



Clustering Data Submission Challenge Call: Contributing Data Sets to Cluster Benchmark Data Repository

The Cluster Benchmarking Task Force of the International Federation of Classification Societies (IFCS) is calling for contributions to their recently launched Cluster Benchmark Data Repository at <http://ifcs.boku.ac.at/repository/>. This submission call is part of a competition connected with the IFCS 2017 conference in Tokyo.

The aim of the repository is to stimulate better practice in benchmarking (performance comparison of methods) for cluster analysis by providing a unique resource comprised of a wide variety of well-documented, high quality datasets and simulation routines for use in practical benchmarking. The repository will include both datasets with and datasets without given “true” clusterings. A unique feature of the repository is that each dataset will be supplied with comprehensive meta-data, including documentation on the specific nature of the clustering problem and on characteristics that useful clusters should fulfill (with scientific justification).

The Task Force is calling for individuals to contribute data sets to this repository as part of a challenge.

Rules for submission:

- Datasets for the challenge may be submitted at <http://ifcs.boku.ac.at/repository/>; please indicate in the corresponding field that your dataset is submitted for the challenge.
- Meta-data are entered via a questionnaire that can be found as part of the submission process on the repository website.
- Each participant in the challenge is allowed to submit several datasets.
- The deadline for dataset submissions is January 15, 2017, 5pm CET.

Competition guidelines:

- Contributions to the challenge will be evaluated by the IFCS Task Force on Benchmarking.
- Apart from technical correctness, the major evaluation criterion will be quality of the meta-data.
- One or two data sets will be selected as winners of the challenge.
- The contributor(s) of the winning data set(s) will be notified no later than January 31, 2017.
- The winning data set(s) will be the object of a second challenge, in which they will be subjected to various types of cluster analysis.
- The winning data set(s) will be introduced during an invited session at the forthcoming IFCS conference in Tokyo, Japan, August 8-10, along with results from the second, data-analytic challenge. The contributor(s) of the winning data set(s) will be welcome to take care of this introduction, yet conference attendance is not a requirement for participation in the challenge.
- The contributor(s) of the winning data set(s) will receive a plaque and a book award.
- Contact for questions: Iven.VanMechelen@kuleuven.be

All researchers interested in clustering and classification are most welcome to participate in this challenge!

Taskforce members: Anne-Laure Boulesteix, Rainer Dangl, Nema Dean, Isabelle Guyon, Christian Hennig, Friedrich Leisch, Douglas Steinley, and Iven Van Mechelen