Bilevel hyperparameter optimization for nonlinear support vector machines

Alain Zemkoho
University of Southampton, United Kingdom
a.b.zemkoho@soton.ac.uk

It has now been well-established that the problem of computing optimal hyperparameters for a support vector classification problem is a bilevel optimization problem. However, so far, focus has essentially been on support vector machines with linear kernels. Unfortunately, linear kernels cannot capture more complex relationships between points of most practical data sets. In this talk, we will present a bilevel optimization model for computing hyperparameters in the context of a nonlinear support vector machine training problem. Unlike in the case with a linear kernel, the dual model of the training problem is needed to get a completely explicit problem in a finite dimensional space by means of the radial basis function. Subsequently, the Karush-Kuhn-Tucker (or mathematical program with equilibrium constraints) reformulation of the problem is introduced and the corresponding theoretical properties studied. In particular, we focus our attention on the fulfilment of the corresponding classical qualification conditions such as the Mangasarian–Fromovitz constraint qualification, linear independence constraint qualification, and strong second order sufficient condition. Some numerical illustrations will also be provided.

Joint work with Stefano Coniglio (University of Bergamo, Italy), Anthony Dunn (Decision Analysis Services Ltd, United Kingdom) and Qing-Na Li (Beijing Institute of Technology, China).