Sequential approaches for learning datum-wise sparse representations

**Gabriel Dulac-Arnold · Ludovic Denoyer ·
Philippe Preux · Patrick Gallinari**

Abstract In supervised classification, data representation is usually considered at the dataset level: one looks for the "best" representation of data assuming it to be the same for all the data in the data space. We propose a different approach where the representations used for classification are tailored to each datum in the data space. One immediate goal is to obtain *sparse datum-wise representations*: our approach learns to build a representation specific to each datum that contains only a small subset of the features, thus allowing classification to be fast and efficient. This representation is obtained by way of a sequential decision process that sequentially chooses which features to acquire before classifying a particular point; this process is learned through algorithms based on Reinforcement Learning.

The proposed method performs well on an ensemble of medium-sized sparse classification problems. It offers an alternative to global sparsity approaches, and is a natural framework for sequential classification problems. The method extends easily to a whole family of sparsity-related problem which would otherwise require developing specific solutions. This is the case in particular for cost-sensitive and limited-budget classification, where feature acquisition is costly and is often performed sequentially. Finally, our approach can handle non-differentiable loss functions or combinatorial optimization encountered in more complex feature selection problems.

G. Dulac-Arnold (✉) · L. Denoyer · P. Gallinari
UPMC, LIP6, Université Pierre et Marie Curie, Case 169, 4 Place Jussieu, 75005 Paris, France
e-mail: gabriel.dulac-arnold@lip6.fr

L. Denoyer
e-mail: ludovic.denoyer@lip6.fr

P. Gallinari
e-mail: patrick.gallinari@lip6.fr

P. Preux
LIFL (UMR CNRS) & INRIA Lille Nord-Europe, Université de Lille, Villeneuve d'Ascq, France
e-mail: philippe.preux@inria.fr