

Stimulating Benchmarking in Classification Research: A Challenge Taken up by the International Federation of Classification Societies

Iven Van Mechelen*

University of Leuven, Belgium Iven.VanMechelen@ppw.kuleuven.be

In the classification domain, very often new methods, algorithms, and data-analytic devices (e.g., procedures to determine the number of clusters in a data set) are being proposed without adequate comparisons with current best practices or gold standards, and without sufficient evidence of good performance on benchmarking data sets. This practice is prohibitive, as it hampers the building of a cumulative body of knowledge. Most recently, the International Federation of Classification Societies has launched a new initiative to stimulate benchmarking in the domain of (unsupervised and supervised) classification methods. In this presentation I will present a brief overview of this initiative, which comprises: (a) a conceptual framework for benchmarking, (b) pointers to a broad range of benchmarking data sets, (c) recommendations for generating appropriate (synthetic and naturalistic) simulated data sets for performance evaluation, (d) recommendations for suitable experimental designs for such evaluations, and (e) an inventory of suitable measures of performance. Special emphasis will be put on a discussion of the concept of quality of a clustering and its evaluation.

Key Words: Cluster analysis, unsupervised and supervised learning, evaluation