Ensemble-Methods for Classification, Regression and Survival-Analysis

Torsten Hothorn

Institut für Medizininformatik, Biometrie und Epidemiologie
Friedrich–Alexander–Universität Erlangen–Nürnberg
Waldstraße 6, D-91054 Erlangen
Torsten.Hothorn@rzmail.uni-erlangen.de

Roughly ten years after the first papers suggesting certain aggregations of an ensemble of multiple weak learners in order to build a strong learner were published, ensemble techniques are a well established tool in supervised learning applications. The methodology leads to improved prediction models in many interesting fields of applied research. Although binary classification problems were addressed most prominently in the past, the methodology is nowadays developed for regression problems as well as for problems with censored data.

In my talk I will give an overview on the most interesting and successful representatives of ensemble methods: Bagging and its stochastic version random forests as well as boosting, an iterative optimisation technique. The basic concepts will be outlined and some modifications, for example a combination of various discriminant analysis models, will be discussed. A bagging version for survival problems which uses an alternative aggregation scheme will be presented.