DataLens: Scalable Privacy Preserving Training via Gradient Compression and Aggregation

0. Abstract

提出

a scalable privacy-preserving generative model DATAENS---->在给出敏感 input data 情况下，以差分隐私DP 的技术，生成 synthetic data。

可以在保护 private information 的情况下，用生成的数据训练模型。

此外，利用：

- the generative adversarial networks (GAN)
- PATE framework

训练多个 discriminators 作为“teacher” model，来利用 gradient vectors 投票，来保证隐私。

Standard PATE privacy preserving framework

允许对于 one-dimensional predictions 进行投票

但对于高纬度的预测不太可行

dimension reduction techniques tradeoff

（1）the improvement of privacy preservation

（2）the slowdown of SGD convergence.

提出

- dimension compression
- aggregation approach TopAGG

结合了 top-k dimension compression 和对应的 noise injection mechanism.

证明了

- DataLens framework 对于 generated data 保证了 differential privacy
- 分析了 convergence

Experiments

- MNIST
- Fashion-MNIST
- high dimensional CelebA and Place365

证明了 DataLens显著地优于 baseline DP data generative models.

1. Introduction

machine learning concerns:

large privacy sensitive information，例如人脸、医疗记录等，这可能在 machine learning models 训练过程中被泄露。

Differential private

给 clipped gradient 添加 Gaussian noise

缺点：Decrease the learning utility.

Semi supervised learning framework PATE

利用的是在 private datasets 上面训练的 teacher model 的 noise 聚合，在 privacy noise情况下提高了 learning effectiveness.

缺点：从 discriminative model 到 generative model 来保证了 generated data 是 differential privacy，这对于给定的 high-dimensional gradient aggregation 是 non-trival.

Improve the flexibility of differentially private

提高 DP generated model 的灵活性

设计了 data generator 和 generated data 都是 differentially private 的，而不仅仅只有 predictions 是 differentially private 的。

这样生成的数据可以被用来训练任意的model tasks.

generative adversarial networks (GAN)

generative adversarial networks (GAN) 可以生成高质量的数据。

现有的 works 只能生成 low dimensional data，有着 weak privacy guarantees.($(\epsilon, \delta) - DP$ with small ϵ)

noisy compression schemes

例如 只保留 gradient 的 top-K elements，可以实现统计上的相似 convergence

noisy compression schemes 引入的 noise 可以与传统的 DP noise mechanism结合来保护 privacy.

这就使得允许使用 fewer noise 来实现相同水平的 DP protection.

Differentially private data generative model DataLens

基于PATE framework 提出了 differential private data generative model DataLens。

DataLens 训练了多个 discriminators 作为不同的 teacher models，来以 differentially private way 给 student generator 提供 back-propagation information

TopAGG

解决 high-dimensional data problem ，提出了an effective noisy gradient compression and aggregation strategy TopAgg

每个 discriminator 去 vote 几个最高的 gradient 的维度，然后聚合 noisy gradient sign，来执行 back-propagation.

证明了：对于 DataLens 的 data generator 和 generated data 提供了 differential privacy.

提供了：对于 gradient compression 和 aggregation strategy 的理论 convergence analysis

结合了：

- coordinate-wise gradient clipping
- gradient compression
- DP noise mechanism

Technical Contributions

- 提出DataLens，可以在有限的 privacy budgets情况下，生成 high-dimensional image data.
- 证明了 privacy guarantees，分析了 DataLens 的 convergence.
- 提出 noisy gradient compression and aggregation 算法， TopAGG。结合了 top-k dimension compression 以及 noise injection.
- 证明了 differential privacy and convergence 的 tradeoff
- 在4个数据集上对DataLens进行了评估，DataLens显著地优于其他的 generative models.

2. Preliminaries

2.1 Differential Privacy

Definition 1 ((ε, δ)-Differential Privacy [16]). A randomized algorithm $\mathcal{M}$ with domain $\mathbb{N}^{|\mathcal{X}|}$ is (ε, δ)-differentially private if for all $\mathcal{S} \subseteq \text{Range}(\mathcal{M})$ and for any neighboring datasets D and D':

$$\Pr[\mathcal{M}(D) \in \mathcal{S}] \leq \exp(\varepsilon) \Pr[\mathcal{M}(D') \in \mathcal{S}] + \delta.$$

PATE framework 通过从几个 在 private data 上训练的teacher models聚合 prediction votes，来实现DP

2.2 Data Generative Models

从类似的 data distribution 中生成 diverse datasets。

可被用于 data augmentation(数据增加)

GAN

generator： Ψ learns to generate synthetic records,

discriminator： 被训练，用来区分 real words 和 fake ones

给出： input x， sampled noise z

训练

discriminator Γ，识别从 generator 生成的 example与真实分布之间的 loss function 最大似然

$$\iota_{\Gamma} = -log\Gamma(x) - log(1 - \Gamma(\Psi(z)))$$

generator Ψ: 最小化生成的数据被 discriminator 识别为 fake 的概率

$$\iota_{\Psi} = -log\Gamma(\Psi(z))$$

GAN 的 generative models 易于泄露训练数据的信息

提出了 DP generative models，可以在保护 sensitive training information 的同时，生成无限制数量的 high-utility data

2.3 Gradient Compression

previous research： quantization、low-rank approximation and sparsification。

通过压缩 gradient，能够在没有显著减小 convergence 的情况下，减小 gradient 的维度。

这可以转换为需要在 DP 中添加的 fewer noise

3. THREAT MODEL & METHOD OVERVIEW

(1) threat model

(2) DataLens framework as a differentially private data generative model.

(3) noisy gradient compression and aggregation method TopAGG.

3.1 Threat Model and Goal

- attacker 可以 train some shadow models 来推断训练的 ”membership“
- attacker 可以通过 data recovery attacks 恢复训练数据

Differential privacy DP

- protect membership inference attacks
- protect training-data memorization

Differentially private data generative model

设计了differentially private data generative model, 确保了 generated data 而不是 model 的 parameter是 differential private。

因此，只要数据生成，就可以在有 differential privacy 保证的情况下，用于别的训练任务

The generated data is of high utility

提供了 visual quality 的 evaluation。

3.2 Method Overview

The goal of DataLens：生成 high-dimensional data，不会在训练中泄露隐私信息。

an overview for the structure of DataLens

结合了：

- TopAGG 对于high dimension DP
- aggregation with GAN
- the PATE framework

DataLens consists:

- 一系列 teacher discriminators (随机 access划分的 non-overlapping sensitive training data)
- a student generator
- TopAGG 对于 high-dimensional DP gradient compression and aggregation

TopAGG consists:

- top-k and sign gradient compression
- DP gradient aggregation for high-dimensional sparse gradients

achieve high data utility，算法需要保护teacher models 的正确的 gradient directions

privacy guarantee， high-dimensional gradient vector 通常花费了较高的 privacy budget， 导致了更弱的 privacy guarantee。

DataLens 解决上述问题的方法是使用 TopAGG算法，对于high-dimensional DP gradient compression and aggregation。

TopAGG

- takes the top-k entries in a gradient vector
- compresses them via stochastic sign gradient quantization
- perform DP gradient aggregation over the sign gradient vectors with a corresponding noise injection mechanism

4 DATALENS: SCALABLE PRIVACY PRESERVING GENERATIVE MODEL

- present DataLens
- analysis on privacy guarantee and convergence
- privacy-utility trade-off
- TopAGG 从 DataLens 到 standard SGD training

4.1 DataLens Training

DataLens algorithm:

- ensemble of teacher discriminators.
- a student generator
- a DP gradient aggregator TopAGG

TopAGG algorithm:

- a top-k gradient compression
- a DP gradient aggregation

Training DP Generator via Teacher Discriminator Aggregation.

teacher discriminator:

在 non-overlapping sensitive data partitions 训练，来区分真实数据和合成数据。

student generator:

生成合成的 records，发送给 teachers 来查询 labels

DP gradient aggregator:

收集 teachers' gradient vectors，并且添加 DP noise

更新 **student generator and teacher discriminator**:

1. Training teacher discriminators
    - student generator 生成 a batch 的合成数据
    - 每个 teacher discriminator 根据减小真实数据和合成数据的loss值来更新 weights，
2. Generating and compressing teacher gradient vectors
    - 每个 teacher discriminator 计算 a gradient vector $g^{(i)}$
    - 根据 gradient vector，student generator 提升合成数据的实用性
3. DP gradient compression and aggregation
    - TopAGG 压缩 teacher gradient vectors，然后 aggregate
    - 对 teachers' gradient 进行聚合的同时，执行了 noise injection
4. Training the student generator
    - 通过 back-propagation 聚合的 DP gradient vectors，提升合成数据的实用性。
    - 定义 student generator loss function:

      $\hat{\iota}_{\Psi(z,\hat{x})} = \frac{1}{m}\sum_{j=1}^{m}(\Psi(z_j) - \hat{x}_j)^2$

      z_j 是 noise sample，$\Psi(z_j)$是合成数据

      $\hat{x}_j = \Psi(z_j) + \gamma\bar{g}_j$是合成数据+聚合的 DP gradient vectors

因为

更新的时候，只需要找到使得DP gradient vectors变小即可

Top-k Gradient Compression via Stochastic Sign Gradient

teacher model

压缩 real-valued gradient vector 到 k 个非零项的sparse sign vector

gradient compression function: TopkStoSignGrad (g, c, k)

1. 对于每个 gradient g，选取 top-k dimensions，设置其余 dimensions 为 0，the j-th dimension 是 $\hat{g}_j$
2. 对于每个 dimension，使用 threshold c，clip the gradient $\hat{g}_j = min(max(\hat{g}_j, -c), c)$

   normalize top-k gradient vector，执行随机梯度 sign quantization。 $\tilde{g}_j$ 是$\hat{g}_j$ 的无偏估计

$$\tilde{g}_j = \begin{cases} 1, & with\ probability\ \frac{1+\hat{g}_j}{2}; \\ -1, & with\ probability\ \frac{1-\hat{g}_j}{2}. \end{cases} \tag{1}$$

通过上述过程，转换一个 real-valued gradient vector 到 一个 sparsified {-1,0,1} value vector。

High Dimensional DP Gradient Aggregation

compression 之后，每个 gradient vector 是一个 k个非零项的 sparse sign vector

每个 teacher votes k 个 gradient dimensions，vote 要么是 positive direction $\tilde{g}_j = 1$，要么是 negative direction $\tilde{g}_j = -1$

Gaussian mechanism with post-processing thresholding.

1. sum of the gradient vectors，注入 Gaussian noise $N(0, \sigma^2)$
2. 检查每个 gradient direction 的 noisy vote是否大于 threshold。保证了只选取高 agreement 的directions

TopAGG

1. Top-k stochastic sign gradient quantization
2. DP gradient aggregation

Algorithm 3 - Differentially Private Gradient Compression and Aggregation (TopAgg). This algorithm takes gradients of teacher models and returns the compressed and aggregated differentially private gradient vector.

Input: Teacher number N, gradient vectors of teacher models $\mathcal{G} = \{\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(N)}\}$, gradient clipping constant c, top-k, noise parameters σ, voting threshold β

▷ *Phase I: Gradient Compression*

for each teacher's gradient $\mathbf{g}^{(i)}$ **do**

  $\tilde{\mathbf{g}}^{(i)} \leftarrow \texttt{TopkStoSignGrad}(\mathbf{g}^{(i)}, c, k)$

end for

▷ *Phase II: Differential Private Gradient Aggregation*

$\tilde{\mathbf{g}}^{*} \leftarrow \sum_{i=1}^{N} \tilde{\mathbf{g}}^{(i)} + \mathcal{N}(0, \sigma^2)$

▷ *Phase III: Gradient Thresholding (Post-Processing)*

for each dimension $\tilde{g}^{*}_{j}$ of $\tilde{\mathbf{g}}^{*}$ **do**

$$\bar{g}_j = \begin{cases} 1, & \text{if} \quad \tilde{g}^{*}_{j} \geq \beta N; \\ -1, & \text{if} \quad \tilde{g}^{*}_{j} \leq -\beta N; \\ 0, & \text{otherwise.} \end{cases}$$

end for

Return: $\bar{\mathbf{g}}$

4.2 Differential Privacy Analysis for DataLens

Rényi Differential Privacy.

Theorem 1 (From RDP to DP [41]). *If a mechanism $\mathcal{M}$ guarantees (λ, α)-RDP, then $\mathcal{M}$ guarantees $(\alpha + \frac{\log 1/\delta}{\lambda - 1}, \delta)$-differential privacy for any $\delta \in (0, 1)$.*

Data-Independent Privacy Bound

gradient aggregation algorithm 保留了 DP 或者 RDP，则基于 post-processing， 同样适用于 student generator

$$\widetilde{G} = (\tilde{g}^{(1)}, \ldots, \tilde{g}^{(N)})$$

$\tilde{g}^{i}$ 是第 i 个 teacher 压缩后的 gradient sum aggregation function

$$f_{sum}(\widetilde{G}) = \sum_{i=1}^{N} \tilde{g}^{(i)}$$

applying Gaussian mechanism

$$\widetilde{G}_{\sigma f_{sum}}(\widetilde{G}) = f_{sum}(\widetilde{G}) + N(0, \sigma^2) = \sum_{\tilde{g} \in \widetilde{G}} \tilde{g} + N(0, \sigma^2)$$

Gaussian 机制提供了如下的 RDP guarantee:

Theorem 2 (RDP Guarantee for Gaussian Mechanism [41]). *If f has ℓ_2-sensitivity s, then the Gaussian mechanism $\mathbf{G}_\sigma f$ satisfies $(\lambda, s^2\lambda/(2\sigma^2))$-RDP.*

Theorem 3. *The TopAgg algorithm (Algorithm 3) guarantees $(\frac{2k\lambda}{\sigma^2} + \frac{\log 1/\delta}{\lambda-1}, \delta)$-differential privacy for all $\lambda \geq 1$ and $\delta \in (0,1)$.*

Data-Dependent Privacy Bound

data-dependent RDP bound for randomized algorithms

Theorem 4 (Data-Dependent RDP Bound [44]). *Let $\mathcal{M}$ be a randomized algorithm with (μ_1, α_1)−RDP and (μ_2, α_2)−RDP guarantees and suppose that there exists a likely outcome $\bar{\mathbf{g}}^*$ given a dataset D and a bound $\tilde{q} \leq 1$ such that $\tilde{q} \geq \Pr[\mathcal{M}(D) \neq \bar{\mathbf{g}}^*]$. Additionally,*

suppose that $\lambda \leq \mu_1$ and $\tilde{q} \leq e^{(\mu_2-1)\alpha_2} / \left(\frac{\mu_1}{\mu_1-1} \cdot \frac{\mu_2}{\mu_2-1}\right)^{\mu_2}$. Then, for any neighboring dataset D' of D, we have:

$$D_\lambda\left(\mathcal{M}(D) \| \mathcal{M}\left(D'\right)\right) \leq \frac{1}{\lambda-1} \log\Big((1-\tilde{q}) \cdot \boldsymbol{A}\left(\tilde{q}, \mu_2, \alpha_2\right)^{\lambda-1} + \tilde{q} \cdot \boldsymbol{B}\left(\tilde{q}, \mu_1, \alpha_1\right)^{\lambda-1}\Big),$$

where

$$\boldsymbol{A}\left(\tilde{q}, \mu_2, \alpha_2\right) \triangleq (1-\tilde{q}) / \left(1-\left(\tilde{q} e^{\alpha_2}\right)^{\frac{\mu_2-1}{\mu_2}}\right), \qquad \boldsymbol{B}\left(\tilde{q}, \mu_1, \alpha_1\right) \triangleq e^{\alpha_1} / \tilde{q}^{\frac{1}{\mu_1-1}}.$$

data-independent privacy bound can achieve better utility with the aggregation and thresholding steps in TopAgg

Theorem 5. *For any* $\bar{\mathbf{g}}^* \in \{0,1\}^d$, *we have*

$$\Pr[\mathcal{M}(\tilde{\mathcal{G}}, N, \beta) \neq \bar{\mathbf{g}}^*] = 1 - \prod_{\{j|\bar{g}_j^*=1\}} \left(1 - \Phi\left(\frac{\beta N - f_j}{\sigma}\right)\right)$$

$$\prod_{\{j|\bar{g}_j^*=-1\}} \Phi\left(\frac{\beta N - f_j}{\sigma}\right) \prod_{\{j|\bar{g}_j^*=0\}} \mathrm{erf}\left(\frac{\beta N - f_j}{\sqrt{2}\sigma}\right)$$

where Φ *is the cumulative distribution function of the normal distribution,* erf *is the error function, and* f_j *is the* j*-th dimension of the gradient vector sum* $\sum_{i=1}^{N} \tilde{\mathbf{g}}^{(i)}$ *without the noise injection.*

5. EXPERIMENTAL EVALUATION

5.1 Experimental Setup

与 3 个最新的 Baselines，比较生成数据的实用性：

DP-GAN，PATE-GAN，GS-WGAN，G-PATE 在4 个 image datasets.

Datasets

high dimensional image datasets

MNIST，Fashion-MNIST

- grayscale images of 28 * 28 dimensions.
- 60,000 training examples
- 10,000 testing examples

CelebA datasets

- 202,599 color images of celebrity faces.
- 64*64 * 3
- CelebA-Gender 是二元分类，gender 是 label

  CelebA-Hair 使用3种颜色作为属性分类label

Places365 dataset

- 1.8 M high resolution color images of categories.
- 64 * 64 * 3

Models

Dimensional

50-dimensional for MNIST

64-dimensional ($\epsilon = 10$) for Fashion-MNIST

100-dimensional for CelebA

100-dimensional for Places365

$\epsilon = 1$

- top-k = 200

  MNIST and Fashion-MNIST

- top-k = 700

  CelebA and Places365

$\epsilon = 10$

- top-k=350

  MNIST and Fashion-MNIST

- top-k = 500

  CelebA

- top-k = 700

  Places365

Evaluation Metrics

评估：

1. data utility

   test accuracy 指示了合成数据的实用性

2. visual quality

   Inception Score(IS)

   Frechet Inception Distance(FID)

5.2 Experimental Results

Data Utility Evaluation

在两种隐私预算设置的情况下比较DataLens 和4个baselines

- $\epsilon = 1, \delta = 10^{(-5)}$
- $\epsilon = 10, \delta = 10^{(-5)}$

DataLens 优于所有的 baselines

尤其是 $\epsilon = 1$ 时候效果最好

GS-WGAN 对于 MNIST and Fashion-MNIST只有在 $\epsilon = 10$ 时候可以收敛

Table 1: Performance of different differentially private data generative models on Image Datasets: Classification accuracy of the model trained on the generated data and tested on real test data under different ε ($\delta = 10^{-5}$).

Dataset \ Methods	DC-GAN ($\varepsilon = \infty$)	ε	DP-GAN	PATE-GAN	G-PATE	GS-WGAN	DataLens
MNIST	0.9653	$\varepsilon = 1$	0.4036	0.4168	0.5810	0.1432	**0.7123**
		$\varepsilon = 10$	0.8011	0.6667	**0.8092**	0.8075	0.8066
Fashion-MNIST	0.8032	$\varepsilon = 1$	0.1053	0.4222	0.5567	0.1661	**0.6478**
		$\varepsilon = 10$	0.6098	0.6218	0.6934	0.6579	**0.7061**
CelebA-Gender	0.8149	$\varepsilon = 1$	0.5330	0.6068	0.6702	0.5901	**0.7058**
		$\varepsilon = 10$	0.5211	0.6535	0.6897	0.6136	**0.7287**
CelebA-Hair	0.7678	$\varepsilon = 1$	0.3447	0.3789	0.4985	0.4203	**0.6061**
		$\varepsilon = 10$	0.3920	0.3900	0.6217	0.5225	**0.6224**
Places365	0.7404	$\varepsilon = 1$	0.3200	0.3238	0.3483	0.3375	**0.4313**
		$\varepsilon = 10$	0.3292	0.3796	0.3883	0.3725	**0.4875**

Evaluation under small privacy budget

privacy budget 越小， protection guarantees 越大

随着 ϵ 的增大，不同的 DP models 逐渐收敛，并且 accuracy 增加

Table 2: Performance Comparison of different differentially private data generative models on Image Datasets under small privacy budget which provides strong privacy guarantees ($\varepsilon \leq 1, \delta = 10^{-5}$).

ε	MNIST					Fashion-MNIST				
	DP-GAN	PATE-GAN	G-PATE	GS-WGAN	DataLens	DP-GAN	PATE-GAN	G-PATE	GS-WGAN	DataLens
0.2	0.1104	0.2176	0.2230	0.0972	**0.2344**	0.1021	0.1605	0.1874	0.1000	**0.2226**
0.4	0.1524	0.2399	0.2478	0.1029	**0.2919**	0.1302	0.2977	0.3020	0.1001	**0.3863**
0.6	0.1022	0.3484	0.4184	0.1044	**0.4201**	0.0998	0.3698	0.4283	0.1144	**0.4314**
0.8	0.3732	0.3571	0.5377	0.1170	**0.6485**	0.1210	0.3659	0.5258	0.1242	**0.5534**
1.0	0.4046	0.4168	0.5810	0.1432	**0.7123**	0.1053	0.4222	0.5567	0.1661	**0.6478**

Visual Quality Evaluation

Table 3: Quality evaluation of images generated by different differentially private data generative models on Image Datasets: we use Inception Score (IS) to measure the visual quality of the generated data under different ε ($\delta = 10^{-5}$).

Dataset	Real data	ε	DP-GAN	PATE-GAN	G-PATE	GS-WGAN	DataLens
MNIST	9.86	1	1.00	1.19	3.60	1.00	**4.37**
		10	1.00	1.46	5.16	**8.59**	5.78
Fashion-MNIST	9.01	1	1.03	1.69	3.41	1.00	**3.93**
		10	1.05	2.35	4.33	**5.87**	4.58
CelebA	1.88	1	1.00	1.15	1.11	1.00	**1.18**
		10	1.00	1.16	1.12	1.00	**1.42**

5.3 Ablation Studies

- the data-dependent and data-independent privacy bounds
- the hyper-parameter impacts
- the comparison with different gradient compression methods

Data-Independent Bound v.s. Data-Dependent Bound

在每个 training epoch 情况下的，privacy budget consumption

data-independent bound is always tighter than the data-dependent

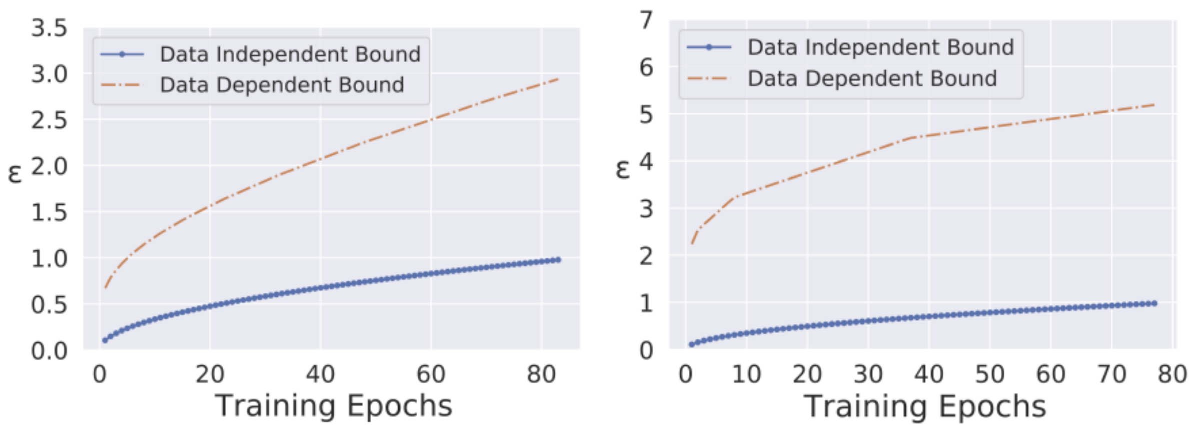

Figure 2: Ablation studies on the data dependent bound v.s. data independent bound on MNIST (left) and CelebA-Hair (right). The data independent bound always yields tighter privacy bound than the data dependent analysis, given high dimensionality of gradients.

Ablation studies on hyper-parameters

- teachers 越多，Performance 越好，save privacy budgets.
- top-k 越小的时候，model 收敛越慢，而且会收敛到一个 bad solution

top-k越大的时候，引入了更大的 DP noise，模型回达到隐私预算的限制

- threshold 越大，容易忽视 top-k voted gradient information.
- clipping value c 越小，容易有更好的收敛性，以及数据实用性。

(a) Hyper-parameters Search for MNIST and Fashion-MNIST

	Top-k			# of Teachers		
	100	200	300	2000	3000	4000
MNIST	0.5889	**0.7123**	0.6753	0.5841	0.7061	**0.7123**
Fashion	0.5738	**0.6478**	0.6088	0.5608	0.5952	**0.6478**
β	0	0.1	0.3	0.5	0.7	0.9
MNIST	0.6361	0.6450	0.6890	0.6921	**0.7123**	0.6956
Fashion	0.5859	0.6103	0.6060	0.6122	0.6213	**0.6478**

但是 teachers 越多，每个 discriminator 分到的 training data 就越少，导致了worse performance (如CelebA-Gender)

(b) Hyper-parameters Search for CelebA-Hair and CelebA-Gender

	Top-k			# of Teachers		
	500	700	900	4000	6000	8000
CelebA-Gender	0.6922	**0.7058**	0.6811	0.6378	**0.7058**	0.6936
CelebA-Hair	0.5792	**0.6061**	0.5769	0.5669	0.5835	**0.6061**
β	0.5	0.6	0.7	0.8	0.85	0.9
CelebA-Gender	0.6440	0.6789	0.6922	0.6861	**0.7058**	0.6381
CelebA-Hair	0.4957	0.5669	0.5612	0.6022	0.5835	**0.6061**

Ablation Studies on the Gradient Compression Methods

D^2P-Fed and FetchSGD

Table 5: Accuracy Comparison of different gradient compression methods (TopAgg, D^{2}P-Fed, FetchSGD). We report the test classification accuracy of models trained with data generated with each technique under $\varepsilon = 1$ and $\delta = 10^{-5}$.

Dataset \ Methods	TopAgg	D^{2}P-Fed	FetchSGD
MNIST	**0.7123**	0.1424	0.6935
Fashion-MNIST	**0.6478**	0.1667	0.6387
CelebA-Gender	**0.7058**	0.4445	0.6552
CelebA-Hair	**0.6061**	0.2893	0.4926

Runtime Analysis

Table 6: Running Time Comparison of different gradient compression methods (TopAgg, D^{2}P-Fed, FetchSGD). We report the average training time per epoch on different datasets under $\varepsilon = 1$ and $\delta = 10^{-5}$.

Dataset \ Methods	TopAgg	D^{2}P-Fed	FetchSGD
MNIST	338.34 s	492.43s	785.34 s
Fashion-MNIST	**340.84s**	471.02s	775.35s
CelebA-Gender	**1196.60s**	3683.22s	2622.40s
CelebA-Hair	**1120.59s**	8092.50 s	2620.63s

Ablation Studies on the Impact of Different Components in DataLens

components:

(1) top-k,

(2) stochastic gradient quantization

(3) gradient thresholding

Table 7: Ablation studies on the impact of different components of DataLens pipeline on Image Datasets: We report the test classification accuracy of models trained with data generated based on different variants of DataLens under $\varepsilon = 1, \delta = 10^{-5}$. The first row of each data groups presents the performance of DataLens.

Dataset \ Component	Top-k	Stochastic Quantization	Aggregation Thresholding	Accuracy
MNIST	✓	✓	✓	**0.7123**
	✗	✓	✓	0.5170
	✓	✗	✓	0.6741
	✓	✓	✗	0.6361
Fashion-MNIST	✓	✓	✓	**0.6478**
	✗	✓	✓	0.4775
	✓	✗	✓	0.6159
	✓	✓	✗	0.5859
CelebA-Gender	✓	✓	✓	**0.7058**
	✗	✓	✓	0.6134
	✓	✗	✓	0.6889
	✓	✓	✗	0.6860
CelebA-Hair	✓	✓	✓	**0.6061**
	✗	✓	✓	0.3318
	✓	✗	✓	0.5325
	✓	✓	✗	0.5504

6. RELATEDWORK

DP Generative Models

现有的 works 被证明能在 Low dimensional datasets 有较好的 performance

但是要么是 low data utility，要么是 high sampling complexity.

应用 DP-SGD 到 GAN，DPGAN 通过给 discriminator gradients 添加 Gaussian noise

DP-CGAN，GS-WGAN 当应用到 high-dimensional datasets，由于 privacy budget 的限制，仍然存在 low data utility

PATE-GAN

结合了 PATE framework 与 GAN。

训练多个 teacher discriminators，并更新 student discriminators。

under limited privacy budget

G-PATE 直接使用 teacher 训练 student，使用 random projection 减小了 gradient dimension

DataLens 在 high dimensional 显著提高了 utility

DP SGD Training

应用 DP 到 SGD

Gradient Compression

DataLens 使用 PATE framework，应用了 sign compression 作为 teacher voting to save privacy budget.

FetchSGD 提出 CountSketch data structure and top-k operation

但是 FetchSGD 缺少 privacy guarantee

TopAGG 结合了 stochastic sign 和 top-k gradient compression

7. Conclusion

- DataLens 应用于 high dimensional data
- TopAGG执行 gradient compression and aggregation
- DP analysis and convergence analysis
- DataLens outperforms 其他的DP generative models，尤其是在 high dimensional 或者 limited privacy budget