ID de Contribution: **74** Type: **Invited talk**

Robust learning with local and global adversarial corruptions

We study learning in an adversarial setting, where an epsilon fraction of samples from a distribution P are globally corrupted (arbitrarily modified), and the remaining perturbations have an average magnitude bounded by rho (local corruptions). With access to n such corrupted samples, we aim to develop a computationally efficient approach that achieves the optimal minimax excess risk. Our approach combines a data-driven cleaning module with a distributionally robust optimization (DRO) framework. We demonstrate that if the data cleaning module is minimax optimal with respect to the Wasserstein loss, solving an optimal transport-based DRO problem ensures a minimax optimal decision. We further provide tractable reformulations for both modules. Specifically, we introduce an optimal filtering algorithm to clean corrupted data by identifying and removing outliers. For the DRO module, we reformulate the problem as a two-player zero-sum game, deriving finite convex formulations. We show that the minimax theorem applies to this game, and Nash equilibria exist. Finally, we present a principled approach for constructing adversarial examples.

Authors: NIETERT, Sloan; SHAFIEE, Soroosh; GOLDFELD, Ziv

Orateur: SHAFIEE, Soroosh

Classification de Session: Contextual stochastic programming

Classification de thématique: Contextual stochastic programming