Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.30, NO.5, Oct. 2020 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2020.30.5.945

차분 프라이버시를 만족하는 안전한 GAN 기반 재현 데이터 생성 기술 연구*

강 준 영,[1†] 정 수 용,[1] 홍 도 원,[2‡] 서 창 호[2]
[1,2]공주대학교 (대학원생, 교수)

A Study on Synthetic Data Generation Based Safe Differentially Private GAN*

Junyoung Kang,[1†] Sooyong Jeong,[1] Dowon Hong,[2‡] Changho Seo[2]
[1,2]Kongju National University (Graduate student, Professor)

요 약

많은 응용프로그램들로부터 양질의 서비스를 제공받기 위해서 데이터 공개는 필수적이다. 하지만 원본 데이터를 그대로 공개할 경우 개인의 민감한 정보(정치적 성향, 질병 등)가 드러날 위험이 있기 때문에 원본 데이터가 아닌 재현 데이터를 생성하여 공개함으로써 프라이버시를 보존하는 많은 연구들이 제안되어왔다. 그러나 단순히 재현 데이터를 생성하여 공개하는 것은 여러 공격들(연결공격, 추론공격 등)에 의해 여전히 프라이버시 유출 위험이 존재한다. 본 논문에서는 이러한 민감한 정보의 유출을 방지하기 위해, 재현 데이터 생성 모델로 주목받고 있는 GAN에 최신 프라이버시 보호 기술인 차분 프라이버시를 적용하여 프라이버시가 보존되는 재현 데이터 생성 알고리즘을 제안한다. 생성 모델은 레이블이 있는 데이터의 효율적인 학습을 위해 CGAN을 사용하였고, 데이터의 유용성 측면을 고려하여 기존 차분 프라이버시보다 프라이버시가 완화된 Rényi 차분 프라이버시를 적용하였다. 그리고 생성된 데이터의 유용성에 대한 검증을 다양한 분류기를 통해 실시하고 비교분석하였다.

ABSTRACT

The publication of data is essential in order to receive high quality services from many applications. However, if the original data is published as it is, there is a risk that sensitive information (political tendency, disease, ets.) may reveal. Therefore, many research have been proposed, not the original data but the synthetic data generating and publishing to privacy preserve. but, there is a risk of privacy leakage still even if simply generate and publish the synthetic data by various attacks (linkage attack, inference attack, etc.). In this paper, we propose a synthetic data generation algorithm in which privacy preserved by applying differential privacy the latest privacy protection technique to GAN, which is drawing attention as a synthetic data generative model in order to prevent the leakage of such sensitive information. The generative model used CGAN for efficient learning of labeled data, and applied Rényi differential privacy, which is relaxation of differential privacy, considering the utility aspects of the data. And validation of the utility of the generated data is conducted and compared through various classifiers.

Keywords: Synthetic Data, CGAN, Differential Privacy, Rényi Differential Privacy, Data Privacy

Received(07. 23. 2020), Accepted(08. 19. 2020)
* 이 논문은 2019년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(2019R1A2C1003146)과 2020년 공주대학교 학술연구지원사업의 연구지원에 의하여 연구되었음
* 본 논문은 2020년도 하계학술대회에 발표한 우수논문을 개선 및 확장한 것임
† 주저자, kjuny0924@smail.kongju.ac.kr
‡ 교신저자, dwhong@kongju.ac.kr(Corresponding author)

I. 서 론

많은 응용프로그램들이 데이터 수집과 분석을 통해 다양한 서비스를 제공하기 위해서는 데이터 공개가 필수적이다. 하지만 원본 데이터를 그대로 공개하는 것은 개인의 민감한 정보(예를 들어 앓고 있거나 앓았던 질병, 정치적 성향, 종교적 성향 등)를 유출할 위험이 있다. 따라서 원본 데이터가 아닌 재현 데이터를 생성하여 공개함으로써 프라이버시를 보존하는 방법에 대한 연구들이 수행되었다. 특히 의료데이터에 대해 Buczak 등[1]은 야토병과 관련된 전자의료기록(EMR)의 재현 데이터 생성 방법에 대한 연구를 수행하였고, McLachlan 등[2]은 다양한 의료 데이터 기반의 현실적인 전자건강기록(EHR)에 대한 재현 데이터 생성 방법인 CoMSER를 제안하였다. 또한 Choi 등[3]은 오토인코더와 GAN을 활용하여 EHR에 대한 재현 데이터 생성 알고리즘인 medGAN을 제안하였다.

그러나 단순히 재현 데이터를 생성하여 공개하는 것은 추가적인 공격에 의해 여전히 프라이버시 취약점이 존재한다. 실제로 익명의 넷플릭스 데이터를 통해 사용자들의 정치적 성향 등 민감한 정보가 유출될 수 있음을 보인 연구가 수행되었다[4]. 또한 Sweeny[5]는 공개된 의료기록 데이터와 주소를 연결 지어 환자의 추가적인 정보를 유추할 수 있음을 보였다.

이와 같이 단순히 재현 데이터를 생성하는 것은 민감한 정보의 유출 위험이 있으므로, 이를 방지하기 위해 프라이버시를 보존하면서 재현 데이터를 생성하는 기술에 대한 연구들이 수행되었다. 특히 수학적으로 엄격한 프라이버시를 보장하는 차분 프라이버시(Differential Privacy, DP)[6][7]를 만족하는 재현 데이터 생성에 대한 연구가 다양한 방면에서 수행되었다. Bowen 등[8]은 차분 프라이버시를 만족하는 다양한 재현 데이터 생성 기술들의 성능을 비교하고 분석하는 연구를 수행하였다. 또한 수집된 개인 데이터에 베이지안 충분통계량(Bayesian sufficient statistics)을 사용하여 차분 프라이버시를 만족하는 재현 데이터를 생성하는 방법인 modips[9]가 제안되었다. 그리고 Li 등[10], Zhang 등[11]은 각각 Copula 함수, 베이지안 네트워크를 활용하여 고차원 데이터에 대해 차분 프라이버시를 만족하는 재현 데이터 생성 기술인 DPCopula와 PrivaBayes를 제안하였다. 그밖에 최근 재현 데이터 생성방법으로 주목받는 딥러닝 모델에 차분 프라이버시를 적용하는 방법들이 제안되었다. 먼저, 오토인코더를 활용하여 차분 프라이버시를 만족하는 데이터 생성 알고리즘인 DP-SYN[12], 그리고 PATE 프레임워크를 수정하여 GAN모델에 적용한 PATE-GAN[13] 등이 제안되었다. 하지만 딥러닝 모델에 차분프라이버시를 적용하는 연구는 일반적인 차분 프라이버시에 대한 연구가 대부분이었으며 데이터의 유용성 향상을 위해 프라이버시를 완화시킨 개념에 대한 연구는 상대적으로 부족하다. 따라서 본 논문에서는 프라이버시가 완화된 개념을 GAN에 적용하여 차분 프라이버시를 만족하는 재현 데이터를 생성하는 방법을 제안한다.

최근 재현 데이터 생성에 있어서 주목받고 있는 GAN은 실제 데이터와 유사한 가짜 데이터를 생성하는 생성자, 실제 데이터와 가짜 데이터의 진위 여부를 판별하는 판별자로 구성되어 있으며, 생성자와 판별자를 번갈아 학습시키면서 실제 데이터와 더욱 유사한 가짜 데이터를 생성할 수 있는 모델이다. 만약 레이블이 있는 데이터에 대한 학습을 진행할 경우 CGAN을 사용하는 것이 효율적이다. CGAN의 판별자는 실제 데이터의 레이블을 포함하여 학습을 진행하고, 생성자는 주어진 레이블에 대한 가짜 데이터를 생성하게끔 학습을 진행함으로써 보다 효율적으로 학습할 수 있다.

최근 '2020년 MIT 10대 혁신 기술'[14]에 선정된 차분 프라이버시는 주목받고 있는 강력한 프라이버시 보호 기술이다. 차분 프라이버시는 데이터베이스에 한 개인의 존재 여부와 상관없이 데이터 분석 결과의 차이가 적다는 개념으로 데이터 수집 및 분석 그리고 기계학습 등 다양한 분야에서 사용되고 있다. 특히, 딥러닝 모델에 차분 프라이버시를 적용하기 위해 확률적 경사 하강법(Stochastic Gradient Decent, SGD)에서 계산되는 기울기에 정규분포를 따르는 잡음을 더해주는 DPSGD(Differential Privacy SGD)[15]가 제안되었다. 하지만 학습이 진행될수록 반복적으로 더해주는 잡음으로 인해 발생하는 프라이버시 비용은 급격하게 증가한다. 이를 해결하기 위해 Abadi 등[16]은 기존의 DPSGD에 기울기의 클리핑(Clipping) 기법을 적용하여 평균 기울기의 민감도의 한계를 제한하고 반복적인 학습을 통해 발생하는 프라이버시 비용을 추적하여 효율적으로 계산하는 방법인 Moments accountant를 제안하였다.

또한 유용성 향상을 위해 기존 차분 프라이버시를 완화시키는 방법에 대한 후속연구들이 많이 수행되었으며, 특히 Mironov[17]는 Rényi 발산을 사용하여 프라이버시를 완화시킨 Rényi 차분 프라이버시를 제안하였다.

본 논문에서는 공개된 데이터로부터 민감한 정보의 유출을 방지하기 위해 데이터 생성 모델인 CGAN에 Rényi 차분 프라이버시를 적용하여 프라이버시가 보존되는 재현 데이터를 생성하는 알고리즘인 Rényi differentially private CGAN을 제안한다. 또한 알고리즘의 학습에서 Rényi 차분 프라이버시를 만족하게 프라이버시 비용을 계산하여 발생하는 프라이버시 비용을 상대적으로 감소시킬 수 있으며, 생성된 데이터의 유용성을 향상시킬 수 있음을 보인다. 이를 검증하기 위해 두 개의 Kaggle 데이터와 한 개의 UCI데이터에 대한 재현데이터를 생성한 후에 8가지의 분류기를 통해 AUROC를 측정하여 비교분석한다.

본 논문의 2장에서는 사용된 개념들에 대해 설명하고, 생성 모델에 차분 프라이버시를 적용하는 방법과 제안한 알고리즘을 3장에서 설명한다. 그리고 4장에서는 생성한 데이터의 유용성을 다양한 분류기를 사용하여 분석하고, 마지막 5장에서 결론짓는다.

II. 배경 지식

이번 장에서는 본 논문에서 사용하는 생성 모델인 GAN과 프라이버시 보호 기법인 차분 프라이버시의 개념과 성질에 대해 설명한다.

2.1 GAN & CGAN

Goodfellow 등[18]이 제안한 GAN은 생성자와 판별자를 번갈아 학습시켜 실제 데이터와 유사한 데이터를 생성할 수 있는 비지도 학습 알고리즘이며, Fig. 1과 같은 구조를 갖는다. GAN의 생성자 G는 무작위 잡음 z를 입력으로 하여 실제 데이터와 유사한 가짜 데이터 $G(z)$를 생성하도록 학습되고, 판별자 D는 실제 데이터 x와 생성된 데이터 $G(z)$를 입력으로 하여 실제 데이터와 생성자가 생성한 가짜 데이터를 잘 판별하도록 학습된다. GAN의 학습은 다음과 같은 목적함수의 최대극소화(minimax)문제를 해결하는 것과 동일하다.

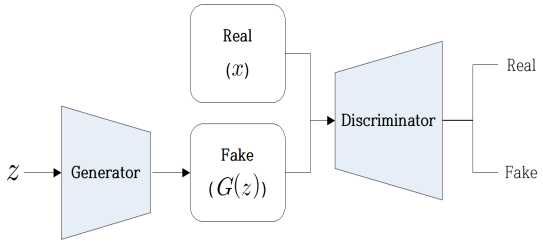

Fig. 1. GAN architecture

$$\min_G \max_D E_{x \sim p_{data}(x)}[\log D(x)] + E_{z \sim p_z(z)}[\log(1 - D(G(z)))]$$

하지만 GAN의 학습이 잘 되었더라도 원하는 레이블에 대한 데이터를 생성할 수 없다는 단점이 있다. 따라서 원하는 레이블에 대한 데이터를 생성하기 위해서는 학습 데이터에 레이블을 포함하여 학습시키는 CGAN[19]을 사용하는 것이 효율적이다. CGAN의 기본 구조는 GAN과 동일하며 생성자 G는 주어진 레이블 y에 대해 실제 데이터와 유사한 가짜 데이터 $G(z|y)$를 생성하도록 학습되고, 판별자 D는 실제 데이터 x와 레이블 y 그리고 생성된 데이터 $G(z|y)$를 입력으로 하여 실제 데이터와 가짜 데이터를 잘 판별하도록 학습된다. CGAN의 학습은 다음과 같은 목적 함수의 최대극소화문제를 해결하는 것과 동일하다.

$$\min_G \max_D E_{x \sim p_{data}(x)}[\log D(x|y)] + E_{z \sim p_z(z)}[\log(1 - D(G(z|y)|y))]$$

2.2 차분프라이버시

차분 프라이버시(Differential Privacy, DP)[6][7]는 수학적으로 엄격한 프라이버시를 보장하는 기술이다. 최근 '2020년 MIT 10대 혁신 기술'에 선정되었으며, 데이터 수집 및 분석, 기계학습 등 다양한 분야에서 사용되고 있다. 이러한 차분 프라이버시는 데이터베이스에 한 개인의 존재 여부와 상관없이 데이터 분석 결과의 차이가 적다는 개념으로 다음과 같이 정의한다.

정의 2.1 ((ϵ, δ)-차분 프라이버시((ϵ, δ)-DP)).
무작위 함수 K가 (ϵ, δ)-차분 프라이버시를 만족한다면, 최대 하나의 요소만 차이나는 인접한 데이터베이스

d, d'과 모든 출력 집합 S에 대해 다음 식을 만족한다.

$$\Pr[K(d)\in S] \le e^{\epsilon}\times\Pr[K(d')\in S]+\delta$$

이때 $\delta=0$이면 ϵ-차분 프라이버시라고 한다.

만약 ϵ이 충분히 작다면 인접한 데이터베이스에 대한 출력값의 차이가 적으므로 엄격한 프라이버시를 보장하고, ϵ이 커질수록 프라이버시가 감소한다. 차분 프라이버시를 만족하기 위한 일반적인 방법은 질의 함수의 출력값에 무작위 잡음을 더해주는 것이다. 이때 잡음의 크기는 해당 함수의 민감도에 의해 결정되며, 특정 함수의 민감도는 다음과 같이 정의한다.

정의 2.2 (l_2-민감도(l_2-sensitivity)).
주어진 함수 f와 모든 인접한 데이터 쌍 (d,d')에 대해, f의 민감도는 다음과 같다.

$$\Delta f = \max_{(d,d')}\|f(d)-f(d')\|_2$$

그리고 차분 프라이버시를 만족하는 데이터의 유용성 향상을 위해 엄격한 프라이버시를 완화시키는 많은 후속연구들이 수행되었다. 특히, Rényi 차분 프라이버시(Rényi Differential privacy, RDP)[17]는 Kullback-Leibler(KL) 발산을 사용한 기존 차분 프라이버시와는 달리 Rényi 발산을 사용하였으며, 다음과 같이 정의한다.

정의 2.3 (Kullback-Leibler(KL) 발산).
확률변수 P와 Q에 대한 KL 발산의 정의는 다음과 같다.

$$D(P\|Q) = E_{x\sim P}\log\frac{P(x)}{Q(x)}$$

정의 2.4 (Rényi 발산).
확률변수 P와 Q에 대한 Rényi 발산의 정의는 다음과 같다.

$$D_{\alpha}(P\|Q) = \frac{1}{\alpha-1}logE_{x\sim Q}\left(\frac{P(x)}{Q(x)}\right)^{\alpha}, (\alpha>1)$$

정의 2.5 ((α,ϵ)-Rényi 차분 프라이버시).
무작위 함수 K가 Rényi 차분 프라이버시를 만족한다면, 최대 하나의 요소만 차이나는 인접한 데이터베이스 d, d'에 대해 다음 식을 만족한다.

$$D_{\alpha}(K(d)\|K(d')) \le \epsilon$$

이때 $P(x)$는 x에 대한 밀도이며, $\alpha\to1$이면 KL 발산과 동치이다. 또한 $\alpha\to\infty$이면 KL 발산의 특수 경우인 최대 발산(Max divergence)이 된다. 최대 발산일 경우 Rényi 차분 프라이버시는 ϵ-차분 프라이버시를 만족한다. 이러한 (ϵ,δ)-차분 프라이버시와 (α,ϵ)-Rényi 차분 프라이버시의 관계는 다음과 같이 정리할 수 있다.

정리 2.1 ((ϵ,δ)-DP와 (α,ϵ)-RDP의 관계[17]).
만약 (α,ϵ)-Rényi 차분 프라이버시를 만족하는 함수를 K라고 한다면, 함수 K는 임의의 $0<\delta<1$에 대해 $(\epsilon+\frac{\log1/\delta}{\alpha-1},\delta)$-차분 프라이버시를 만족한다.

III. 재현 데이터 생성

차분 프라이버시를 만족하는 재현 데이터를 생성하는 방법에 대한 연구는 다방면에서 수행되었다. 하지만 일반적인 차분 프라이버시에 대한 연구가 대부분이었으며 데이터의 유용성 향상을 위해 프라이버시를 완화시킨 개념에 대한 연구는 상대적으로 부족하다. 따라서 이번 장에서는 프라이버시가 완화된 Rényi 차분 프라이버시를 데이터 생성 모델인 CGAN에 적용하여 재현 데이터를 생성하는 방법을 제안한다.

3.1 차분 프라이버시를 만족하는 GAN

GAN에 차분 프라이버시를 적용하기 위해 딥러닝 모델의 학습에서 사용되는 확률적 경사하강법(Stochastic Gradient Decent, SGD)에서 계산되는 기울기에 무작위 잡음을 추가하는 방법인 DPSGD(Differential Privacy SGD)가 제안되었다[15]. DPSGD는 딥러닝 모델에 차분 프라이버시를 적용하는 핵심 기술로 사용되고 있지만, 모델의 학습이 반복될수록 발생하는 프라이버시 비용은 급격하게 증가한다. 이를 해결하기 위해 Abadi 등[16]은 기울기 클리핑(Clipping) 기법을 통해 평균 기울기의 민감도의 한계를 제한하고 더욱 엄격하게 프라이버시

비용을 계산하는 방법인 Moments accountant를 제안하였다. Moments accountant를 사용한다면 기존에 반복적으로 발생하는 프라이버시 비용을 상대적으로 감소시킬 수 있다. 하지만 RDP accountant로 Rényi 차분 프라이버시를 만족하게 프라이버시 비용을 계산한다면, 발생하는 프라이버시 비용을 더욱 감소시킬 수 있다. 따라서 본 논문에서는 RDP accountant로 계산하여 Moments accountant보다 발생하는 프라이버시 비용을 감소시킬 수 있음을 보인다.

프라이버시 비용은 프라이버시 손실(loss)에 대한 추정값을 계산하는 것으로, 인접한 데이터베이스 d,d'과 메커니즘 K, 보조 입력값 aux, 출력값 $o\in R$가 주어졌을 때의 프라이버시 손실은 다음과 같이 정의한다.

정의 3.1 (프라이버시 손실(Privacy loss)[20]).

$$c(o;K,aux,d,d')\triangleq\log\left(\frac{\Pr[K(aux,d)=o]}{\Pr[K(aux,d')=o]}\right)$$

Moments accountant로 프라이버시 비용을 계산할 때 프라이버시 손실의 점근적인 한계는 표준편차 σ, 클리핑 변수 C, 데이터로부터의 무작위 추출 확률 $q<1/16\sigma$ 그리고 임의의 양수 $\alpha\le 1+\sigma^2\log(1/q\sigma)$에 대해 다음 식을 만족해야 한다[17].

$$\epsilon \le q^2\alpha(\alpha+1)/(1-q)C^2\sigma^2+O(q^3/C^3\sigma^3) \quad (1)$$

동일한 조건에서 RDP accountant의 프라이버시 손실의 점근적인 한계는 다음 식을 만족해야 한다[21].

$$\epsilon \le q^2\alpha/(1-q)C^2\sigma^2+O(q^3\alpha^3/C^3\sigma^3) \quad (2)$$

다시 말해, 식 (2)와 같이 계산한다면, 상수 부분을 제외하고 식 (1)보다 $(\alpha+1)$만큼의 인수를 절약할 수 있다. 따라서 RDP accountant로 프라이버시 손실의 한계를 계산하여 Moments accountant를 사용하는 것보다 발생하는 프라이버시 비용을 더욱 감소시킬 수 있다.

그리고 차분 프라이버시를 만족하는 함수가 임의의 함수와 결합하더라도 차분 프라이버시를 만족하게 되는 후처리 성질에 따라 GAN의 판별자 학습에만 Rényi 차분 프라이버시를 적용하더라도 생성된 데이터는 Rényi 차분 프라이버시를 만족한다.

Algorithm 1. Rényi differentially private CGAN

Input N : number of samples, L : group size, $q=L/N$: sampling probability, C : gradient norm bound, m : batch size, η : learning rate, σ : noise scale, (ϵ,δ) : total privacy budget

Output Rényi differentially private generator G

Initialize discriminator parameters d_0, generator parameters θ_0, n = 1

while $\hat{\epsilon}<\epsilon$ **do**
  for $t\in[T]$
    pick a random sample $L_t^x=\{x^{(j)}\}_{j=1}^{L}, L_t^y=\{y^{(j)}\}_{j=1}^{L}\sim P(X)$ from the real data with sampling probability $q=\frac{L}{N}$
    sample $\{z^{(j)}\}_{j=1}^{L}\sim P(Z)$ a batch of prior samples
    <compute the per-example gradient>
    $g_t(x^{(j)}|y^{(j)})\leftarrow\nabla_t D(x^{(j)}|y^{(j)})$ for $x^{(j)}\in L_t^x, y^{(j)}\in L_t^y$
    $g_t(z^{(j)}|y^{(j)})\leftarrow\nabla_t D(G(z^{(j)}|y^{(j)})|y^{(j)};\theta))$ for $j\in[L], y^{(j)}\in L_t^y$
    <clip gradient>
    $\bar{g}_t(x^{(j)}|y^{(j)})\leftarrow g_t(x^{(j)}|y^{(j)})/\max(1,\|g_t(x^{(j)}|y^{(j)})\|_2/C)$ for $x^{(j)}\in L_t^x, y^{(j)}\in L_t^y$
    $\bar{g}_t(z^{(j)}|y^{(j)})\leftarrow g_t(z^{(j)}|y^{(j)})/\max(1,\|g_t(z^{(j)}|y^{(j)})\|_2/C)$ for $j\in[L], y^{(j)}\in L_t^y$
    <add noise>
    $\tilde{g}_t\leftarrow\frac{1}{L}\sum_{j=1}^{L}\bar{g}_t(x^{(j)}|y^{(j)})+N(0,\sigma^2C^2I)-\frac{1}{L}\sum_{j=1}^{L}\bar{g}_t(z^{(j)}|y^{(j)})$
    $d\leftarrow d+\eta\times Adam(d,\tilde{g}_t)$
    <update RDP account>
    $\hat{\epsilon}\leftarrow\frac{q^2n\alpha}{(1-q)C^2\sigma^2}+O\left(\frac{nq^3\alpha^3}{C^3\sigma^3}\right)+\frac{\log 1/\delta}{\alpha-1}$
    $n\leftarrow n+1$
    pick a random sample $L_t^y=\{y^{(j)}\}_{j=1}^{m}\sim P(X)$
    sample $\{z^{(j)}\}_{j=1}^{m}\sim P(z)$ a batch of prior samples
    $g_\theta\leftarrow\nabla_\theta\frac{1}{m}\sum_{j=1}^{m}D(G(z^{(j)}|y^{(j)})|y^{(j)};\theta))$
    $\theta\leftarrow\theta+\eta\times Adam(\theta,g_\theta)$
  end for
end while
return G

Fig. 2. Rényi differentially private CGAN

3.2 Rényi differentially private CGAN

이번 절에서는 Rényi 차분 프라이버시를 만족하는 재현 데이터 생성 알고리즘인 Rényi differentially private CGAN을 제안한다. Fig. 2는 Rényi differentially private CGAN의 의사 코드(pseudo code)이다. 알고리즘은 전체 데이터의 수

N, 추출할 그룹의 크기 L, 그룹이 추출될 확률 $q=L/N$, 기울기 클리핑 변수 C, 배치 크기 m, 학습률 η, 표준편차 σ, 그리고 총 프라이버시 비용 (ϵ,δ)를 입력으로 하여 Rényi 차분 프라이버시를 만족하는 생성 모델 G를 출력한다. 모델의 학습에 앞서 판별자와 생성자의 가중치 변수 d_0, θ_0 그리고 학습 반복횟수 n을 초기화하고 판별자에 대한 학습을 먼저 진행한 후, 생성자에 대한 학습을 진행한다.

먼저 실제 데이터로부터 무작위 데이터 그룹 L_t^x와 레이블 그룹 L_t^y, 그리고 무작위 잡음 z를 L만큼 추출한다. 이때 그룹이 추출될 확률은 q이다. 그리고 각각의 기울기인 $g_t(x^{(j)}|y^{(j)})$와 $g_t(z^{(j)}|y^{(j)})$를 계산한다. line 6, 7에서 계산된 기울기를 변수 C로 클리핑한 값인 $\bar{g}_t(x^{(j)}|y^{(j)})$와 $\bar{g}_t(z^{(j)}|y^{(j)})$를 계산하고, line 9, 10에서 계산된 값의 평균을 구한 다음 가우시안 잡음 $N(0,\sigma^2C^2I)$을 더해준다. 계산된 기울기 값을 기반으로 Adam 최적화 알고리즘을 통해 판별자의 변수 d를 갱신하고 지금까지 소모된 프라이버시 비용을 RDP accountant로 계산한다. 다음으로 생성자의 학습을 진행하기 위해 실제 데이터의 레이블 배치 L_t^y와 무작위 잡음 z를 임의의 배치 크기 m만큼 추출한다. 그리고 추출된 데이터에 대한 기울기를 계산하고 Adam최적화 알고리즘을 통해 생성자 변수 θ를 갱신한다. 모든 학습은 사용된 프라이버시 비용이 주어진 총 프라이버시 비용을 초과할 때까지 반복한다.

Rényi differentially private CGAN은 판별자의 학습 과정에 Rényi 차분 프라이버시를 적용하였으며 차분 프라이버시의 후처리 성질에 의해 출력된 생성 모델 G로 생성한 데이터는 Rényi 차분 프라이버시를 만족한다.

3.3 프라이버시 증명

이번 절에서는 Rényi differentially private CGAN이 (α,ϵ)-Rényi 차분 프라이버시와 (ϵ,δ)- 차분 프라이버시를 만족함을 증명한다.

정리 3.1 표준편차 σ, 클리핑 변수 C에 대해 만약 민감도가 1이라면 가우시안 메커니즘은 $(\alpha,\alpha/2C^2\sigma^2)$-Rényi 차분 프라이버시를 만족한다.

증명 $D_\alpha(N(0,C^2\sigma^2)\|N(1,C^2\sigma^2))$

$$= \frac{1}{\alpha-1}log\int_{-\infty}^{\infty}\frac{1}{C\sigma\sqrt{2\pi}}exp(-\alpha x^2/(2C^2\sigma^2)) \times \exp(-(1-\alpha)(x-1)^2/(2C^2\sigma^2))dx$$

$$= \frac{1}{\alpha-1}log\frac{1}{C\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp[(-x^2+2(1-\alpha)x -(1-\alpha))/(2C^2\sigma^2)]dx$$

$$= \frac{1}{\alpha-1}log\{\frac{C\sigma\sqrt{2\pi}}{C\sigma\sqrt{2\pi}}exp[(\alpha^2-\alpha)/(2C^2\sigma^2)]\}$$

$$= \alpha/2C^2\sigma^2$$ □

정리 3.2 표준편차 σ, 클리핑 변수 C, 그리고 전체 학습 반복횟수 n에 대해 Rényi differentially private CGAN은 $(\alpha, q^2n\alpha/((1-q)C^2\sigma^2)+O(nq^3\alpha^3/(C^3\sigma^3)))$-Rényi 차분 프라이버시를 만족한다.

증명 RDP accountant의 프라이버시 손실의 점근적인 한계는 식 (2)를 만족해야한다. 또한 정리 3.1에 의해 한 번의 학습을 진행한 Rényi differentially private CGAN은 $(\alpha, q^2\alpha/((1-q)C^2\sigma^2)+O(q^3\alpha^3/(C^3\sigma^3)))$- Rényi 차분 프라이버시를 만족한다. 그리고 n번의 학습을 진행한 Rényi differentially private CGAN은 Rényi 차분 프라이버시의 결합성[17]에 의해 $(\alpha,n\alpha/(2C^2\sigma^2))$-Rényi 차분 프라이버시를 만족하며, 누적 프라이버시 비용은 $\epsilon \le q^2n\alpha/(1-q)C^2\sigma^2+O(nq^3\alpha^3/C^3\sigma^3)$으로 제한된다. 따라서 Rényi differentially private CGAN은 $(\alpha, q^2n\alpha/((1-q)C^2\sigma^2)+O(nq^3\alpha^3/(C^3\sigma^3)))$-Rényi 차분 프라이버시를 만족한다. □

정리 3.3 Rényi differentially private CGAN은 (ϵ,δ)-차분 프라이버시를 만족한다. 여기에서 ϵ은 $q^2n\alpha/((1-q)C^2\sigma^2)+O(nq^3\alpha^3/(C^3\sigma^3))+\log(1/\delta)/(\alpha-1)$이다.

증명 정리 3.2에 의해 Rényi differentially private CGAN은 $(\alpha, q^2n\alpha/((1-q)C^2\sigma^2)+O(nq^3\alpha^3/(C^3\sigma^3)))$-Rényi 차분 프라이버시를 만족한다. 그리고 정리 2.1의 관계에 따라 Rényi differentially private CGAN은 $(q^2n\alpha/((1-q)C^2\sigma^2)+O(nq^3\alpha^3/(C^3\sigma^3))+\log(1/\delta)/(\alpha-1),\delta)$-차분 프라이버시를 만족한다. □

Table 1. AUROC comparison across 8 classifiers trained on the synthetic credit card dataset generated each generative models(with ϵ=8, $\delta=10^{-5}$)

Algorithm / Classifier	Real	CGAN	dp_cgan	rdp_cgan
Logistic Regression	0.9797	0.9388	0.8959	0.9319
Decision Tree	0.9064	0.9012	0.8449	0.8479
Bagging	0.9481	0.9481	0.9027	0.9138
Random Forest	0.9823	0.9526	0.9242	0.9355
Gradient Boosting	0.9529	0.9185	0.8801	0.9067
Adaboost	0.9884	0.9546	0.8919	0.9335
Bernoulli Naive Bayes	0.9595	0.9091	0.8651	0.8941
XGBoost	0.9911	0.9551	0.9189	0.9452
Average	0.9635	0.9348	0.8905	0.9136

IV. 실 험

4.1 데이터셋

실험에 사용된 데이터셋은 총 3개이며 각각 다음과 같다.

- Kaggle credit card[22] : Kaggle credit card 데이터셋은 2013년 9월 유럽인들을 대상으로 수집한 신용카드 데이터로 레이블은 이상거래 여부이다. 총 29개의 특성을 갖는 284,807개의 데이터이며 492(0.2%)개만 이상거래 이력이 있는 데이터이다.
- Kaggle cervical cancer[23] : Kaggle cervical cancer 데이터셋은 베네수엘라의 카라카스 대학 병원에서 수집한 자궁경부암 데이터로 레이블은 자궁경부암 양성검사 결과이다. 총 35개의 특성을 갖는 858개의 데이터이며 55(6.4%)개만 양성판정을 받은 데이터이다.
- UCI Adult[24] : UCI Adult 데이터셋은 1994년 인구조사 데이터베이스에서 추출한 미국 성인의 소득 데이터로 레이블은 연 소득이 50,000 달러를 초과하는지에 대한 여부이다. 총14개의 특성을 갖는 32,561개의 데이터이며 7841(24.08%) 개만 연 소득이 50,000달러를 초과하는 데이터이다.

4.2 실험 과정

3장에서 제안한 알고리즘의 성능을 평가하기 위해 8개의 분류기를 통해 AUROC를 측정한다. 실험에 사용한 분류기는 로지스틱 회귀(Logistic Regression), 의사결정트리(Decision Tree), 배깅(Bagging), 랜덤 포레스트(Random Forest), 그래디언트 부스팅(Gradient Boosting), 에이다부스트(Adaboost), 베르누이 나이브 베이즈(Bernoulli Naive Bayes), XG부스트(XGBoost)이다. 이때 분류기의 학습 데이터는 재현 데이터를, 실험 데이터는 실제 데이터를 사용하며 학습 데이터와 실험 데이터의 비율은 8:2로 설정한다. 분류기의 학습 데이터로 사용할 재현 데이터는 CGAN과 (ϵ,δ)-차분 프라이버시가 적용된 dp_cgan, 그리고 본 논문에서 제안한 Rényi differentially private CGAN(rdp_cgan)으로 생성한 3개의 재현 데이터를 사용하며, CGAN으로 생성한 데이터는 차분 프라이버시가 보장되지 않는(non-private) 재현 데이터이다. 그리고 실제 데이터와 재현 데이터에 대해 각 분류기별 AUROC를 측정하여 알고리즘의 성능을 평가한다. 각 측정값들은 10번 반복 측정한 결과의 평균을 나타낸다.

4.3 실험 결과

Table. 1은 각각의 알고리즘으로 생성한 credit card 재현 데이터에 대해 프라이버시 비용(ϵ)이 8이고, δ는 10^{-5}일 때 분류기별 AUROC를 측정한 것이다. 각 측정값을 통해 rdp_cgan으로 생성한 데이터가 dp_cgan으로 생성한 데이터보다 AUROC값이 크므로 유용성 측면에서 더 좋은 결과를 나타내는 것을 확인할 수 있다. 또한 rdp_cgan으로 생성한 데이터에 대한 AUROC의 평균은 0.9136으로 프라이버시가 보장되지 않는 재현 데이터(CGAN)에 대한 AUROC의 평균인 0.9348과 근소한 차이를 나

Table 2. AUROC comparison across 8 classifiers trained on the synthetic credit card dataset generated by dp_cgan and rdp_cgan with various ϵ(with $\delta=10^{-5}$)

Algorithm	dp_cgan					rdp_cgan				
Epsilon / Classifier	ϵ = 1	ϵ = 2	ϵ = 4	ϵ = 6	ϵ = 8	ϵ = 1	ϵ = 2	ϵ = 4	ϵ = 6	ϵ = 8
Logistic Regression	0.6355	0.7486	0.7953	0.8644	0.8959	0.6502	0.7892	0.8048	0.8954	0.9319
Decision Tree	0.5184	0.6606	0.7641	0.8058	0.8449	0.5684	0.6723	0.7805	0.8039	0.8479
Bagging	0.7371	0.8125	0.8657	0.8998	0.9027	0.7506	0.8554	0.8879	0.9309	0.9138
Random Forest	0.7389	0.7818	0.8645	0.9071	0.9242	0.7555	0.7966	0.8838	0.9123	0.9355
Gradient Boosting	0.5511	0.8242	0.8389	0.8273	0.8801	0.6415	0.7478	0.8506	0.8339	0.9067
Adaboost	0.5381	0.8013	0.8105	0.8786	0.8919	0.6845	0.7802	0.8214	0.8842	0.9335
Bernoulli Naive Bayes	0.5265	0.5751	0.7992	0.8591	0.8651	0.6341	0.7054	0.8061	0.8928	0.8941
XGBoost	0.4343	0.7999	0.8484	0.8883	0.9189	0.5599	0.7031	0.8561	0.9389	0.9452
Average	0.5849	0.7505	0.8233	0.8663	0.8905	0.6556	0.7563	0.8364	0.8866	0.9136

Table 3. AUROC comparison across 8 classifiers trained on the synthetic cervical cancer and adult dataset generated each generative models(with ϵ=8, $\delta=10^{-5}$)

Dataset	Cervical cancer				Adult			
Algorithm / Classifier	Real	CGAN	dp_cgan	rdp_cgan	Real	CGAN	dp_cgan	rdp_cgan
Logistic Regression	0.9875	0.9596	0.9182	0.9376	0.9099	0.8838	0.8450	0.8621
Decision Tree	0.9787	0.9684	0.8853	0.9378	0.9123	0.8816	0.8301	0.8675
Bagging	0.9956	0.9465	0.9105	0.9278	0.9111	0.9005	0.8050	0.8356
Random Forest	0.9783	0.9499	0.9174	0.9328	0.9268	0.9081	0.8645	0.8715
Gradient Boosting	0.9898	0.9781	0.9087	0.9312	0.9462	0.9241	0.8201	0.8581
Adaboost	0.9901	0.9889	0.9241	0.9518	0.9222	0.8985	0.7664	0.8148
Bernoulli Naive Bayes	0.9466	0.9213	0.8955	0.9216	0.8873	0.8802	0.8541	0.8538
XGBoost	0.9913	0.9574	0.9134	0.9271	0.9424	0.9117	0.8147	0.8551
Average	0.9822	0.9587	0.9092	0.9335	0.9198	0.8986	0.8250	0.8523

타내며, 이는 프라이버시를 보존하면서 높은 수준의 유용성을 갖는 데이터를 생성할 수 있음을 의미한다. Table. 2는 dp_cgan과 rdp_cgan으로 생성한 credit card 재현 데이터에 대해 ϵ에 따른 분류기별 AUROC를 측정한 것이다. 표를 통해 rdp_cgan으로 생성한 데이터가 dp_cgan을 생성한 데이터보다 유용성이 향상됨을 확인할 수 있다. 또한 대부분의 경우 ϵ이 증가함에 따라 측정값이 증가하는데, 이는 ϵ이 증가할수록 생성된 데이터의 유용성은 증가하지만 프라이버시는 감소하는 것을 나타낸다. 그리고 rdp_cgan으로 생성한 데이터에 대한 측정에서 ϵ이 6, 8일 때의 Bagging을 사용한 측정, ϵ이 4, 6일 때 Gradient Boosting을 사용한 측정과 같이 측정값이 감소하는 경우가 드물게 발생할 수 있다.

Table. 3은 각각의 알고리즘으로 생성한 cancer와 adult 재현 데이터에 대해 $\epsilon=8, \delta=10^{-5}$일 때 분류기별 AUROC를 측정한 것이다. 데이터별 AUROC값을 통해 rdp_cgan으로 생성한 데이터가 dp_cgan으로 생성한 데이터보다 유용성이 향상되는 것을 확인할 수 있다. 또한 cancer와 adult 데이터의 경우 rdp_cgan으로 생성한 데이터에 대한 AUROC의 평균은 각각 0.9335, 0.8523으로 실제 데이터에 대한 AUROC의 평균인 0.9822, 0.9198과 비교하여도 높은 수준의 유용성 갖는다. 물론, 프라이버시를 보장하지 않는 재현 데이터에 대한 AUROC의 평균인 0.9587, 0.8986보다 작게 측정되지만 프라이버시가 보존된다는 점에서 높은 수준의

Table 4. AUROC comparison across 8 classifiers trained on the synthetic cervical cancer dataset generated by dp_cgan and rdp_cgan with various ϵ(with $\delta=10^{-5}$)

Algorithm	dp_cgan					rdp_cgan				
Epsilon / Classifier	ϵ=1	ϵ=2	ϵ=4	ϵ=6	ϵ=8	ϵ=1	ϵ=2	ϵ=4	ϵ=6	ϵ=8
Logistic Regression	0.6451	0.7496	0.7978	0.8733	0.9182	0.6611	0.7604	0.8114	0.8827	0.9376
Decision Tree	0.6204	0.7315	0.7887	0.8517	0.8853	0.6331	0.7478	0.8156	0.8620	0.9378
Bagging	0.6805	0.7617	0.8179	0.8736	0.9105	0.6957	0.7861	0.8339	0.8874	0.9278
Random Forest	0.6711	0.7567	0.8271	0.9121	0.9174	0.6822	0.7676	0.8633	0.9211	0.9328
Gradient Boosting	0.5577	0.7035	0.8343	0.8777	0.9087	0.5903	0.6607	0.7623	0.8804	0.9312
Adaboost	0.5871	0.7331	0.7571	0.8076	0.9241	0.5684	0.7481	0.8291	0.8491	0.9518
Bernoulli Naive Bayes	0.5733	0.7275	0.7947	0.8112	0.8955	0.5959	0.8102	0.7692	0.8607	0.9216
XGBoost	0.5841	0.6761	0.8288	0.8733	0.9134	0.5612	0.7537	0.7571	0.8741	0.9271
Average	0.6149	0.7300	0.8058	0.8601	0.9092	0.6235	0.7543	0.8053	0.8772	0.9335

Table 5. AUROC comparison across 8 classifiers trained on the synthetic adult dataset generated by dp_cgan and rdp_cgan with various ϵ(with $\delta=10^{-5}$)

Algorithm	dp_cgan					rdp_cgan				
Epsilon / Classifier	ϵ=1	ϵ=2	ϵ=4	ϵ=6	ϵ=8	ϵ=1	ϵ=2	ϵ=4	ϵ=6	ϵ=8
Logistic Regression	0.4775	0.6022	0.7212	0.7864	0.8450	0.5399	0.6579	0.7294	0.8225	0.8621
Decision Tree	0.4671	0.6123	0.7094	0.7812	0.8301	0.5436	0.6998	0.7499	0.8399	0.8675
Bagging	0.4401	0.6054	0.7087	0.7621	0.8050	0.5652	0.6607	0.7686	0.8078	0.8356
Random Forest	0.4403	0.6579	0.7543	0.8087	0.8645	0.6349	0.7435	0.8007	0.8483	0.8715
Gradient Boosting	0.4610	0.5756	0.6988	0.8402	0.8201	0.5396	0.6554	0.7497	0.8506	0.8581
Adaboost	0.4637	0.5756	0.6955	0.7639	0.7664	0.6192	0.6679	0.7399	0.7628	0.8148
Bernoulli Naive Bayes	0.5396	0.5416	0.7051	07579	0.8541	0.5509	0.7531	0.8070	0.8416	0.8538
XGBoost	0.4606	0.5709	0.6497	0.7998	0.8147	0.5473	0.6734	0.7561	0.8594	0.8551
Average	0.4687	0.5927	0.7053	0.7875	0.8250	0.5676	0.6890	0.7627	0.8291	0.8523

유용성을 갖는다고 볼 수 있다. Table. 4와 5는 dp_cgan과 rdp_cgan으로 생성한 cancer와 adult 재현 데이터에 대해 ϵ에 따른 분류기별 AUROC를 측정한 것이다. 표를 통해 cancer 데이터와 adult 데이터에 대해서도 rdp_cgan을 사용하여 재현 데이터를 생성하는 것이 dp_cgan을 사용하는 것보다 유용성이 향상되는 것을 확인할 수 있다. 또한 대부분의 경우 ϵ이 증가할수록 ARUOC가 증가하는데, 이는 ϵ이 작다면 생성된 데이터의 프라이버시는 증가하지만 유용성은 감소하고, ϵ이 크다면 그만큼 프라이버시는 감소하지만 유용성이 향상되는 것을 의미한다.

Fig. 3은 세 개의 데이터에 대해 8개의 분류기로 측정한 AUROC값의 평균을 그래프로 나타낸 것으로 왼쪽부터 순서대로 credit card, cancer, adult 데이터에 대한 그래프이다. 분석을 통해 rdp_cgan으로 생성한 데이터에 대한 AUROC가 dp_cgan으로 생성한 데이터에 대한 AUROC보다 크므로 더욱 유용성이 향상되는 것을 알 수 있다. 그리고 ϵ이 감소할수록 프라이버시가 증가하지만 데이터의 유용성은 급격하게 감소하는 것을 확인할 수 있다.

V. 결 론

데이터를 공개할 때 원본 그대로의 데이터를 공개하는 것은 개인의 민감한 정보를 유출할 위험이 있다. 하지만 민감한 정보의 유출을 막기 위해 재현 데

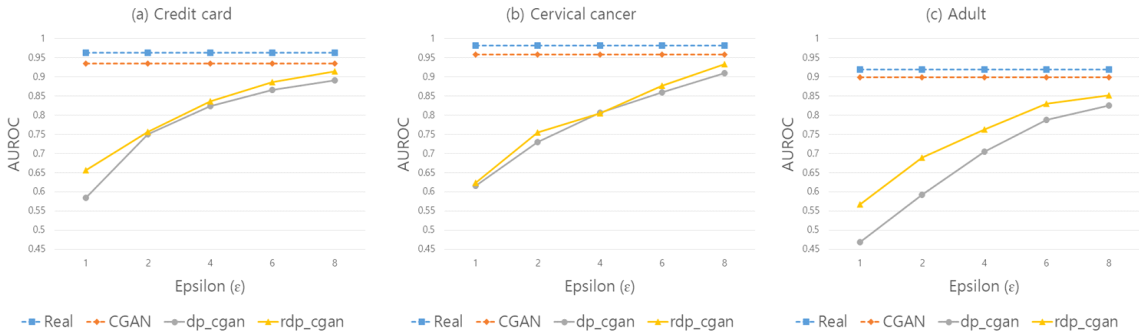

Fig. 3. Average of AUROC across 8 classifiers with various ϵ(with $\delta = 10^{-5}$)

이터를 생성 하더라도 추가적인 공격들로 인한 위험이 여전히 존재하는 것을 이전 연구들을 통해 확인하였다. 이를 방지하기 위해 차분 프라이버시를 만족하는 재현 데이터를 생성하는 방법에 대한 연구들이 많이 수행되었다. 하지만 일반적인 차분 프라이버시에 대한 연구가 대부분이었으며 유용성 향상을 위해 프라이버시를 완화시킨 차분 프라이버시에 대한 연구는 상대적으로 부족하다. 만약 프라이버시가 완화된 개념을 적용한다면, 일반적인 차분 프라이버시를 적용하는 것보다 효율적으로 활용 가능한 재현 데이터를 생성할 수 있다. 따라서 본 논문에서는 데이터 생성 모델인 CGAN에 프라이버시가 완화된 개념인 Rényi 차분 프라이버시를 적용하여 프라이버시가 보존되는 재현 데이터 생성 알고리즘인 Rényi differentially private CGAN을 제안하였다. 또한 알고리즘의 성능을 확인하기 위해 세 가지의 현실적인 데이터를 사용하여 재현 데이터를 생성하고, 이에 대한 유용성 분석을 위해 8가지의 분류기를 사용하였다. 실험을 통해 ϵ에 따른 프라이버시와 유용성의 상충 관계를 확인하였다. 또한, 대부분의 경우 동일한 ϵ에 대해 (α,ϵ)-Rényi 차분 프라이버시를 적용하는 것이 (ϵ,δ)-차분 프라이버시를 적용하는 것보다 각 분류기별 AUROC값이 크므로, 생성된 재현 데이터의 유용성이 향상되는 것을 확인하였다. 따라서 Rényi differentially private CGAN를 사용하여 재현 데이터를 생성한다면, 프라이버시가 보존되는 유용한 데이터를 공개할 수 있다.

References

[1] Buczak, Anna L., Steven B., and Linda M. "Data-driven approach for creating synthetic electronic medical records." BMC medical informatics and decision making 10(1), 59. Oct. 2010

[2] McLachlan, S., Kudakwashe D., and Thomas G. "Using the caremap with health incidents statistics for generating the realistic synthetic electronic healthcare record." IEEE International Conference on Healthcare Informatics (ICHI). 2016 IEEE, 2016. pp. 439-448. Oct. 2016

[3] Choi, E., Biswal, S., Malin, B., Duke, J., Stewart, W. F. and Sun, J. "Generating multi-label discrete patient records using generative adversarial networks." arXiv preprint arXiv:1703.06490. Mar. 2017

[4] Narayanan, A, and Vitaly S. "Robust de-anonymization of larg sparse datasets." 2008 IEEE Symposium on Security and Privacy, 2008 IEEE, pp. 111-125, May. 2008

[5] Sweeney, L. "Matching known patients to health records in Washington State data." Available at SSRN 2289850, Jul. 2013

[6] Dwork, C., McSherry, F., Nissim, K. and Smith, A. "Calibrating noise to sensitivity in private data analysis." Journal of Privacy and Confidentiality, 7(3), pp. 17-51, May. 2016

[7] Dwork, C. "Differential privacy: A

survey of results." International conference on theory and applications of models of computation. Springer, pp. 1-19, Apr. 2008
[8] Bowen, C. M., and Liu, F. "Comparative study of differentially private data synthesis methods." arXiv preprint arXiv: 1602.01063. Feb. 2016
[9] Liu, F. "Model-based differentially private data synthesis." arXiv preprint arXiv:1606.08052, Jun. 2016
[10] Li, H., Xiong, L. and Jiang, X. "Differentially private synthesization of multi-dimensional data using copula functions." InAdvances in database technology: proceedings. International conference on extending database technology, vol. 2014. NIH Public Access, pp. 475, Nov. 2014
[11] Zhang, J., Cormode, G., Procopiuc, C. M., Srivastava, D. and Xiao, X. "Privbayes: Private data release via bayesian networks." ACM Transactions on Database Systems (TODS), 42(4), pp. 1-41, Oct. 2017
[12] Abay, N. C., Zhow, Y., Kantarcioglu, M., Thuraisingham, B. and Sweeney, L. "Privacy preserving synthetic data release using deep learning." In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer, Cham, pp. 510-526, Jan. 2018
[13] Jordon, J., Yoon, J. and van der Scharr, M. "PATE-GAN: Generating synthetic data with differential privacy guarantees." In International Conference on Learning Representations. Sep. 2018
[14] http://www.technologyreview.com/10-breakthrough-technologies/2020/
[15] Song, S., Chaudhuri, K. and Sarwate, A. D. "Stochastic gradient descent with differentially private updates." In 2013 IEEE Global Conference on Signal and Information Processing, IEEE, pp. 245-248, Dec. 2013
[16] Abadi, M., Chu, A., Goodfellow, I., Mamahan, H. B., Mironov, I., Talwar, K. and Zhang, L. "Deep learning with differential privacy." In Proceedings of the 2016 ACM SIGSAC Confernece on Computer and Communications Security, pp. 308-318, Oct. 2016
[17] Mironov, I. "Rényi differential privacy." In 2017 IEEE 30th Computer Security Foundations Symposium (CSF), IEEE, pp. 263-275, Aug. 2017
[18] Goodfellow, I., Pouget_Abadie, J., Mirza, N., Xu, B. WardeFarley, D., Ozair, S., Courville, A. and Bengio Y. "Generative adversarial nets." In Advances in neural information processing systems, pp. 2672-2680, Jun. 2014
[19] Mirza, M and Ssindero, S. "Conditional generative adversarial nets." arXiv preprint arXiv:1411.1784, Nov. 2014
[20] Dwork, C. and Roth, A. "The algorihmic foundations of differential privacy." Foundations and Trends in Theoretical Computer Science, 9(3-4), pp. 211-407, Aug. 2014
[21] Mironov, I., Talwar, K. and Zhang, L. "Rényi Differential Privacy of the Sampled Gauusian Mechnism." arXiv preprint arXiv:1908.10530, Aug. 2019
[22] Dal Pozzolo, A., Caelen, O., Johnson, R. A. and Bontempi, G. "Calibrating probability with undersampling for unbalanced calssification." In 2015 IEEE Symposium Series on Computational Intelligence, IEEE, pp. 159-166, Jan. 2015
[23] Fernandes, K., Cardoso, J. S. and Fernandes, J. "Transfer learning with partial observability applied to cervical cancer screening." In Iberian conference on pattern recognition and

image analysis, Springer, Cham, pp. 243-250, May. 2017
[24] Asuncion, A. and Newman, D. "UCI machine learning repository." http://archive.ics.uci.edu/ml, 2007

〈저자소개〉

강 준 영(Junyoung Kang) 학생회원
2019년 2월: 공주대학교 응용수학과 학사
2019년 3월~현재: 공주대학교 융합과학과 석사과정
〈관심분야〉 암호 구현, 프라이버시 보호기술

정 수 용 (Sooyong Jeong) 학생회원
2018년 2월: 공주대학교 응용수학과 학사
2020년 2월: 공주대학교 융합과학과 석사
2020년 3월~현재: 공주대학교 융합과학과 박사과정
〈관심분야〉 암호모듈 구현, 데이터 보안

홍 도 원 (Dowon Hong) 종신회원
1994년 2월: 고려대학교 수학과 학사
2000년 2월: 고려대학교 수학과 박사
2000년 4월~2012년 2월: 한국전자통신연구원 팀장, 책임연구원
2012년 3월~현재: 공주대학교 응용수학과 교수
〈관심분야〉 암호기술, 프라이버시 보호기술

서 창 호 (Changho Seo) 종신회원
1990년: 고려대학교 수학과 학사
1992년: 고려대학교 수학과 이학석사
1996년: 고려대학교 수학과 이학박사
1996년~1996년: 국방과학연구소 선임연구원
1996년~2000년: 한국전자통신연구원 선임연구원, 팀장
2000년~현재: 공주대학교 응용수학과 교수
〈관심분야〉 암호알고리즘, PKI, 무선인터넷 보안 등