

基于深度学习的动态差分隐私保护算法

何佳阳[†]

(西安华为云计算技术有限公司, 陕西 西安 710119)

摘 要: 深度学习中使用的训练的数据集中可能存在一些用户的敏感信息, 在模型训练过程中, 这些敏感信息将会隐性地存在于模型参数中, 从而存在隐私泄露的风险。本文在 AdamP 优化器的基础上引入高斯机制的差分隐私, 提出了一种基于一次幂函数的动态隐私预算分配算法来更合理地分配差分隐私的隐私预算, 即 DP-AdamP, 以更好地平衡隐私性和模型准确性。实验结果表明, 本文的 DP-AdamP 算法相比于传统的 DP-SGD 算法, 在相同隐私预算下具有更好的准确性, 低隐私预算情况下高出约 7.7%, 高隐私预算情况下高出约 3.9%, 并且针对更具有实际意义的低隐私预算的情况进行单独实验, 进一步验证了 DP-AdamP 算法的有效性。

关键词: 深度学习; 差分隐私; 动态隐私预算; AdamP 算法

中图分类号: TP391

文献标识码: A

Dynamic Differential Privacy Preserving Algorithm Based on Deep Learning

HE Jiayang[†]

(Xi'an Huawei Cloud Computing Technology Co., Ltd., Xi'an, Shaanxi 710119, China)

Abstract: There may be some sensitive information of users in the training dataset used in deep learning, and during the model training process, these sensitive information will implicitly exist in the model parameters, thus there is a risk of privacy leakage. In this paper, we introduce differential privacy of Gaussian mechanism on the basis of AdamP optimizer, and propose a dynamic privacy budget allocation algorithm based on primary power function to allocate the privacy budget of differential privacy more reasonably, i. e. DP-AdamP, so as to better balance the privacy and model accuracy. Experimental results show that the DP-AdamP algorithm in this paper has better accuracy than the traditional DP-SGD algorithm under the same privacy budget, about 7.7% higher in the case of low privacy budget, and about 3.9% higher in the case of high privacy budget, and separate experiments are conducted for the more practically relevant case of low privacy budget, which further validates the effectiveness of the DP-AdamP algorithm.

Key words: deep learning; differential privacy; adaptive privacy budget; AdamP algorithm

近些年随着硬件水平的提升以及时时刻刻都在产生的海量数据, 人工智能技术正在逐渐改变人们的生活, 例如学校、地铁等使用的人脸识别装置^[1], 医学诊断中使用人工智能进行辅助诊断^[2],

智能驾驶中使用的驾驶辅助、路况分析等^[3], 都在人们的生活中发挥着重要作用。人工智能算法的实现需要大量的数据作为训练数据集, 这样不可避免地会产生数据隐私泄露的风险, 如人脸识别系统

中的用户人脸数据,医学诊断中的医学图像都属于需要进行保护的隐私数据^[4, 5]。

深度学习中的隐私泄露主要存在于两个方面:数据集的直接泄露、模型参数中隐含隐私信息造成的泄露^[6]。本文的研究主要是用于防止训练完成后的模型被恶意攻击者通过模型提取攻击^[7]或者模型反演攻击^[8]获得数据集中某些用户的隐私信息。面对上述针对模型的攻击方法,差分隐私是一种能够为数据提供强隐私保护的隐私保护方法,尤其针对人脸图像等图像数据可以行之有效去保护用户的隐私安全^[9]。

在 2006 年的时候差分隐私技术就已经被 Dwork 等人^[10]提出,但是提出之后并没有在学界造成什么影响,这是由差分隐私的原理导致的,使用差分隐私技术对隐私数据进行保护时,隐私保护水平和模型准确率呈负相关的关系。本文研究的重点是在使用大规模的图像隐私数据进行模型训练时使用差分隐私技术来解决随机优化器中的隐私泄露问题。Song 等^[11]提出了在 SGD(Stochastic Gradient Descent)优化器中添加差分隐私噪声的算法,这是最初的差分隐私深度学习算法。Pan 等^[12]对该算法进行改进,提出了一种对差分隐私的隐私预算进行动态分配的机制,但是它进行动态分配的过程计算开销较为昂贵。之后李敏等^[13]设计了基于 Adam(Adaptive Moment Estimation)优化器的差分隐私算法,该算法能够改善 SGD 优化器带来的容易陷入局部最优的问题,但是 Adam 优化器额外引入动量会导致尺度不变权重的有效步长的快速减少。

在 2021 年 Heo 等^[14]提出了减慢基于动量的优化器的权重增长的 AdamP(Adaptive Moment Estimation Projection)算法,在本文在 AdamP 算法的基础上设计了 DP-AdamP 算法,其中引入高斯机制的差分隐私,并使用基于一次幂函数进行隐私预算的动态分配。算法在模型训练过程的反向传播中使用了动态隐私预算分配算法,以实现隐私预算随着训练轮数的增加而增加,模型能够较好地进行收敛,该算法会在 2.2 节中详细介绍。在第 3 节中进行对比实验,证明 DP-AdamP 算法相较于 DP-SGD 算法在相同隐私预算下,模型结果具有更好的准确性;在相同的模型准确性下,需要的隐私预算更小,具有更高的隐私保护性。

1 相关算法概述

1.1 差分隐私

ϵ -差分隐私技术最早是由 Dwork 提出的,其中 ϵ 表示隐私预算,差分隐私中有几个相关的定义:

定义 1 邻近数据集:设 D^n 为所在数据集的域,也成为数据集的集合。对于 $D, D' \in D^n$,如果 D 和 D' 只有一个条目是不同的,则称它们为相邻项,即 $\|D - D'\| = 1$ 。

定义 2 灵敏度:给定两个邻近数据集 D 和 D' , f 是数据集上的随机算法, Δf 为 f 的全局灵敏度。 $\|\cdot\|_1$ 表示 l_1 范数。

$$\Delta f = \max_{D, D'} \|f(D) - f(D')\|_1 \quad (1)$$

定义 3 (ϵ, δ) -差分隐私^[15]:对于给定的随机算法 M 在数据集 D, D' 上的任意输出结果 $S = (S \in \text{range}(M))$ 满足下式,则称 M 满足 (ϵ, δ) -差分隐私。

$$\Pr[M(D) \in S] \leq e^\epsilon [M(D') \in S] + \delta \quad (2)$$

式中, $\Pr$ 表示概率, ϵ 表示隐私预算, δ 表示隐私泄露的概率。理论上, ϵ 的值越小,隐私保护程度越高。

实现差分隐私常用的机制有拉普拉斯机制^[16]、高斯机制^[17]及其变体的采样高斯机制^[18]等。本文采用的机制为高斯机制,高斯机制拥有更宽松的定义,在实际情况中更容易实现,尽管允许有一定的隐私泄露风险,但注入的噪声量更少,因此结果的可用性更好。

定义 4 高斯分布:若随机变量 X 服从一个位置参数为 μ 、尺度参数为 σ 的概率分布,且其概率密度函数为:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (3)$$

记作 $X \sim N(\mu, \sigma^2)$ 。

$X \sim N(\mu, \sigma^2)$ 的高斯分布概率密度函数如下:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (4)$$

定义 5 高斯噪声机制:对于任意函数 f ,如果一个算法满足 $A = f(D) + N(0, \Delta f^2 \sigma^2 I)$,其中 Δf 为 $L2$ 全局灵敏度,并且对于任意的 $\epsilon \in (0, 1)$ 有

$$\delta \geq \frac{4}{5} \exp\left\{-\frac{(\epsilon\sigma)^2}{2}\right\}, \sigma \geq \frac{\sqrt{2\ln(1.25/\delta)}\Delta f}{\epsilon}$$

则算法满足 (ϵ, δ) -差分隐私。

1.2 DP-SGD 算法

DP-SGD 算法是在 SGD 算法的基础上引入了差分隐私算法以保护数据隐私。

在 DP-SGD 算法中使用的是非动态的平均分配隐私预算的方法,这种方法在训练后期可能造成模型无法找到最优解的情况。本文使用动态隐私预算分配的算法进行改进。而且 SGD 算法下降速度慢,可能会在沟壑间持续振荡,使用 AdamP 算法则可以很好地避免上述问题。第 2 节将会介绍基于 AdamP 算法设计的 DP-AdamP 算法。

DP-SGD 算法伪代码过程如算法 1 所示。

算法 1:DP-SGD

Input: 样本 $\{x_1, x_2, \dots, x_N\}$, 损失函数 $L(w) = \frac{1}{N} \sum_i L(w, x_i)$ 。学习率 η , 噪声尺度 σ , 分组大小 L , 梯度裁剪范数 C , 随机初始化参数 w_0 , 迭代总数 T , 时间步长为 t 。

Output: w_T

```

while  $t \in [T]$  do
  1) 计算梯度: 对每个样本  $i \in L_t$ , 计算
     $g_t(x_i) \leftarrow \nabla_{w_t} L(w_t, x_i)$ 
  2) 梯度裁剪:
     $\bar{g}_t(x_i) \leftarrow g_t(x_i) / \max\left(1, \frac{\|g_t(x_i)\|_2}{C}\right)$ 
  3) 添加噪声:
     $\tilde{g}_t \leftarrow \frac{1}{L} (\sum_i \bar{g}_t(x_i) + N(0, \sigma^2 C^2 I))$ 
  4) 更新参数:  $w_{t+1} \leftarrow w_t - \eta \tilde{g}_t$ 
End while

```

2 动态隐私预算的 DP-AdamP 算法

该部分的 DP-AdamP 算法是在 AdamP 算法中引入高斯机制的差分隐私,并且提出了动态隐私预算分配算法。AdamP 算法相较于 SGD 算法,引入了一阶矩估计、二阶矩估计以及动量,使得振荡方向的步长很小,梯度小的方向步长变大的同时逐渐遗忘历史梯度,针对动量造成的有效学习率降低的问题,使用投影更新的方式进行改进,防止诱导有效学习率降低的情况发生,之后再对矩估计结果进行修正,避免冷启动。从而解决了 SGD 算法下降速度慢,可能会在沟壑间持续振荡无法找到最优解的缺点。

训练过程中的差分隐私噪声具有较强的随机性,所以训练过程中差分隐私的隐私预算分配是会直接影响到模型训练结果的一个重要因素。结合图像深度学习模型的训练需要,设计了一次幂函数动态隐私预算分配算法对训练过程中的隐私预算

进行动态分配。该分配算法会随着模型训练的进行,不断增加隐私预算的量,从而实现不断减少噪声的添加量,以提高模型准确率。

DP-AdamP 算法伪代码如下算法 2 所示:

算法 2:DP-AdamP

Input: 样本 $\{x_1, x_2, \dots, x_N\}$, 损失函数 $L(w) = \frac{1}{N} \sum_i L(w, x_i)$ 。学习率 η (默认为 0.001), 噪声尺度 σ , 小常数 δ (用于修正矩估计,并且用于高斯机制中), 分组大小 L , 梯度裁剪范数 C , 随机初始化参数 w_0 , 动量 $\beta_1, \beta_2 \in [0, 1]$, 一阶矩估计 v , 二阶矩估计 r , 迭代总数 T , 时间步长为 t 。

Output: w_T

```

while  $t \in [T]$  do
  1) 计算梯度: 对每个样本  $i \in L_t$ , 计算
     $g_t(x_i) \leftarrow \nabla_{w_t} L(w_t, x_i)$ 
  2) 梯度裁剪:
     $\bar{g}_t(x_i) \leftarrow g_t(x_i) / \max\left(1, \frac{\|g_t(x_i)\|_2}{C}\right)$ 
  3) 计算加噪后的平均梯度:
     $\tilde{g}_t \leftarrow \frac{1}{L} (\sum_i \bar{g}_t(x_i) + N(0, \sigma^2 C^2 I))$ 
  4) 计算一阶矩估计:
     $v_t = \beta_1 v_{t-1} + (1 - \beta_1) \tilde{g}_t(x_i)$ 
  5) 计算二阶矩估计:
     $r_t = \beta_2 r_{t-1} + (1 - \beta_2) \tilde{g}_t(x_i) \odot \tilde{g}_t(x_i)$ 
  6) 修正:
     $\hat{v}_t = \frac{v_t}{1 - \beta_1}, \hat{r}_t = \frac{r_t}{1 - \beta_2}$ 
  7) 动量更新:
     $p_t \leftarrow \frac{\hat{v}_t}{\sqrt{\hat{r}_t} + \delta}$ 
  8) 投影更新:
     $q_t = \begin{cases} \prod_{w_t} (p_t), & \text{如果 } \cos[w_t, \hat{g}_t(x_i)] < \frac{\delta}{\sqrt{\dim(w)}} \\ p_t, & \text{其他情况} \end{cases}$ 
  9) 更新参数:
     $w_t \leftarrow w_{t-1} - \eta \cdot q_t$ 
End while

```

下面对上述算法的关键步骤进行解释。

2.1 梯度裁剪

为了让差分隐私充分发挥作用以实现隐私保护,需要对每个样本的模型参数的影响进行限制,以此控制灵敏度,使用 l_2 范数来对模型训练梯度进行约束。

2.2 动态噪声添加

为了解决非动态平均隐私预算分配算法对模型精度的影响,本文在平衡模型准确率和隐私保护性的基础上,设计了一种基于一次幂函数的动态隐私预算分配算法,该算法可以随着训练轮数的增加动态地分配隐私预算,下面会对该算法进行详细介绍,是否还有别的分配算法更加适用于深度学习训

练过程中的隐私预算分配,这将会是之后的研究内容。

该算法将动态的随机噪声加入裁剪后的梯度中,保证黑客无法推断出特定用户的某些数据是否存在于数据集中。此处算法中使用高斯机制产生噪声。加噪后的平均梯度计算式为

$$\tilde{g}_t \leftarrow \frac{1}{L} [\sum_i \bar{g}_t(x_i) + N(0, \sigma^2 C^I)] \quad (5)$$

其中噪声服从均值为 0,标准差为 $C \cdot \sigma$, I 为维度与样本数量以及梯度数量相关的一个单位矩阵,添加的噪声通过这种方式进行矩阵运算。由定义 5 可知标准差:

$$\sigma \geq \frac{\sqrt{2\ln(1.25/\delta)}\Delta f}{\epsilon} \quad (6)$$

这意味着相同的数据集,灵敏度 Δf 是相同的,允许的隐私泄露率 δ 一般也是个很小的固定值,因此隐私预算 ϵ 的大小便是决定模型隐私保护性和模型结果准确性的唯一参数。 ϵ 越小,高斯机制的方差相应的也就越小,高斯分布也就越分散,加入的噪声也就越大,对数据的隐私保护性就越强,但正因加入噪声量大,模型结果准确性也会随之降低;反之, ϵ 越大,高斯机制的方差越小,生成的高斯分布越集中,则加入噪声量越小,对数据隐私保护性就越弱,模型结果的准确性越高。

下面为本文提出的动态隐私预算分配算法,使用一次幂函数作为隐私预算的分配函数,如下所示:

设总隐私预算 ϵ ,迭代总数为 T ,第 $i(i>0)$ 轮所分配的隐私预算为:

$$\epsilon_i = \epsilon \cdot \frac{i}{\sum_{i=1}^T i} \quad (7)$$

假设 $\epsilon=3$, $T=300$,隐私预算取 3 的原因是隐私预算的大小基本不会影响分配曲线的趋势, T 取 300 的原因是观察训练轮数对分配曲线的影响,取一次幂函数是由实验得出,在 3.1 节中进行验证。则隐私预算 ϵ_i 的分配曲线如图 1 所示。

图 1 中纵轴为隐私预算 ϵ 的值,是一个常数,横轴为训练轮数 epoch,单位为轮。从图中可以看出,隐私预算随着训练轮数的增加不断增加,保证其在训练中添加的噪声量不断减少,使得算法能够有足够的时间进行收敛,以达到较高的模型准确率。在第 3 节中会进行对比验证。

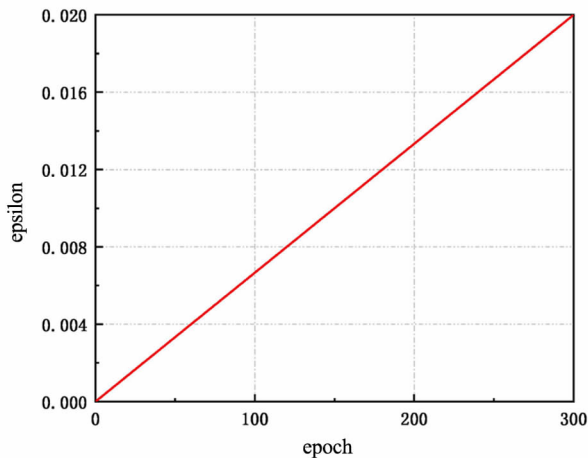


图 1 一次幂函数预算分配曲线

3 实验与结果分析

本文的实验环境为 Linux 操作系统,内存为 32 GB, CPU 为 Xeon Gold 5400, GPU 为 NVIDIA V100,在 Python3.8 编程环境下实现。

本节将会对动态隐私预算分配的 DP-AdamP 算法与非动态隐私预算分配的 DP-SGD 算法在低、中和高的隐私预算下进行对比。并且将对比在 AdamP 算法下不同幂次动态隐私预算分配函数的训练过程。

3.1 数据集介绍

本文实验使用的数据集为机器学习中常用的两种用于分类的数据集 MNIST^[19] 与 CIFAR10^[20],MNIST 数据集总共包含 7 万张图像,其中有 6 万张图像用于模型的训练,剩余 1 万张图片用于验证。该数据集中的图像格式均为 28×28 像素的图像。其中有各种形态的手写数字图像供模型训练算法使用,是机器学习常用的实验数据集。CIFAR10 数据集则包含了 6 万张 32×32 的彩色图像,分为 10 个类别,分别为飞机、汽车、鸟、猫、鹿、狗、蛙、马、船、卡车的数据集,且每个类别有 6 千张图像,其中包含有 5 万张训练图像和 1 万张验证图像。CIFAR10 数据集相较于 MNIST 更为复杂也更贴近现实场景,使用两个数据集能够更好地验证本文算法的有效性。

3.2 不同幂次预算分配算法对比

为比较 DP-AdamP 算法中不同幂次的模型准确率,分别选取了 1~3 三个幂次进行实验,实验数据集为 CIFAR10,训练轮数为 200 轮,隐私预算取 2,隐私泄露率取 10^{-5} ,训练情况如图 2 所示。

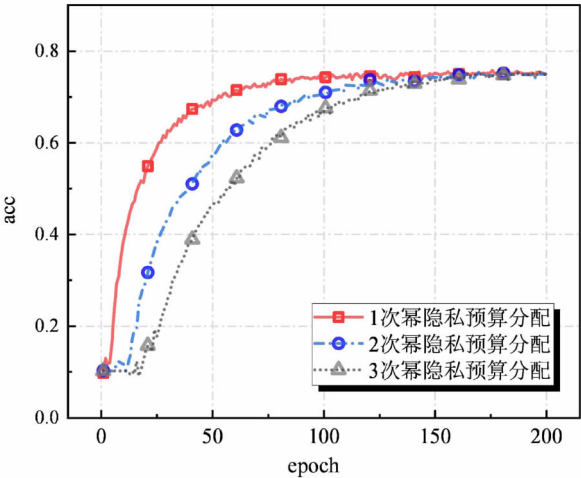


图 2 不同幂次模型训练情况

3.3 DP-AdamP 与 DP-SGD 对比

本实验为 DP-AdamP 算法与 DP-SGD 算法在不同隐私预算下的对比,训练轮数均为 300 轮,隐私泄露率取 10^{-5} ,隐私预算均从 0.5 取到 3,实验结果如图 3、图 4 所示。

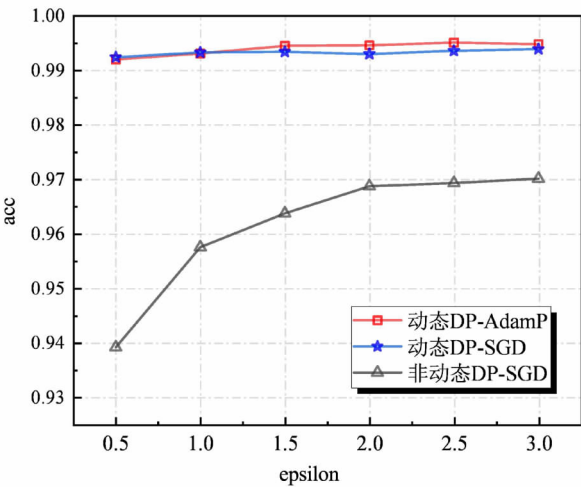


图 3 MNIST 下不同隐私预算结果对比

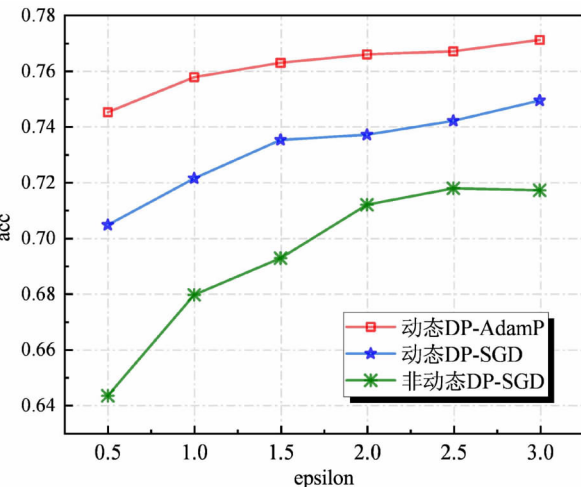


图 4 CIFAR10 下不同隐私预算结果对比

本节实验中对使用幂函数动态分配的 DP-AdamP 算法、动态分配的 DP-SGD 算法以及非动态分配的 DP-SGD 算法进行了对比。从图 3、图 4 中可以发现,无论是简单的 MNIST 还是较为复杂的 CIFAR10,使用动态分配的 DP-AdamP 算法和 DP-SGD 算法的模型准确率都要优于非动态分配的 DP-SGD 算法,验证了动态隐私预算分配算法的可靠性,且在同样使用动态分配的算法中,DP-AdamP 相较于 DP-SGD 又有较高的准确率,验证了 AdamP 算法相较于 SGD 的优势性。

MNIST 数据集上,在 0.5 的低隐私预算的情况下,DP-AdamP 算法和非动态分配的 DP-SGD 算法准确率分别为 99.20%、93.91%,高了 5.29%;3 的高隐私预算下准确率分别为 99.48%、97.01%,也高了 2.47%。CIFAR10 数据集上,0.5 的低隐私预算下,DP-AdamP 算法和非动态分配的 DP-SGD 算法准确率分别为 74.52%、64.32%,高了 10.20%;3 的高隐私预算下,准确率分别为 77.12%、71.71%,也高了 5.41%。

上述实验结果说明 DP-AdamP 算法相较于 DP-SGD 算法有着较高的模型准确率,在保证隐私性的同时提升了模型准确率。并且 DP-AdamP 算法在低隐私预算的情况下模型结果有着远远优于 DP-SGD 算法的准确率。

3.4 低隐私预算下的 DP-AdamP 算法

因为隐私预算较高时,隐私保护性较弱,安全性难以保证,因此重点考虑隐私预算小于 1 的情况,本节在 CIFAR10 数据集上,隐私预算为 0.1~1 之间的隐私预算情况下进行实验,观察 DP-AdamP 算法在低隐私预算下的表现。实验结果如图 5 所示。

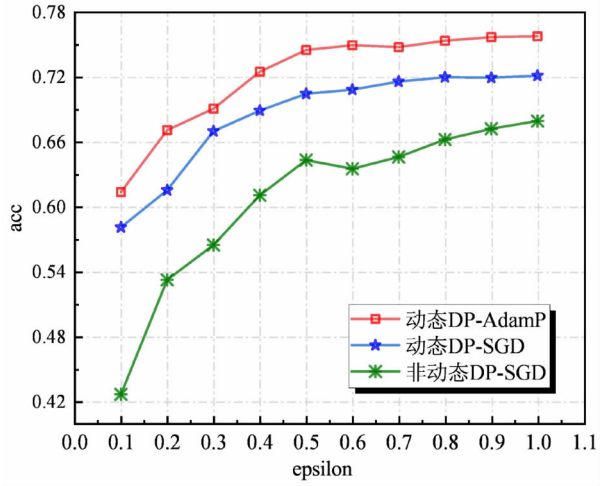


图 5 低隐私预算下 DP-AdamP 与 DP-SGD 结果对比

如图 5 所示,在 0.1 到 1 的低隐私预算的情况下,DP-AdamP 算法相较于 DP-SGD 算法模型准确率也更高,具体模型准确率如表 1 所示。

表 1 低隐私预算下的模型准确率

算法	epsilon=0.1	epsilon=1	变化量
动态 DP-AdamP	61.35%	75.78%	14.43%
动态 DP-SGD	58.1%	72.14%	14.04%
非动态 DP-SGD	42.63%	67.95%	25.32%

由表 1 可知,在隐私预算为 0.1 和 1.0 的情况下,动态 DP-AdamP 算法的模型准确率分别为 61.35%、75.78%,提高了 14.43%;动态 DP-SGD 算法的模型准确率分别为 58.1%、72.14%,提高了 14.04%,非动态 DP-SGD 算法的模型准确率分别为 42.63%、67.95%,提高了 25.32%。通过该实验可以说明,DP-AdamP 算法的在低隐私预算的情况下,相较于同样使用动态隐私预算分配算法的 DP-SGD 可以达到更高的准确率,并且动态 DP-SGD 算法相较于非动态的 DP-SGD 算法在隐私预算从 0.1 提高到 1 的过程中模型准确率波动较小,说明动态隐私预算分配算法有着较好的抗干扰能力,在低隐私预算的情况下也能获得较好的模型准确率。

4 结 论

本文在 AdamP 算法中引入高斯机制的差分隐私,并根据差分隐私的隐私预算分配问题,设计了一种基于一次幂函数的动态隐私预算分配算法,这种算法可以随着训练轮数的迭代动态地对隐私预算进行分配,从而平衡模型训练中隐私性和准确性。相较于非动态分配的 DP-SGD 算法,本文算法克服了 SGD 算法和非动态隐私预算分配的缺陷,进一步地提高了模型准确率,并且用实验对比证明 DP-AdamP 算法相较于 DP-SGD 算法具有较好的隐私性和模型准确率。

参考文献

[1] 黄帅凤, 汤丽娟, 梁龙兵. 一种改进的 LeNet-5 嵌入式人脸识别方法[J]. 计算技术与自动化, 2021, 40(4): 91-96.

[2] MANICKAM P, MARIAPPAN S A, MURUGESAN S M, et al. Artificial intelligence (AI) and internet of medical

things (IoMT) assisted biomedical systems for intelligent healthcare[J]. Biosensors, 2022, 12(8): 562.

[3] YANG M. Research on vehicle automatic driving target perception technology based on improved MSRPN algorithm[J]. Journal of Computational and Cognitive Engineering, 2022, 1(3): 147-151.

[4] KAISSIS G A, MAKOWSKI M R, RÜCKERT D, et al. Secure, privacy-preserving and federated machine learning in medical imaging[J]. Nature Machine Intelligence, 2020, 2(6): 305-311.

[5] LI T, LIN L. Anonymousnet: natural face de-identification with measurable privacy[C]. Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, USA. IEEE, 2019: 56-65.

[6] LIU X M, XIE L H, WANG Y P, et al. Privacy and security issues in deep learning: A survey[J]. IEEE Access, 2020, 9: 4566-4593.

[7] GONG X L, WANG Q, CHEN Y J, et al. Model extraction attacks and defenses on cloud-based machine learning models[J]. IEEE Communications Magazine, 2020, 58(12): 83-89.

[8] REN H C, DENG J J, XIE X H. Grnn: generative regression neural network—a data leakage attack for federated learning[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2022, 13(4): 1-24.

[9] EL OUADRHIRI A, ABDELHADI A. Differential privacy for deep and federated learning: a survey[J]. IEEE Access, 2022, 10: 22359-22380.

[10] DWORK C, MCSHERRY F, NISSIM K, et al. Calibrating noise to sensitivity in private data analysis[C]. Proceedings of the Theory of Cryptography Conference. Springer, 2006: 265-284.

[11] SONG S, CHAUDHURI K, SARWATE A D. Stochastic gradient descent with differentially private updates[C]. Proceedings of the 2013 IEEE Global Conference on Signal and Information Processing. IEEE, 2013: 245-248.

[12] PAN K, GONG M G, FENG K Y, et al. Differentially private regression analysis with dynamic privacy allocation[J]. Knowledge-Based Systems, 2021, 217: 106795.

[13] 李敏, 李红娇, 陈杰. 差分隐私保护下的 Adam 优化算法研究[J]. 计算机应用与软件, 2020, 37(6): 253-258+296.

[14] HEO B, CHUN S, OH S J, et al. AdamP: Slowing down the slowdown for momentum optimizers on scale-invariant weights[J]. arXiv preprint arXiv:200608217, 2020.

[15] DWORK C. A firm foundation for private data analysis[J]. Communications of the ACM, 2011, 54(1): 86-95.

[16] SARATHY R, MURALIDHAR K. Evaluating Laplace noise addition to satisfy differential privacy for numeric data[J]. Transactions on Data Privacy, 2011, 4(1): 1-17.

[17] DWORK C, ROTH A. The algorithmic foundations of differential privacy[J]. Foundations and Trends in Theoretical Computer Science, 2014, 9(3-4): 211-407.

[18] MIRONOV I, TALWAR K, ZHANG L. R'enyi differential privacy of the sampled gaussian mechanism[J]. arXiv preprint arXiv:190810530, 2019.

[19] LECUN Y. The MNIST database of handwritten digits[EB/OL]. <http://yann.lecun.com/exdb/mnist/>, 1998.

[20] LI H M, LIU H C, JI X Y, et al. Cifar10-dvs: an event-stream dataset for object classification[J]. Frontiers in Neuroscience, 2017, 11: 309.