EDITORIAL

Proceedings of Reisensburg 2011

**Harald Binder · Hans A. Kestler ·
Matthias Schmid**

The conference "Statistical Computing" is held annually at the Reisensburg castle (Germany) by the working group "Biostatistics" of the German Classification Society and by the working group "Statistical Computing" of the German Region of the International Biometric Society and of the German Society of Medical Informatics, Biometry and Epidemiology (GMDS). The conference covers recent topics in Biostatistics and Bioinformatics, with a special focus on applications regarding the computational aspects of these fields of research. In 2009, the conference organizers started publishing selected articles of the conference as a special issue of Computational Statistics: Volume 26, Issue 2 (2011) for the 2009 conference and Volume 28, Issue 1 (2013) for the 2010 conference. Being either research or tutorial papers on recent software developments, all articles featured in a Reisensburg special issue undergo the regular reviewing process of Computational Statistics. This year's special issue features the highlights of the 2011 Statistical Computing conference, which took place from May 22 to May 25, 2011. The main topics of this year's workshop were "Clustering",

H. Binder (✉)
Division Biostatistics, Institute of Medical Biostatistics, Epidemiology and Informatics,
University Medical Center Johannes Gutenberg University Mainz,
Obere Zahlbacher Str. 69, 55131 Mainz, Germany
e-mail: binderh@uni-mainz.de

H. A. Kestler
Research Group Bioinformatics and Systems Biology, Institute of Neural Information Processing,
University of Ulm, 89069 Ulm, Germany
e-mail: hans.kestler@uni-ulm.de

M. Schmid
Institut für Statistik, Working Group Computational Statistics, Ludwig-Maximilians-
Universität München, Ludwigstraße 33, München 80539, Germany
e-mail: matthias.schmid@stat.uni-muenchen.de

"Integrating and analyzing multiple data sources", "Regularized estimation techniques", "Visualization", "Information fusion", and "Model evaluation".

The conference provided a tutorial on boosting, which resulted in the paper by Hofner et al. (2014) that illustrates the use of the various functionalities provided by the "mboost" package. The paper on the contribution of Krey et al. (2014) nicely extends the work presented in a previous Reisensburg special issue. Starting from the observation, made in the earlier paper, that clustering is useful for extracting representative features of sound, e.g. for instrument classification, the new paper incorporates the temporal structure into clustering, i.e. adapts clustering using domain knowledge for increasing stability. This is an instance where the data at hand is enriched by additional knowledge. Another type of external knowledge is considered by Hieke et al. (2014), where genomic measurements on single nucleotide polymorphisms are aggregated according to the genes they are mapped to. To subsequently test for differences between cases and controls, a minimal p-value approach is chosen and implemented in as R package. The contributions of Schmidt et al. (2014) and Lausser et al. (2014) present specific approaches for compressing the information given in a data set, while still retaining the most important structure. Schmidt et al. (2014) consider ways for making graphs accessible for machine learning tasks. Specifically, they illustrate how principal eigenvalues of adjacency graphs can be used for classification. This makes it feasible to incorporate 2D images that are described by sets of points into standard statistical approaches, such as Gaussian mixture models. Lausser et al. (2014) specifically consider nearest neighbor classifiers as a machine learning approach. There they identify predictive hubs that still represent the most important structure, greatly simplifying computation and subsequent data analysis. In Maucher et al. (2014), both the merits and the limitations of higher-order correlations in the inference of Boolean networks are presented.

References

Hofner B, Mayr A, Robinzonov N, Schmid M (2014) Model-based boosting in R: a hands-on tutorial using the R package mboost. Comput Stat 29. doi:10.1007/s00180-012-0382-5

Hieke S, Binder H, Nieters A, Schumacher M (2014) minPtest: a resampling based gene region-level testing procedure for genetic case-control studies. Comput Stat 29. doi:10.1007/s00180-012-0391-4

Krey S, Ligges U, Leisch F (2014) Music and timbre segmentation by recursive constrained K-means clustering. Comput Stat 29. doi:10.1007/s00180-012-0358-5

Lausser L, Müssel C, Melkozerov A, Kestler HA (2014) Identifying predictive hubs to condense the training set of k-nearest neighbour classifiers. Comput Stat 29. doi:10.1007/s00180-012-0379-0

Maucher M, Kracht DV, Schober S, Bossert M, Kestler HA (2014) Inferring Boolean functions via higher-order correlations. Comput. Stat 29. doi:10.1007/s00180-012-0385-2

Schmidt M, Palm G, Schwenker F (2014) Spectral graph features for the classification of graphs and graph sequences. Comput Stat 29. doi:10.1007/s00180-012-0381-6