***Nicole Mücke***

*(TU Berlin)*

Title:

***"Efficiency of Kernel Methods: Recent Developments and Future Perspectives"***

Abstract:

For solving supervised learning tasks such as non-parametric least-squares regression in reproducing kernel Hilbert spaces, kernel methods like Support vector machines are well established tools. Experimental and theoretical results show that they handle well small- and medium sized datasets. However, many learning tasks demand learning methods that handle large scale data sets, where observations have high dimension and/or the number of observations is large. One of the limiting factors using traditional kernel methods in the large scale setting are their super-linear computational requirements in terms of the number of training samples. In this talk I give an overview of methods to speed up the traditional setting. In particular, I present the learning properties of stochastic gradient descent (SGD) in least squares learning, in many cases far outperforming more traditional algorithms (which in many theoretical papers are used as toy models). We focus on various techniques of using data most efficiently, depending on computational resources, in particular the effect and interplay of multiple passes over the data, mini-batching and most importantly tail averaging. The results show how these different flavors of SGD can be combined to achieve optimal learning errors, also providing practical insights. Finally, I shall give an outlook how kernel methods can be extended to a more modern machine learning setting, including for instance Neural Network learning.