Kernel deep stacking networks [1] (KDSNs) are a novel method for supervised learning in biomedical research and belong to the class of deep learning methods. Deep learning uses multiple layers of non-linear transformations to derive higher abstractions of the input features [2]. These can more efficiently represent complex dependencies of joint distributions [3]. Training of deep artificial neural networks is a non-convex optimization problem, which may result in local optima and slow convergence. Kernel deep stacking networks are a computationally faster alternative, which is based on solving multiple convex optimization problems by combined kernel ridge regressions and random Fourier transformations.

Tuning of KDSNs is a challenging task, as there are multiple hyper parameters to tune. We propose a new data-driven tuning strategy for KDSNs using model based optimization (MBO) [4]. The performance criterion is RMSE on cross validation samples, and noisy Kriging is used as surrogate model. New design points are chosen by maximisation of the expected improvement criterion.

Numerical studies show, that the MBO approach is substantially faster than traditional grid search strategies. Further analysis of real data sets demonstrates, that tuned KDSNs are competitive to other state-of-the-art machine learning techniques in terms of prediction accuracy. The fitting and tuning procedures are implemented in the R package kernDeepStackNet.

New developments extend the KDSN tuning framework to include variable selection, dropout [5] and L1 penalization. Variable selection is based on the non-linear randomized dependence coefficient [6] in combination with sure independence screening [7]. Numerical simulations show that these sparse kernel deep stacking networks (SKDSN) improve generalization error compared to KDSNs. SKDSNs are also competitive to other state-of-the-art supervised learning methods in high dimensional environments, like random forests and gradient boosting.
References


