Advanced Topics in Machine Learning 2024

Amartya Sanyal

Home Assignment 5

Deadline: 23:59, Thursday, 24th October 2024

The assignments must be submitted individually – each student must write and submit a personal solution but we do not prevent you from discussing ***high-level*** *ideas in small groups. If you use any LLM tool such as ChatGPT, please specify the* ***purpose*** *and* ***manner*** *in which you have utilized it.*

We are interested in how *you solved the problems, and not just in the final answers. Please explain your solutions and ideas as clearly as you can.*

Late Penalty and multiple Submissions *Late submissions will incur a penalty of 10% of the total marks for every hour of delay (rounded up) with a maximum allowed delay of 5 hours after which the submission server will close. If you submit multiple submissions, only the last submission will be considered relevant both for grading answers as well as late penalty.*

Submission format: *Please upload your answers in a single* `.pdf` *file. If you have created code please include at least the key parts in the PDF as well as a link to the full code (on Github, Colab, or similar).*

We have studied the definition for Differential Privacy (DP)-Probably Approximately Correct (PAC) learning. We will now dive deep into it and design a generic DP-PAC learner. But first, lets define it again.

Definition 1. *For any $\alpha, \beta, \epsilon > 0$, an algorithm $\mathcal{A}_{DP}^{\alpha,\beta,\epsilon} : (\mathcal{X} \times \mathcal{Y})^n \to \mathcal{H}$[1] is said to be an $(\alpha, \beta, \epsilon)$-DP-PAC learner for hypothesis class $\mathcal{H}$ with private sample complexity n, if for all $h^* \in \mathcal{H}$ and for all distributions $\mathcal{D}$ over $\mathcal{X}$, the algorithm satisfies the following two properties.*

- ***Privacy*** *For any dataset S of size n sampled from $\mathcal{D}$ and labelled by h^*, $\mathcal{A}_{DP}^{\alpha,\beta,\epsilon}$ is $\epsilon - DP$ with respect to S.*
- ***Utility*** *Define $\mathcal{R}\left(\hat{h}; \mathcal{D}\right)$ to be the 0-1 error of $\hat{h}$ on distribution $\mathcal{D}$ labelled by h^*. Then,*

$$\Pr_S\left[\mathbb{E}_{\hat{h}\sim\mathcal{A}_{DP}(S)}\left[\mathcal{R}\left(\hat{h}; \mathcal{D}\right)\right] \leq \alpha\right] \geq 1 - \beta$$

where the outer probability is over all datasets S of size n sampled from $\mathcal{D}$ and labelled with h^*.

Exercise 1 (50 points)**.** *For any $d > 0$, let $\mathcal{H}_d$ be any finite hypothesis class such that $|\mathcal{H}_d| \leq \exp(\text{poly}(d))$. For any such $\mathcal{H}_d$, design a generic $(\alpha, \beta, \epsilon)$- DP-PAC learner with sample size polynomial in $d, \frac{1}{\alpha}, \frac{1}{\epsilon}, \log\left(\frac{1}{\beta}\right)$. Your solution does not need to be computationally efficient. You must write down the algorithm as well as the proof for its privacy and utility guarantee.*

1. *10 points are reserved for a clear pseudocode of the algorithm.*
2. *10 points are reserved for the proof of privacy.*
3. *30 points are reserved for the proof utility. This proof contains two main steps which are described in the hint below and each carries equal weight of 15 points each.*

[1] For simplicity, we will shorthand $\mathcal{A}_{\text{DP}}^{\alpha,\beta,\epsilon}$ with $\mathcal{A}_{\text{DP}}$.

Hint: *First step involves showing that no matter what $\hat{h}$ is output, its error on the train set will be close to the error on the distribution. This argument involves a combination of Chernoff bound and union bound similar to what we did in class. Second step involves showing that we can privately find a $\hat{h}$ with good training error. Each of these two steps carry 20 points each.*

The above result dealt with the case of finite hypothesis class $\mathcal{H}$ and any domain $\mathcal{X}$. Now we will look at **any** hypothesis class with finite VC dimension and finite input domain.

Exercise 2 (35 points)**.** *For any $d, p > 0$, let $\mathcal{X}_p$ be a finite input domain such that $|\mathcal{X}_p| \leq \exp(p)$ and $\mathcal{H}_d$ be any hypothesis class on $|\mathcal{X}_p|$ with VC dimension d. For any such $\mathcal{H}_d, \mathcal{X}_p$, design a generic $(\alpha, \beta, \epsilon)$- DP-PAC learner with sample size polynomial in $d, p, \frac{1}{\alpha}, \frac{1}{\epsilon}, \log\left(\frac{1}{\beta}\right)$. Your solution does not need to be computationally efficient. You must write down the algorithm as well as the proof for its privacy and utility guarantee.*

Hint: *First step involves finding a finite set of hypothesis which is sufficient to output from. You will have to use the Sauer-Shelah lemma in this step. Second step is applying the previous theorem in Exercise 1.*

Exercise 3 (15 points)**.** *Based on the above results and non-private learning theory, comment on the differences between private and non-private PAC learning from both statistical (amount of data required for small error) as well as computational (amount of time taken) perspective.*

Do not write more than 100 words.

Good luck!

Amartya