Query Complexity of Bayesian Private Learning

Kuang Xu Stanford Graduate School of Business kuangxu@stanford.edu

STANFORD GRADUATE SCHOOL OF BUSINESS

Summary

We study the query complexity of Bayesian Private Learning [1]. A learner wishes to locate a random target within an interval by submitting queries, in the presence of an adversary who observes all of her queries but not the responses. How many queries are necessary and sufficient in order for the learner to accurately estimate the target, while simultaneously concealing the target from the adversary?

The main result is a query complexity lower bound that is tight up to the first order. We show that if the learner wants to estimate the target within an error of ϵ, while ensuring that no adversary estimator can achieve a constant additive error with probability greater than $1/L$, then the query complexity is on the order of $L\log(1/\epsilon)$ as $\epsilon \to 0$. Our result demonstrates that increased privacy, as captured by L, comes at the expense of a *multiplicative* increase in query complexity. The proof builds on Fano's inequality and properties of certain proportional-sampling estimators.

Bayesian Private Learning - Learner

• A model for studying the **privacy vs. query complexity trade-off in active learning**.

• A learner wants to find a target, X^*, distributed uniformly at random in $[0,1)$.

• Learner submits N queries sequentially, Q_k, $k=1,\ldots,N$, and receives a response for each query, indicating target's relative position:

$$R_k = \mathbb{I}\{X^* \geq Q_k\}$$

• Learner constructs estimator $\widehat{X}$ for the target using the responses. A learner strategy is ϵ**-accurate** if

$$|\widehat{X} - X^*| \leq \epsilon/2, \quad \text{with probability 1.}$$

Bayesian Private Learning - Adversary

• An adversary observes all of learner's queries, $\{Q_k\}_{k=1,\ldots,N}$, but not responses.

• Adversary constructs an estimator $\widehat{X}^a$ using the observed queries. Adversary estimator is (δ, L)**-accurate** if

$$\mathbb{P}(|\widehat{X}^a - X^*| \leq \delta/2) > \frac{1}{L},$$

• **Privacy**: A learner query strategy is (δ, L)**-private**, if no adversary estimator is (δ, L)-accurate. Here, L corresponds to the level of privacy.

Metric: Query Complexity

Complexity of Bayesian Private Learning:

The **query complexity**, $N^*(\epsilon, \delta, L)$, is defined as the minimum number of queries for there to exist an ϵ-accurate learner strategy that is also (δ, L)-private.

Main Result: Query Complexity Lower Bound

Theorem (Query Complexity of Bayesian Private Learning) Fix ϵ and δ in $(0,1)$ and $L \in \mathbb{N}$, such that $\epsilon < \delta/4$ and $\delta < 1/L$.

1. Upper bound ([1])

$$N(\epsilon, \delta, L) \leq L\log(1/\epsilon) - L(\log L - 1).$$

2. Lower bound (this work)

$$N(\epsilon, \delta, L) \geq L\log(1/\epsilon) - L\log(2/\delta) - 3L\log\log(\delta/\epsilon).$$

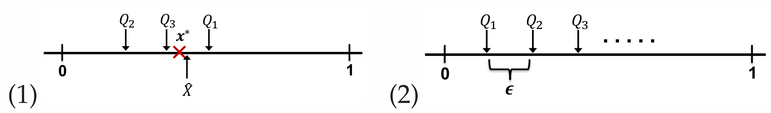
Applying the theorem in the regime where δ and L stay fixed, while the learner's error tolerance, $\epsilon \to 0$, we obtain the following corollary in which the upper and lower bounds **coincide**.

Corollary (Multiplicative Price of Privacy): Fix $\delta \in (0,1)$ and $L \in \mathbb{N}$, such that $\delta < 1/L$.

$$N(\epsilon, \delta, L) \sim L\log(1/\epsilon), \quad \text{as } \epsilon \to 0.$$

Our results demonstrate that there is a hefty price to pay in exchange for privacy, as the query complexity depends *multiplicatively* on the level of privacy, L.

Two (Extreme) Examples



(1) **Bisection search**: recursively halving the size of the interval known to contain the target.

• Complexity: optimal, with $\log(1/\epsilon)$ queries.

• Privacy: **never** private. Adversary can infer target by simply looking at the last query.

(2) ϵ**-Uniform**: uniformly placing $1/\epsilon$ queries across the domain.

• Complexity: highly inefficient, with $(1/\epsilon)$ queries.

• Privacy: **always** private. Queries do not depend on target value.

Key Proof Ideas

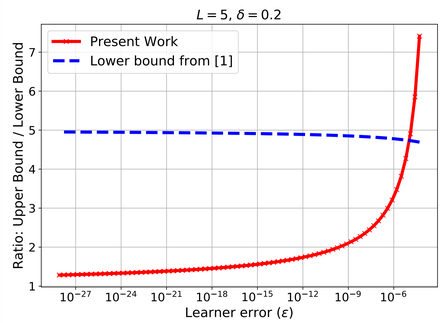
• Main challenge is to characterize the posterior distribution of the target for an arbitrary learner strategy.

• Key idea: analyze instead a (sub-optimal) adversary inference algorithm that uses *proportional sampling* (PS).

• Under PS, the adversary essentially chooses uniformly at random one of the queries and uses it as its target estimator. We show PS is sufficient to force any accurate learner to use a large number of queries.

• Proof hinges on a local complexity lemma derived using Fanno's inequality, which states that the number of queries in the **vicinity of the target** much be large under any accurate learner strategy, hence exposing the learner to attacks under PS.

Numerical Example



• Comparison between the tightness of the lower bound in the present work and that in [1]. Note that the ratio between the upper and lower bound in the present work converges to 1 as $\epsilon \to 0$.

Conclusions

• Privacy requirement can lead to a substantial, multiplicative overhead in the query complexity of active learning.

• Future work:

◇ Generalizing results to high-dimensional active learning, as well as learning with observational noise.

◇ Exploring implications of privacy in learning problems where actions (queries) and information (responses) are highly correlated.

References

[1] J.N. Tsitsiklis, K. Xu and Z. Xu, Private sequential learning, *Conference on Learning Theory (COLT)*, Stockholm, July 2018.