Mach Learn
DOI 10.1007/s10994-013-5395-y

Metacluster-based Projective Clustering Ensembles

Francesco Gullo · Carlotta Domeniconi · Andrea Tagarelli

Received: 25 May 2012 / Accepted: 11 June 2013

Abstract The Projective Clustering Ensemble (PCE) problem is a recent clustering advance aimed at combining the two powerful tools of *clustering ensembles* and *projective clustering*. PCE has been formalized as either a two-objective or a single-objective optimization problem. Two-objective PCE has been recognized as more accurate than its single-objective counterpart, although it is unable to jointly handle the object-based and feature-based cluster representations.

In this paper, we push forward the current PCE research, aiming to overcome the limitations of all existing PCE formulations. We propose a novel single-objective PCE formulation so that (i) the object-based and feature-based cluster representations are jointly considered, and (ii) the resulting optimization strategy follows a metacluster-based methodology borrowed from traditional clustering ensembles. As a result, the proposed formulation features best suitability to the PCE problem, thus guaranteeing improved effectiveness. Experiments on benchmark datasets have shown how the proposed approach achieves better average accuracy than all existing PCE methods, as well as efficiency superior to the most accurate existing metacluster-based PCE method on larger datasets.

Keywords Clustering · Clustering ensembles · Projective clustering · Subspace clustering · Dimensionality reduction · Optimization

Editors: Emmanuel Müller, Ira Assent, Stephan Günnemann, Thomas Seidl, Jennifer Dy.

F. Gullo (✉)
Yahoo! Research, 08018 Barcelona, Spain
e-mail: gullo@yahoo-inc.com

C. Domeniconi
Dept. of Computer Science, George Mason University, 22030 Fairfax, VA, USA
e-mail: carlotta@cs.gmu.edu

A. Tagarelli
DIMES Dept., University of Calabria, 87036 Rende (CS), Italy
e-mail: tagarelli@deis.unical.it

Published online: 29 June 2013