Collective Outlier Detection and Enumeration with Conformalized Closed Testing

Chiara G. Magnani c.magnani9@campus.unimib.it
Department of Economics, Management and Statistics, University of Milano-Bicocca, Milano, Italy
Matteo Sesia sesia@marshall.usc.edu
Department of Data Sciences and Operations, University of Southern California, Los Angeles, California, USA
Aldo Solari aldo.solari@unive.it
Department of Economics, Ca' Foscari University of Venice, Venice, Italy; Department of Economics, Management and Statistics, University of Milano-Bicocca, Milano, Italy

Editor: Simone Vantini, Matteo Fontana, Aldo Solari, Henrik Boström and Lars Carlsson

Abstract

This paper develops a distribution-free method for collective outlier detection and enumeration, designed for situations in which the precise identification of individual outliers may be impractical due to the sparsity or weakness of their signals. This method builds upon the latest developments in conformal inference and blends them with more classical ideas from other areas, including multiple testing, rank tests, permutations, and non-parametric large-sample asymptotics. Key innovations include an extension of the Wilcoxon-Mann-Whitney test, which may be of some independent interest, and a principled algorithm for tuning the choices of machine learning classifier and two-sample testing procedure utilized by our method, yielding to an adaptive strategy. Assuming to have a control sample where all the observations are drawn independently from the same distribution (*inlier* distribution) and a test sample where possibly some observations are extracted from a different distribution (*outlier* distribution), our methodology implements the closed testing procedure providing simultaneous inference on the number of outliers in the test sample or in any subset of the test set. The inferential result produced by our method is a $(1-\alpha)$-confidence lower bounds for the number of true outliers after any selection of the data in the test set. Further, we motivate theoretically the choice of the extended Wilcoxon-Mann-Whitney tests as local test in the closed testing procedure, studying their optimality and deriving interesting findings under distribution-free alternatives. Delving into how local optimality transfers to the closed testing procedure is prompt for future research. The effectiveness of our method is highlighted through extensive empirical demonstrations, including an analysis of the LHCO high-energy particle collision data set.

Keywords: Closed testing; Conformal inference; Novelty detection; Rank tests.