Differential Privacy

1. Introduction

The differential privacy algorithm is very important in protecting the data privacy of individuals. The algorithm is used in data analysis, machine learning, and statistics.

- DP-SGD: Adding noises to the gradient while training the model
- US used differential privacy to protect the 2020 census data [1].
- Uber used differential privacy to protect the data of the riders and drivers, preventing the data analyst from raw data [1].

2. Motivation

Why we use the word "differential" in the term? "Differential" here means the slightly difference in the dataset.

Imagine such a scenario, we have a dating website where the developer could see the number of people who have particular interests. For example, you could see the number of people who like to read books, watch movies, or play sports. When it comes to some sensitive interests, such as people who have a particular disease, if the bad guys track whether a person signs up for the website or not, they could infer whether the person has the disease or not by observing the change in the number of people who have the disease. This is where the differential privacy algorithm comes in. We need to ensure that individual's behaviour is not revealed by the data.

3. Definition of Differential Privacy

3.1. Task

In the case mentioned before, our challenge is to enable the dating website to publish accurate aggregated statistics about user interests while ensuring that individual user data cannot be inferred from these statistics, even by attackers with substantial background knowledge. Generally speaking, we need to develop a mechanism that allows us to answer questions about a dataset without revealing information about individual contributors to the dataset [2].

3.2. Definition

3.2.1. Preliminary Concepts

To obtain such a mechanism, we need to define the following mathmatical concepts:

- Query: A query is a function that takes a dataset as input and returns a real number, denoted by

  $$f : \boldsymbol{x} \mapsto \mathbb{R}$$

  Example: The number of people who like to read books in the dataset, the mean of the ages of the people in the dataset.

- Noise:

  - Laplace Mechanism: $f(x) = \frac{1}{2b} \exp\left(-\frac{|x-\mu|}{b}\right)$,where b is the scale parameter and μ is the mean.
  - Gaussian Mechanism: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, where σ is the standard deviation and μ is the mean.
- Randomized Algorithm: A randomized algorithm is an algorithm that for specific inputs, the output is not deterministic while it may follow a probability distribution.

  Example: We could obtain a randomized algorithm A by adding noise to the query function: $A(\boldsymbol{T}) = f(\boldsymbol{T}) + x$, where $x \sim \mathcal{N}(0,1)$
- Adjacent Datasets: Two datasets are adjacent if they differ by only one element, mathematically, two datasets $\boldsymbol{X}$ and $\boldsymbol{X}'$ are adjacent if

  $$|\boldsymbol{X} \Delta \boldsymbol{X}'| = 1$$

  Note: $A \Delta B = (A \setminus B) \cup (B \setminus A)$

3.2.2. Definition of Differential Privacy

Now we could define the differential privacy as follows:

A randomized algorithm $A : \boldsymbol{D} \mapsto \mathbb{R}$ satisfies ε-indistinguishable if for two adjacent dataset $\boldsymbol{D}$ and $\boldsymbol{D}'$ and any output O we have

$$\Pr\{A(\boldsymbol{D}) = O\} \leq e^{\varepsilon} \Pr\{A(\boldsymbol{D}') = O\}$$

$$\left|\log\left(\frac{\Pr\{A(\boldsymbol{D}) = O\}}{\Pr\{A(\boldsymbol{D}') = O\}}\right)\right| \leq \varepsilon$$

where ε is called the privacy budget or leakage [2].

Intuition: The differential privacy algorithm ensures that the probability of the output of the algorithm is really similar for two adjacent datasets, which means the attacker could not tell anything about whether a person is in the dataset or not. The graph below shows the intuition of the differential privacy algorithm.

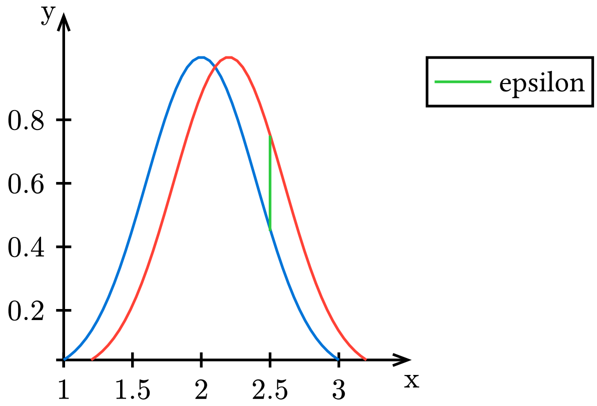

Figure 1: The PDF of two randomized algorithm

The green line here plot out the ε value, which means the maximum difference between the two PDF. The larger ε is, the more privacy we could bear to leak.

4. Design a Differential Privacy Algorithm

Here we introduce the laplace mechanism to design a differential privacy algorithm.

1. Define the query function f. f is used to analyze the dataset and return the result.
2. Compute the sensitivity Δf of the query function. The sensitivity of the query function is the maximum change in the output between any two adjacent datasets.
3. Calculate the scale parameter b of the laplace mechanism: $b = \frac{\Delta f}{\varepsilon}$
4. Add noise to the query function: $A(D) = f(D) + \text{Laplace}(0, b)$

Intution about scale parameter: Firstly, as the scale parameter b increases, the laplace distribution is more likely to output large noise. Secondly, if the sensitivity is large, more noises are needed to protect from the changes of slightly difference in the dataset. On the other hand, if the privacy budget is small, which means we want a safer privacy level, we need to add more noise to the query function.

5. Significance & Drawbacks

5.1. Advantages

- The ability to defend against new attacks: In contrast to traditional privacy algorithms, DP assume the attackers has the maximum background knowledge, which makes it more robust to defend against new attacks [3].
- Balance between privacy and utility: DP provides a trade-off between privacy and utility. The privacy budget could be adjusted to meet the requirement of privacy and utility [3].
- Theoretical guarantee: DP provides a theoretical guarantee that the privacy of the dataset is protected [3].

5.2. Drawbacks

- Although the design of the privacy budget is subtle, in real life, it is hard to determine the perfect privacy budget [1].
- Reconstruction the noise model: If the distribution of the noise model is determined, through numerous queries, the attacker could reconstruct the noise model and infer the individual's data. So in practice, the privacy budget could be set to decrease each time and as it reaches zero, no queries are allowed [1].

6. Python Demo

Bibliography

[1] "Differential Privacy and Applications." [Online]. Available: https://digitalprivacy.ieee.org/publications/topics/differential-privacy-and-applications

[2] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*, 2006, pp. 265–284.

[3] 熊平, 朱天清, 王晓峰, and others, "差分隐私保护及其应用," 计算机学报, vol. 37, no. 1, pp. 101–122, 2014.