y=1/(1+e^{-z})$；总内积值$z=W\times X_i$，$W\times X\_i=W\_a\times X\_ai+W\_b\times X\_bi$，W_a为模型发起方的特征权重，W_b为数据参与方的特征权重，X_ai为模型发起方的特征值，X_bi为数据参与方的特征值。

[0129] 步骤303、对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值；

[0130] 步骤304、由模型发起方，根据梯度中间值计算模型发起方的梯度值，利用模型发起方的梯度值更新样本特征的特征权重，并且，对梯度中间值进行加密得到加密梯度中间值，并发送至数据参与方；

[0131] 步骤305、由数据参与方，根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度并发送至模型发起方；

[0132] 步骤306、由模型发起方，对加密掩码梯度利用私钥解密，得到掩码梯度并发送至数据参与方；

[0133] 步骤307、由数据参与方，对掩码梯度去除掩码，得到数据参与方的梯度值，并利用数据参与方的梯度值更新特征权重。

[0134] 本申请实施例中，随着迭代次数的增加，特征权重逐步找到全局最优或者局部最优点，预测值与真实值直接的误差越来越小，此时梯度信息能够提供的有效信息较多，如果

泄露梯度信息，将会产生数据信息泄露的安全隐患，因此，在后一阶段的联邦学习过程中，模型发起方和数据参与方采用加密方式通信，以保护双方隐私数据。

[0135] 在一些可选的实施方式中，有加密机制的联邦学习迭代训练采用分批训练的方式。对应的，根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度并发送至模型发起方，包括：根据加密梯度中间值计算数据参与方的梯度后，对各对应样本特征的加密梯度值做聚合操作，并增加掩码，得到加密掩码聚合梯度并发送至模型发起方。同样的，利用梯度值更新特征权重，包括：计算每一样本特征的梯度值均值，利用梯度值均值更新特征权重。

[0136] 本申请实施例中，有加密机制的联邦学习迭代训练采用分批训练的方式，分批训练做聚合能够提高安全性，适用于一些对安全要求更高的场景下。

[0137] 在一些可选的实施方式中，加密的方式包括半同态加密、全同态加密或mpc秘密共享。

[0138] 本申请实施例中，模型发起方和数据参与方之间使用半同态加密、全同态加密或者mpc秘密共享等方式传输梯度信息，实现双方特征权重的隐私计算。

[0139] 在一些可选的实施方式中，还包括：模型发起方基于真实标签和预测值计算模型的损失值，根据损失值判断模型是否收敛：

[0140] 若收敛，则确定模型训练完成，输出模型参数；

[0141] 若不收敛，则继续迭代更新。

[0142] 请参照图8，图8为本申请实施例提供的一种纵向联邦学习模型的训练方法中数据参与方和模型发起方的工作流程图。

[0143] 本申请实施例提供的一种纵向联邦学习模型的训练方法，应用于数据参与方，包括：计算数据参与方的每一样本特征的特征值与特征权重的内积，向模型发起方发送数据参与方的内积；接收模型发起方发送的梯度中间值；根据梯度中间值计算数据参与方的梯度值，并更新样本特征的特征权重；以及，数据参与方判断每一样本特征的梯度值变化趋势是否满足预设条件，向模型发起方发送数据参与方的梯度值变化趋势满足预设条件的梯度值数量；重复迭代更新，直到接收模型发起方发送的用于进行有加密机制的联邦学习迭代训练的指令；计算数据参与方的每一样本特征的特征值与特征权重的内积，向模型发起方发送数据参与方的内积；接收模型发起方发送的加密梯度中间值；根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度，并向模型发起方发送加密掩码梯度；接收模型发起方发送的掩码梯度；对掩码梯度去除掩码，得到数据参与方的梯度值，并利用数据参与方的梯度值更新特征权重。

[0144] 本实施例中，数据参与方，在前一阶段的训练中，在梯度中间值采取明文的方式接收模型发起方发送的梯度中间值，该训练过程不采用加密方式通信，具有运行速度快的优点，并且由于是前一训练阶段，模型信息量少且不稳定，几乎不泄露有效信息。随着迭代次数的增加，特征权重逐步找到全局最优或者局部最优点，预测值与真实值直接的误差越来越小，此时梯度信息能够提供的有效信息较多，如果泄露梯度信息，将会产生数据信息泄露的安全隐患，因此，数据参与方，在后一阶段的联邦学习过程中，接收模型发起方发送的加密梯度中间值，以保护双方隐私数据。

[0145] 本申请实施例提供的一种纵向联邦学习模型的训练方法，应用于模型发起方，包

括:计算模型发起方每一样本特征的特征值与特征权重的内积;接收数据参与方的样本特征的特征值与特征权重的内积;对每一样本特征,将数据参与方和模型发起方的内积相加,得到总内积值,使用sigmoid函数对总内积值进行转换得到预测值;对每一样本特征,将预测值与对应的模型发起方的真实标签做差,得到梯度中间值,向数据参与方发送梯度中间值;根据梯度中间值计算模型发起方的样本特征的梯度值,并更新样本特征的特征权重;模型发起方接收数据参与方发送的数据参与方的梯度值变化趋势满足预设条件的梯度值数量;重复迭代上述过程,直到判断出模型发起方和数据参与方的所有样本特征中梯度值变化趋势满足预设条件的数量在特征总数的占比超过预设阈值时,向数据参与方发送用于进行有加密机制的联邦学习迭代训练的指令;其中,特征总数为模型发起方和数据参与方的样本特征数之和;计算模型发起方的每一样本特征的特征值与特征权重的内积;接收数据参与方的内积;对每一样本特征,将数据参与方和模型发起方的内积相加,得到总内积值,使用sigmoid函数对总内积值进行转换得到预测值;对每一样本特征,将预测值与对应的模型发起方的真实标签做差,得到梯度中间值;根据梯度中间值计算模型发起方的梯度值,利用模型发起方的梯度值更新样本特征的特征权重,并且,对梯度中间值进行加密得到加密梯度中间值,向数据参与方发送加密梯度中间值;接收数据参与方发送的加密掩码梯度;对加密掩码梯度利用私钥解密,得到掩码梯度,向数据参与方发送掩码梯度。

[0146] 本实施例中,模型发起方,在前一阶段的训练中,在梯度中间值采取明文的方式发送的梯度中间值至数据参与方,该训练过程不采用加密方式通信,具有运行速度快的优点,并且由于是前一训练阶段,模型信息量少且不稳定,几乎不泄露有效信息。随着迭代次数的增加,特征权重逐步找到全局最优或者局部最优点,预测值与真实值直接的误差越来越小,此时梯度信息能够提供的有效信息较多,如果泄露梯度信息,将会产生数据信息泄露的安全隐患,因此,模型发起方,在后一阶段的联邦学习过程中,发送加密梯度中间值至数据参与方,以保护双方隐私数据。

[0147] 请参照图9,图9为另一实施例提供的纵向联邦学习模型的训练方法,具体如下:

[0148] 本申请实施例提供的一种纵向联邦学习模型的训练方法,应用于数据参与方,包括:计算数据参与方的每一样本特征的特征值与特征权重的内积,向模型发起方发送数据参与方的内积;接收模型发起方发送的梯度中间值;根据梯度中间值计算数据参与方的梯度值,并更新样本特征的特征权重;以及,数据参与方判断每一样本特征的梯度值变化趋势是否满足预设条件,向模型发起方发送数据参与方的梯度值变化趋势满足预设条件的梯度值数量;重复迭代更新,直到接收模型发起方发送的用于进行有加密机制的联邦学习迭代训练的指令;计算数据参与方的每一样本特征的特征值与特征权重的内积,向模型发起方发送数据参与方的内积;接收模型发起方发送的加密梯度中间值;根据加密梯度中间值计算数据参与方的加密梯度值做聚合操作,并增加掩码,得到加密掩码聚合梯度,并向模型发起方发送加密掩码聚合梯度;接收模型发起方发送的掩码聚合梯度;对掩码聚合梯度去除掩码,计算数据参与方的梯度值均值,并利用数据参与方的梯度值均值更新特征权重。

[0149] 本申请实施例提供的一种纵向联邦学习模型的训练方法,应用于模型发起方,包括:计算模型发起方每一样本特征的特征值与特征权重的内积;接收数据参与方的样本特征的特征值与特征权重的内积; 对每一样本特征,将数据参与方和模型发起方的内积相加,得到总内积值,使用sigmoid函数对总内积值进行转换得到预测值;对每一样本特征,将

预测值与对应的模型发起方的真实标签做差，得到梯度中间值，向数据参与方发送梯度中间值；根据梯度中间值计算模型发起方的样本特征的梯度值，并更新样本特征的特征权重；模型发起方接收数据参与方发送的数据参与方的梯度值变化趋势满足预设条件的梯度值数量；重复迭代上述过程，直到判断出模型发起方和数据参与方的所有样本特征中梯度值变化趋势满足预设条件的数量在特征总数的占比超过预设阈值时，向数据参与方发送用于进行有加密机制的联邦学习迭代训练的指令；其中，特征总数为模型发起方和数据参与方的样本特征数之和；计算模型发起方的每一样本特征的特征值与特征权重的内积；接收数据参与方的内积；对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值；根据梯度中间值计算模型发起方的梯度值，利用模型发起方的梯度值更新样本特征的特征权重，并且，对梯度中间值进行加密得到加密梯度中间值，向数据参与方发送加密梯度中间值；接收数据参与方发送的加密掩码聚合梯度；对加密掩码聚合梯度利用私钥解密，得到掩码聚合梯度，向数据参与方发送掩码聚合梯度。

[0150] 本实施例中，模型训练需要采用分批训练的模式，不允许单条执行。数据参与方(即仅有部分特征数据)在加密梯度中间值与特征的计算得到加密梯度值后需要做求和操作，得到每一个特征对应的加密的梯度之和，然后再加上随机数掩码。通过该聚合操作，可以有效避免原始样本数据的信息暴露。在这里，一个批次的大小需要大于某一量级，比如100或者1000一个批次，批次中数量越大安全性越高。

[0151] 请参照图10，图10为本申请实施例提供的一种纵向联邦学习模型的训练系统功能模块图，包括联邦学习模块1、趋势获取模块2和加密学习模块3。

[0152] 其中，联邦学习模块1，用于进行梯度下降的联邦学习迭代训练，并获取所有特征每次迭代所使用的梯度值；趋势获取模块2，用于根据每一特征的梯度值，获取每一特征的梯度值变化趋势；加密学习模块3，用于判断出梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值时，进行有加密机制的联邦学习迭代训练。

[0153] 本申请实施例中，在纵向联邦学习场景，将模型的训练流程拆分为前后两个阶段，利用联邦学习模块1进行前一训练阶段，该训练阶段模型信息量少且不稳定，在梯度中间值采取明文的方式由带有标签数据的发起方通信给数据参与方进行模型的学习，该过程几乎不泄露有效信息，因此，前一训练阶段进行不加密的联邦学习迭代训练，重复迭代更新，直到判断出趋势获取模块2所得到的梯度值变化趋势满足预设调节的特征在特征总数的占比超过预设阈值时，利用加密学习模块3开始后一训练阶段进行有加密机制的联邦学习迭代训练。通过对模型训练过程的差异化处理，在保护数据安全的前提下，能够加快联邦学习算法的运行速度，明显提升算法性能。

[0154] 在一些可选的实施方式中，所述联邦学习模块1还用于：数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积；对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值；以及，根据梯度中间值分别计算数据参与方和模型发起方的样本特征的梯度值，并更新样本特征的特征权重。

[0155] 利用联邦学习模块1，在梯度中间值采取明文的方式由模型发起方通信给数据参与方进行模型的学习，该训练过程不采用加密方式通信，具有运行速度快的优点，并且由于是前一训练阶段，模型信息量少且不稳定，几乎不泄露有效信息。

[0156] 在一些可选的实施方式中，加密学习模块3还用于：数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积；对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值；对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值；由模型发起方，根据梯度中间值计算模型发起方的梯度值，利用模型发起方的梯度值更新样本特征的特征权重，并且，对梯度中间值进行加密得到加密梯度中间值，并发送至数据参与方；由数据参与方，根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度并发送至模型发起方；由模型发起方，对加密掩码梯度利用私钥解密，得到掩码梯度并发送至数据参与方；以及，由数据参与方，对掩码梯度去除掩码，得到数据参与方的梯度值，并利用数据参与方的梯度值更新特征权重。

[0157] 随着迭代次数的增加，特征权重逐步找到全局最优或者局部最优点，预测值与真实值直接的误差越来越小，此时梯度信息能够提供的有效信息较多，如果泄露梯度信息，将会产生数据信息泄露的安全隐患，因此，利用加密学习模块3，在后一阶段的联邦学习过程中，模型发起方和数据参与方采用加密方式通信，以保护双方隐私数据。

[0158] 在本申请所提供的实施例中，应该理解到，所揭露装置和方法，可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，又例如，多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些通信接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

[0159] 另外，作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0160] 再者，在本申请各个实施例中的各功能模块可以集成在一起形成一个独立的部分，也可以是各个模块单独存在，也可以两个或两个以上模块集成形成一个独立的部分。

[0161] 在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。

[0162] 以上所述仅为本申请的实施例而已，并不用于限制本申请的保护范围，对于本领域的技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

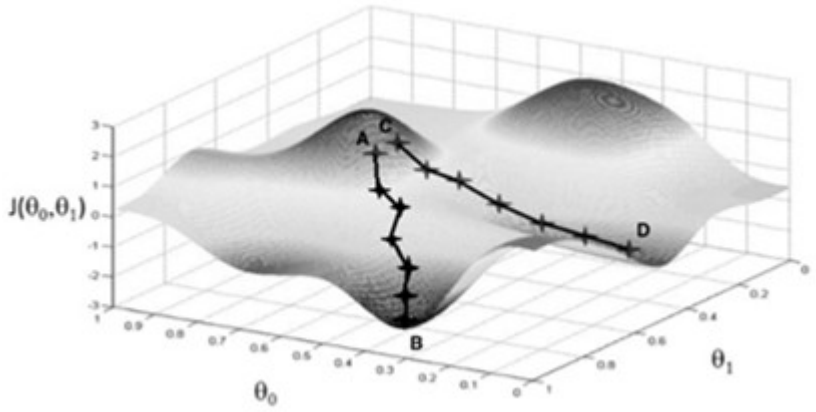


图1

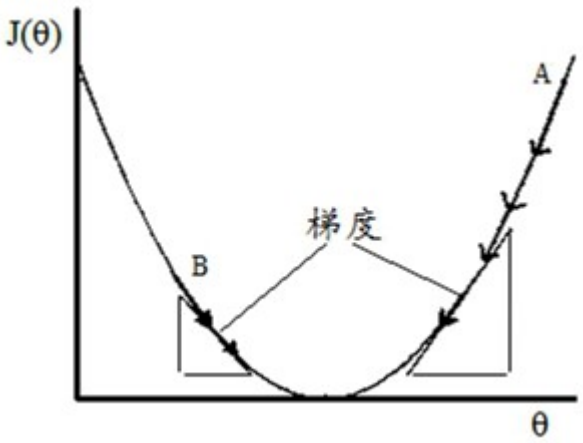


图2

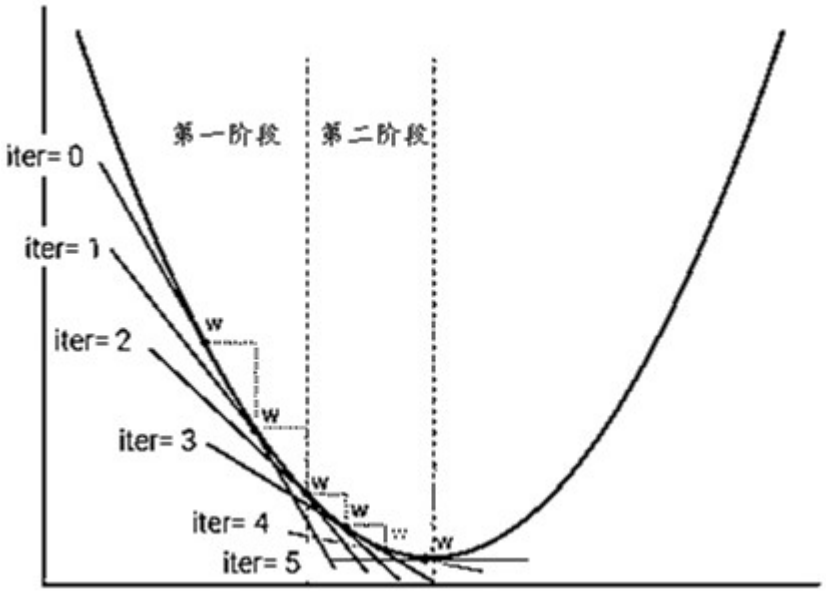


图3

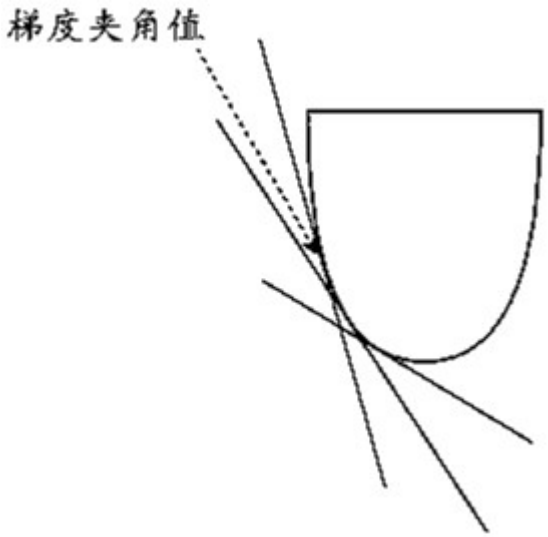


图4

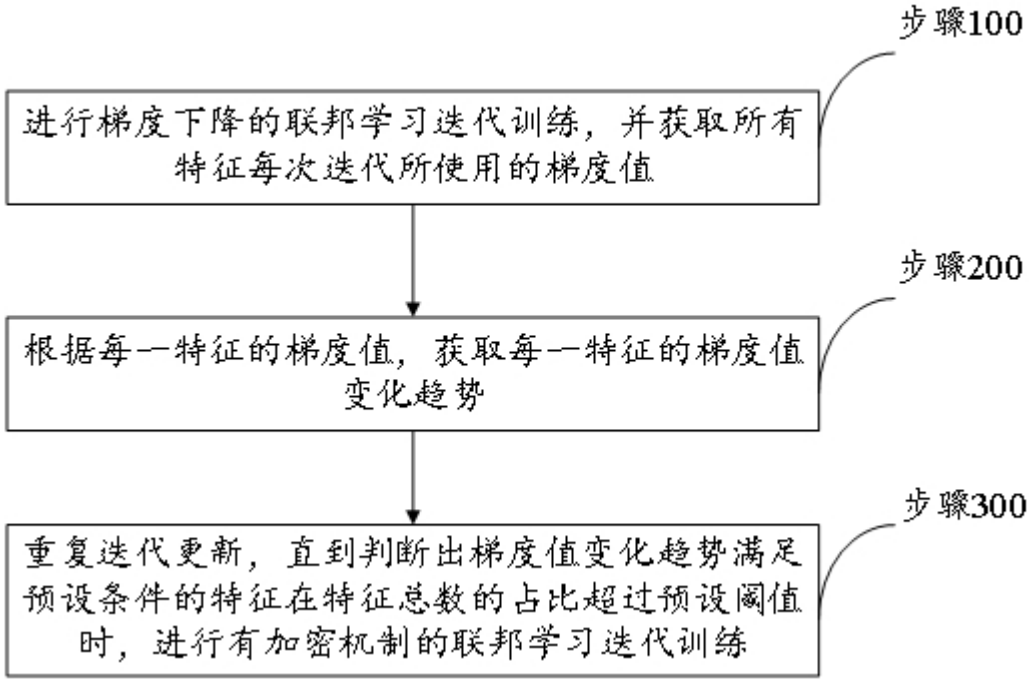

步骤100
进行梯度下降的联邦学习迭代训练，并获取所有特征每次迭代所使用的梯度值
步骤200
根据每一特征的梯度值，获取每一特征的梯度值变化趋势
步骤300
重复迭代更新，直到判断出梯度值变化趋势满足预设条件的特征在特征总数的占比超过预设阈值时，进行有加密机制的联邦学习迭代训练

图5

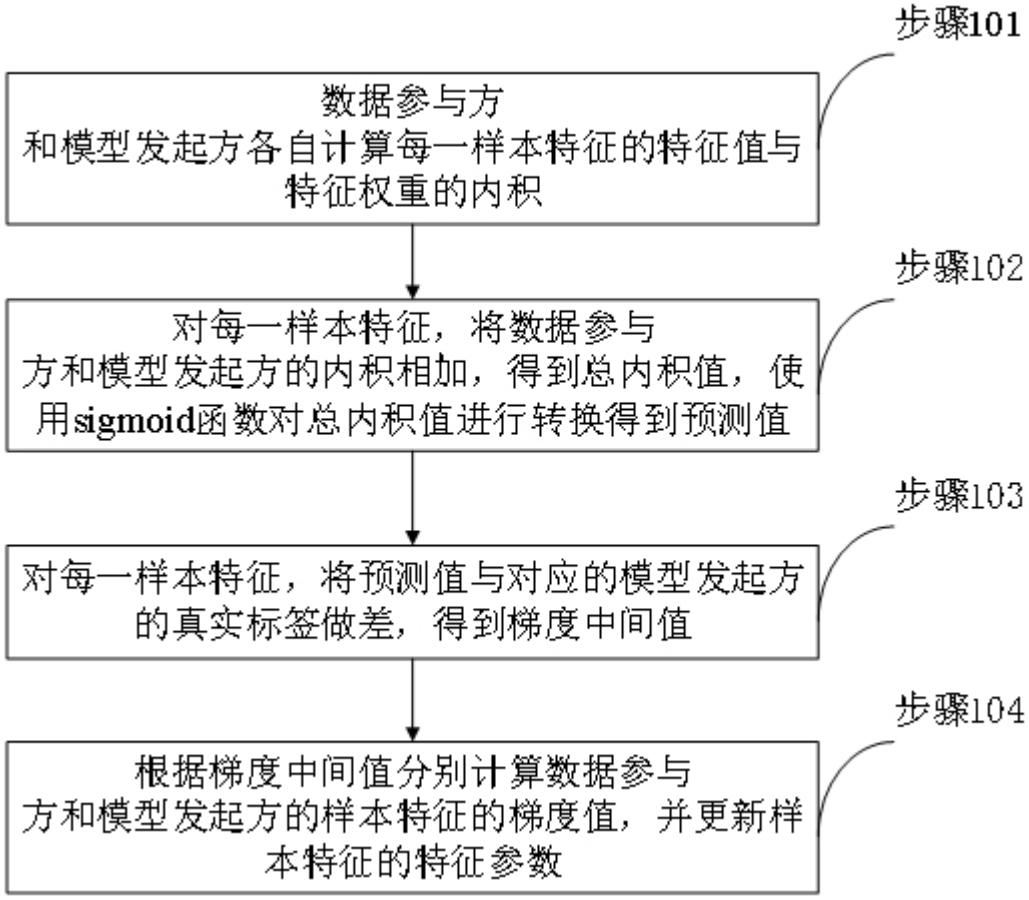

步骤101
数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积
步骤102
对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值
步骤103
对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值
步骤104
根据梯度中间值分别计算数据参与方和模型发起方的样本特征的梯度值，并更新样本特征的特征参数

图6

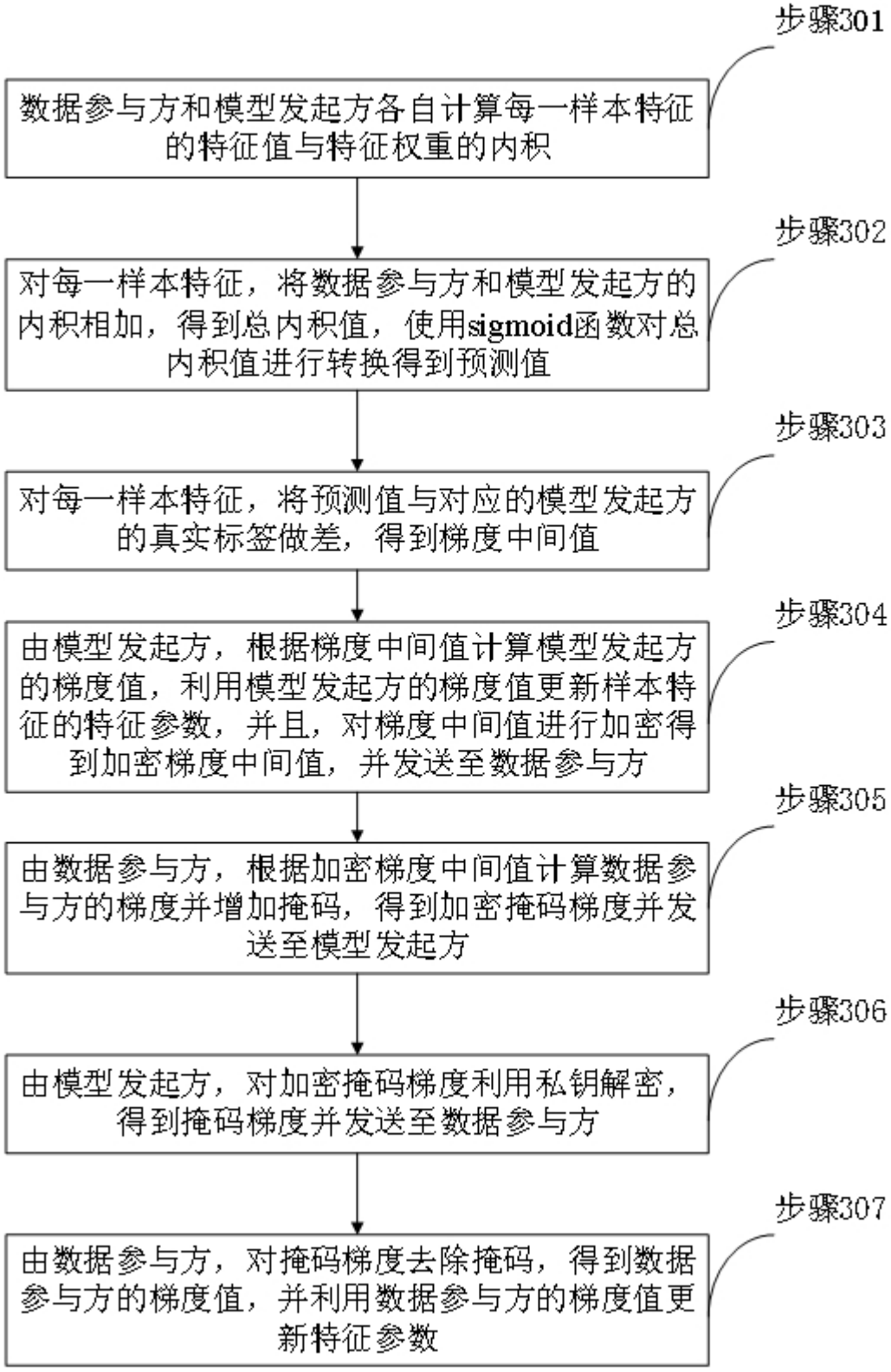

步骤301
数据参与方和模型发起方各自计算每一样本特征的特征值与特征权重的内积
步骤302
对每一样本特征，将数据参与方和模型发起方的内积相加，得到总内积值，使用sigmoid函数对总内积值进行转换得到预测值
步骤303
对每一样本特征，将预测值与对应的模型发起方的真实标签做差，得到梯度中间值
步骤304
由模型发起方，根据梯度中间值计算模型发起方的梯度值，利用模型发起方的梯度值更新样本特征的特征参数，并且，对梯度中间值进行加密得到加密梯度中间值，并发送至数据参与方
步骤305
由数据参与方，根据加密梯度中间值计算数据参与方的梯度并增加掩码，得到加密掩码梯度并发送至模型发起方
步骤306
由模型发起方，对加密掩码梯度利用私钥解密，得到掩码梯度并发送至数据参与方
步骤307
由数据参与方，对掩码梯度去除掩码，得到数据参与方的梯度值，并利用数据参与方的梯度值更新特征参数

图7

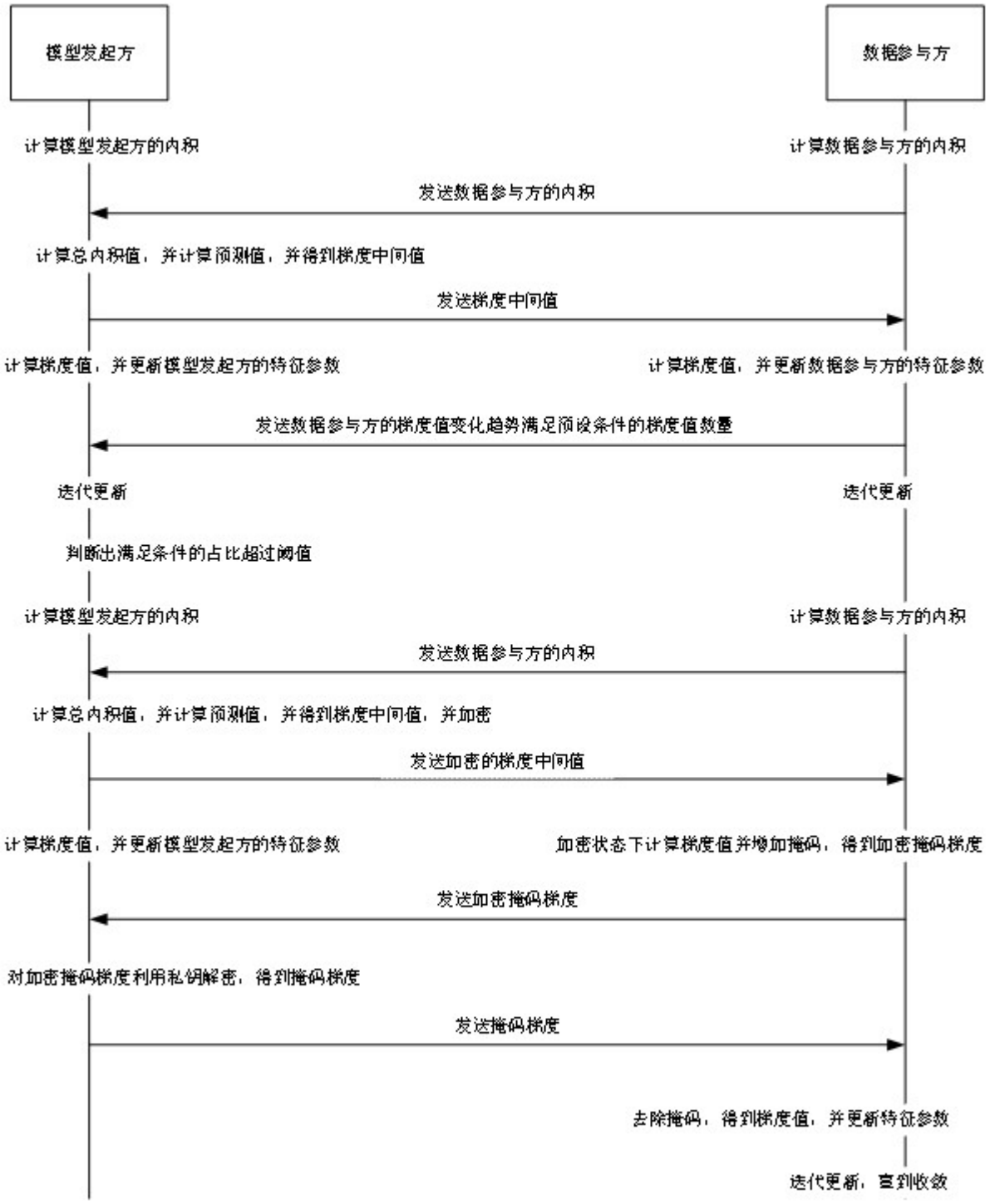

模型发起方
数据参与方
计算模型发起方的内积
计算数据参与方的内积
发送数据参与方的内积
计算总内积值，并计算预测值，并得到梯度中间值
发送梯度中间值
计算梯度值，并更新模型发起方的特征参数
计算梯度值，并更新数据参与方的特征参数
发送数据参与方的梯度值变化趋势满足预设条件的梯度值数量
迭代更新
迭代更新
判断出满足条件的占比超过阈值
计算模型发起方的内积
计算数据参与方的内积
发送数据参与方的内积
计算总内积值，并计算预测值，并得到梯度中间值，并加密
发送加密的梯度中间值
计算梯度值，并更新模型发起方的特征参数
加密状态下计算梯度值并增加掩码，得到加密掩码梯度
发送加密掩码梯度
对加密掩码梯度利用私钥解密，得到掩码梯度
发送掩码梯度
去除掩码，得到梯度值，并更新特征参数
迭代更新，直到收敛

图8

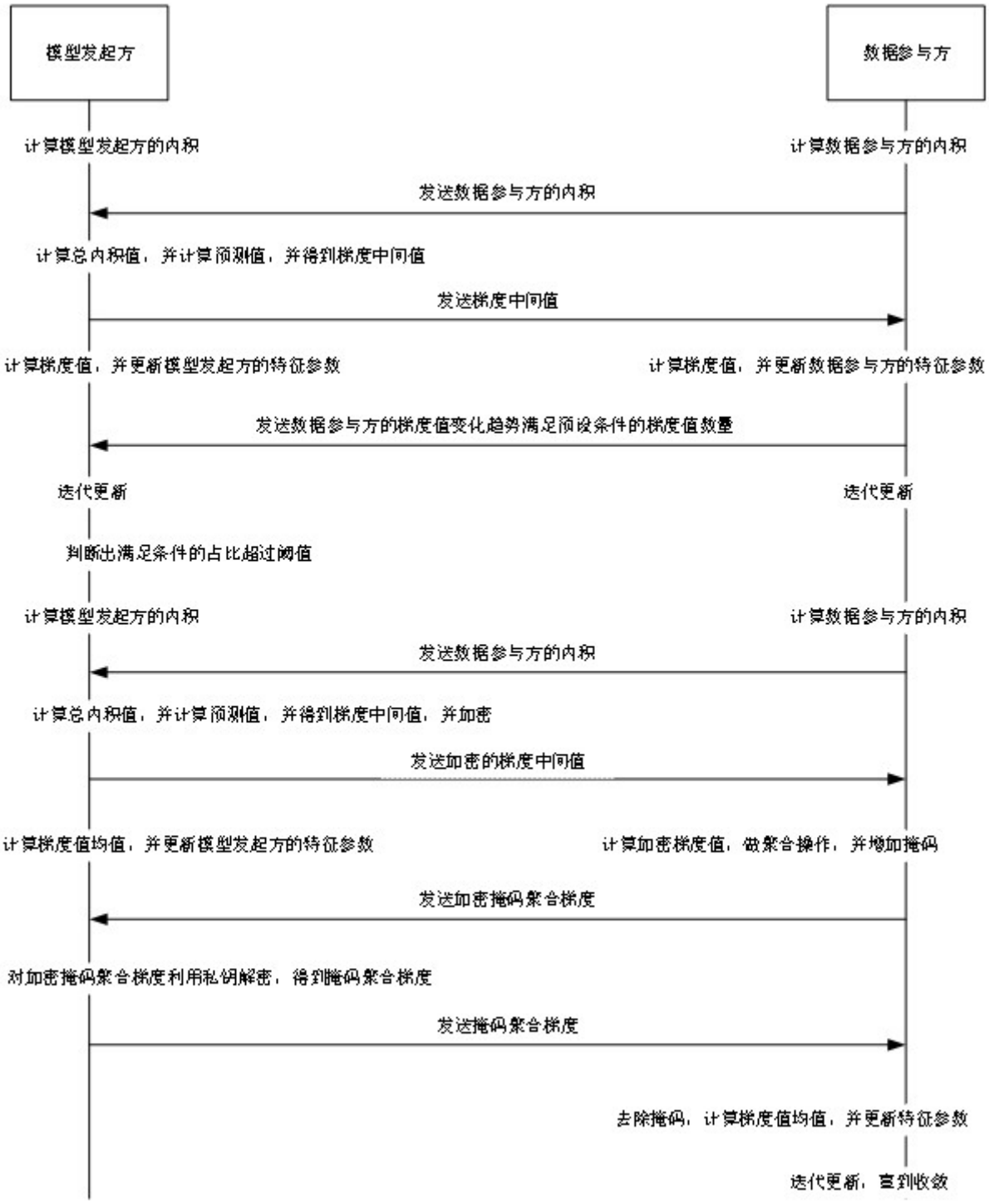

模型发起方
数据参与方
计算模型发起方的内积
计算数据参与方的内积
发送数据参与方的内积
计算总内积值，并计算预测值，并得到梯度中间值
发送梯度中间值
计算梯度值，并更新模型发起方的特征参数
计算梯度值，并更新数据参与方的特征参数
发送数据参与方的梯度值变化趋势满足预设条件的梯度值数量
迭代更新
迭代更新
判断出满足条件的占比超过阈值
计算模型发起方的内积
计算数据参与方的内积
发送数据参与方的内积
计算总内积值，并计算预测值，并得到梯度中间值，并加密
发送加密的梯度中间值
计算梯度值均值，并更新模型发起方的特征参数
计算加密梯度值，做聚合操作，并增加掩码
发送加密掩码聚合梯度
对加密掩码聚合梯度利用私钥解密，得到掩码聚合梯度
发送掩码聚合梯度
去除掩码，计算梯度值均值，并更新特征参数
迭代更新，直到收敛

图9

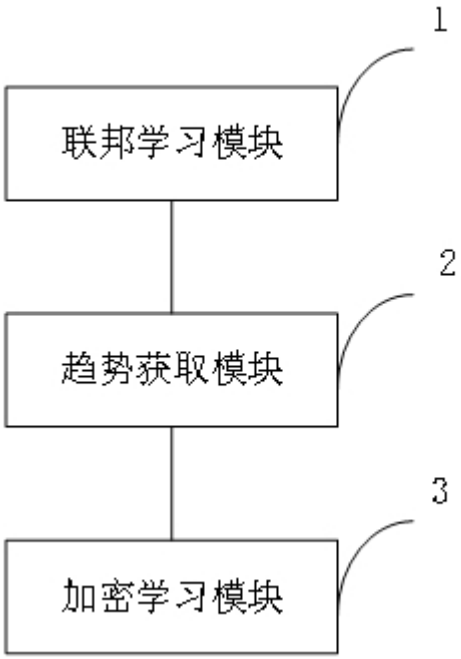


图10