Nonparametric classification with missing data

Abstract:
Missing data are ubiquitous in modern statistics, posing a major challenge in a plethora of applications. In the first half of the talk, I will firstly introduce the general missing data problem and describe different approaches to deal with it. I will focus in particular on classification problems, where a practitioner is presented with the task of assigning a new observation to one of two classes, based on a training set of labelled data. In the second half of the talk, I will motivate a new nonparametric framework for classification problems in the presence of missing data, and propose a new method, called the Hard-thresholding Anova Missing data (HAM) classifier, which not only has better theoretical properties than off-the-shelf classifiers, but also performs well in numerical experiments.

Torben Sell
School of Mathematics, University of Edinburgh
E-Mail: torben.sell@ed.ac.uk
Webpage: https://www.maths.ed.ac.uk/~tsell/

Kontakt:
Dr. Leonie Selk
E-Mail: leonie.selk@uni-hamburg.de
Webpage: http://www.math.uni-hamburg.de/home/selk