Machine Learning and Interpretable Machine Learning with R

Carolin Strobl, Mirka Henninger, & Yannick Rothacher

Full day short course

The aim of this course is to provide the audience with a general introduction to machine learning (ML) techniques and principles, including means for enhancing the interpretability of ML results.

We will start with a "grand tour" of supervised and unsupervised ML, touching upon a variety of ML methods, such as k-means and hierarchical clustering, k-NN, and support vector machines, as well as important concepts and terminology, such as "black box" models, cross validation and parameter tuning.

In the next part of the course we will introduce two of the most widely used families of ML methods in more detail: the ensemble methods bagging, random forests, and boosting, including their construction principles and properties, as well as neural networks, focusing on single hidden-layer, feed-forward networks, the role of activation functions, parameter tuning, and the related dangers of over- and underfitting.

Using these methods for illustration, we will further present several graphical and numeric techniques from the field of Interpretable Machine Learning (IML) that allow us to assess the importance and shape of the effect of the predictor variables. Besides presenting the techniques, we will also discuss potential caveats and risks of misinterpretation.

The course lectures will be interspersed with practical exercises, in which participants learn how to apply the presented techniques in the free, open-source software R. The participants will receive detailed instructions on how to install the free software before the course. Previous experience with R is a plus, but the course materials and presenters are prepared in a way that makes it possible to follow even for R novices.

At the end of the course participants will understand key principles of ML and IML, be able to apply several widely used machine learning methods in R, know where to be careful not to mis- or overinterpret results, and be able to judge if and how machine learning could contribute to their own research.