Comp 585 Scribe 4/11

Joshua Yaffee

Spring 2024

Differential Privacy

In many contexts, data may be highly sensitive, yet the data needs to be used for some sort of public report. For example, an individual's hospital data should be kept confidential; meanwhile, hospitals or insurance companies may be inclined to publish deindividualized data. However, the deindividualized data may unexpectedly be less private than expected. In fact, in 1997, Massachusetts Governor William Weld's medical data was identified in such a fashion. Additionally, we have seen in this class how the Statistical Query Model can reconstruct data with enough queries. Due to this, it is crucial that algorithms are mathematically private. But how can usefully we define privacy? It's crucial that a malicious agent cannot guess with high accuracy the data of a single user even if given all other users' data. This motivates the following definition:

Definition 1. ***Neighboring Data Sets****: We say x and x' are neighboring data sets if and only if the hamming distance between x and x' is one. i.e. $||x - x'||_0 = 1$.*

This would correspond to a single user's data being changed. And so, we are ready to define privacy:

Definition 2. ***Differential Privacy****: We say a mechanism $M : \S \Rightarrow \mathcal{Y}$ is ϵ differentiably private if and only if $\forall S \subset \mathcal{Y}$ and for every neighboring pair (x, x') taken from $\S$,*

$$Pr[M(x) \in S] \leq e^{\epsilon} Pr[M(x') \in S]$$

By the symmetry of neighboring data sets, we can bound this expression on both sides:

$$e^{-\epsilon} Pr[M(x') \in S] \leq Pr[M(x) \in S] \leq e^{\epsilon} Pr[M(x') \in S]$$

Additionally, we define ϵ, δ Differential Privacy:

Definition 3. ***Differential Privacy****: We say a mechanism $M : \S \Rightarrow \mathcal{Y}$ is ϵ, delta differentiably private if and only if $\forall S \subset \mathcal{Y}$ and for every neighboring pair (x, x') taken from $\S$,*

$$Pr[M(x) \in S] \leq e^{\epsilon} Pr[M(x') \in S] + \delta$$

However, we use the first definition of differential privacy for the remainder of the lecture.

Let $f(x) : \mathcal{X} \Rightarrow \mathbb{R}$ be some function that we want to make private. Perhaps, $f(x_i)$ represents user i's income. Then, we can define sensitivity:

Definition 4. ***Sensitivity****: Let $Z = \{x, x' \in \mathcal{X} : X \text{ and } X' \text{are neighboring}\}$ be the set of all The sensitivity of a function f, $\Delta(f) = \max_{x,x' \in Z} |f(x) - f(x')|$*

If it is obvious which function we are concerned with, we may abbreviate $\Delta(f)$ as simply, Δ. Sensitivity is analogous to Lipschitz Continuity in a sense.

The first method we discussed is using the Laplace distribution to add appropriate noise to the data. the mechanism is given by $M(x) = f(x) + Lap(\frac{\Delta}{\epsilon})$ where the pdf of $Lap(t)$ is given by $\frac{e^{-|t|/b}}{2b}$. Then, $\forall S \subset \mathbb{R}$

$$\frac{Pr[M(x) \in S]}{Pr[M(x') \in S]} = \frac{\int_{y \in S} Pr[M(x) = y] dy}{\int_{y \in S} Pr[M(x') = y] dy} = \frac{\int_{y \in S} e^{\frac{|f(x)-y}{\Delta}} \frac{2\epsilon}{\Delta} dy}{\int_{y \in S} e^{\frac{|f(x')-y}{\Delta}} \frac{2\epsilon}{\Delta} dy} \leq \frac{\int_{y \in S} e^{\epsilon} e^{\frac{|f(x')-y}{\Delta}} dy}{\int_{y \in S} e^{\frac{|f(x')-y}{\Delta}} dy} \leq e^{\epsilon}$$

as desired.

Next, we explore the Exponential Mechanism. Suppose $S = \{s_1, ..., s_n\}$ and $u : S \times \mathcal{X} \Rightarrow \mathbb{R}$ is a utility function of s_i given the data, x. The objective is to maximize utility while still operating privately. The given solution is to pick s_i with likelihood

$$e^{\frac{u(s_i, x)\epsilon}{2\Delta}}$$

where $\Delta(u) = \max_{s_i \in S} \max_{x,x' \in Z} |u(s_i, x) - u(s_i, x')|$.
Then, given an outcome s_i

$$\frac{\Pr[M(x) = s_i]}{\Pr[M(x') = s_i]} = \frac{e^{\frac{u(s_i,x)\epsilon}{2\Delta}}}{\sum_{s_j \in S} e^{\frac{u(s_j,x)\epsilon}{2\Delta}}} \cdot \frac{\sum_{s_j \in S} e^{\frac{u(s_j,x')\epsilon}{2\Delta}}}{e^{\frac{u(s_i,x')\epsilon}{2\Delta}}} \leq e^{\frac{|u(s_i,x)-u(s_i,x')|\epsilon}{2\Delta}}$$

So for all s_j,

$$e^{-\epsilon/2} \leq \frac{e^{\frac{u(s_i,x)\epsilon}{2\Delta}}}{e^{\frac{u(s_j,x')\epsilon}{2\Delta}}} \leq e^{\epsilon/2}$$

Finally, we discussed randomized response mechanisms. In this setting, users give their data through a private protocol to a public server where the algorithm takes place. The user does not want to simply give their data as is, as it is entering a public stage. For simplicity, suppose a user's data is binary. That is, $x_i \in \{0, 1\}$. Then, an appropriate protocol would y_i like so:

$$y_i = \begin{cases} x_i & \text{with probability } \frac{e^\epsilon}{1+e^\epsilon}, \\ 1 - x_i & \text{with probability } \frac{1}{1+e^\epsilon}. \end{cases}$$

This ratio of probabilities will ensure privacy. It remains to show how well y_i estimates of x_i. It is rather straightforward:

$$\mathbb{E}[y_i] = \frac{e^\epsilon}{1+e^\epsilon}\mathbb{E}_S[x_i] + \frac{1}{1+e^\epsilon}(1 - \mathbb{E}_S[x_i]) = \frac{e^\epsilon - 1}{1+e^\epsilon}\mathbb{E}_S[x_i] + \frac{1}{1+e^\epsilon}$$

$$\implies (\mathbb{E}[y_i] - \frac{1}{1+e^\epsilon}) \cdot \frac{1+e^\epsilon}{1-e^\epsilon} = \mathbb{E}_S[x_i]$$