Minisymposium 6 - Positive definite functions and applications

Denoising and Dimension Reduction in Feature Space

KLAUS-ROBERT MÜLLER
(FRAUNHOFER FIRST.IDA, BERLIN & UNIVERSITY OF POTSDAM)

joint work with
Mikio L. Braun, Fraunhofer FIRST.IDA, Berlin
Joachim Buhmann, Department of Computer Science, ETH Zürich

We prove that the relevant information about a classification problem in feature space is contained in a rather small number of leading kernel PCA components. This new theoretical insight means that kernels not only transform data sets such that they become ideally linearly separable in feature space. Rather kernelizing is done in a manner which makes economic use of feature space dimensions, i.e. well chosen kernels provide efficient representations of the data that are effective for classification. Thus our work provides another contribution for explaining why kernel-based learning methods work well. Practically we propose an algorithm which enables us to estimate the subspace and dimensionality relevant for good classification. Our algorithm can therefore be applied (1) to analyze the interplay of data set and kernel better, (2) to help in model selection, and to (3) de-noising in feature space in order to yield better classification results. Simulations underline these claims.