Privately Learning Thresholds: Closing the Exponential Gap

Abstract:

We study the sample complexity of learning threshold functions under the constraint of differential privacy. Unlike the non-private case, where the sample complexity is independent of the domain size, it turns our that for private learning the sample complexity must depend on the domain size $X$. Our current understanding of this task places its sample complexity somewhere between $\log^*|X|$ and $2^{\log^*|X|}$, where at least three different algorithms are known with sample complexity exponential in $\log^*|X|$. In this work we reduce this gap significantly, and show that the sample complexity is at most polynomial in $\log^*|X|$.

Joint work with Haim Kaplan, Katrina Ligett, Yishay Mansour, and Moni Naor.