מכון ויצמן למדע

THE WEIZMANN INSTITUTE OF SCIENCE

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Machine Learning and Statistics Seminar

Room 1 ,Ziskind Building
on Wednesday, Mar 27, 2019
at 11:15

Roi Livni
TAU

Passing Tests Without Memorizing

Abstract:

Generative Adversarial Networks (GANs) is a recent algorithmic framework that has won considerable attention. In a nutshell, GANs receive as input an IID sample and outputs synthetic data that should resemble data from the true underlying distribution. For example, consider an algorithm that receives as input some tunes from a specific music genre (e.g. jazz, rock, pop) and then outputs a new, original, tune from that genre.

From a theoretical perspective, the distinction between algorithms that genuinely generate original new examples vs. algorithms that perform naive manipulations (or even merely memorization) of the input sample is an elusive distinction. This makes the theoretical analysis of GANs algorithms challenging.

In this work we introduce two mathematical frameworks for the task of generating synthetic data. The first model we consider is inspired by GANs, and the learning algorithm has only an indirect access to the target distribution via a discriminator. The second model, called DP-Foolability, exploits the notion of differential privacy as a criterion for "non-memorization".

We characterize learnability in each of these models as well as discuss the interrelations. As an application we prove that privately PAC learnable classes are DP-foolable. As we will discuss, this can be seen as an analogue of the equivalence between uniform convergence and learnability in classical PAC learning.

Joint work with Olivier Bousquet and Shay Moran.
https://arxiv.org/pdf/1902.03468.pdf