ID de Contribution: **7** Type: **Non spécifié**

Datasets dictionary learning and applications to frugal AI

vendredi 6 juin 2025 15:30 (20 minutes)

Empirical risk minimization learning paradigm works under the assumption that training and test data are identically distributed. However, this hypothesis is seldom met in practice, due to several factors including changes in the underlying physical process generating the data, data acquisition conditions or sensors drifts. This problem is known as distributional shift between the reference and target data. A straightforward generalization of this problem is when training with multiple datasets exhibiting distributional shifts.

In this context, optimal transport provides useful tools for comparing and manipulating probability measures. In particular, a Wasserstein barycenter is a meaninful way of averaging datasets in the space of probability measures equipped with the optimal transport based distance. Leveraging this concept, we introduce a novel dictionary learning problem over empirical probability measures and present its applications to different learning setting where no labels are available on the target dataset.

Orateur: NGOLÈ, Fred (CEA Paris-Saclay)

Classification de Session: COLOURS School