Mach Learn (2012) 87:357–380
DOI 10.1007/s10994-012-5286-7

Meta-learning for evolutionary parameter optimization of classifiers

Matthias Reif · Faisal Shafait · Andreas Dengel

Received: 16 September 2010 / Accepted: 16 March 2012 / Published online: 13 April 2012

Abstract The performance of most of the classification algorithms on a particular dataset is highly dependent on the learning parameters used for training them. Different approaches like grid search or genetic algorithms are frequently employed to find suitable parameter values for a given dataset. Grid search has the advantage of finding more accurate solutions in general at the cost of higher computation time. Genetic algorithms, on the other hand, are able to find good solutions in less time, but the accuracy of these solutions is usually lower than those of grid search.

This paper uses ideas from meta-learning and case-based reasoning to provide good starting points to the genetic algorithm. The presented approach reaches the accuracy of grid search at a significantly lower computational cost. We performed extensive experiments for optimizing learning parameters of the Support Vector Machine (SVM) and the Random Forest classifiers on over 100 datasets from UCI and StatLib repositories. For the SVM classifier, grid search achieved an average accuracy of 81 % and took six hours for training, whereas the standard genetic algorithm obtained 74 % accuracy in close to one hour of training. Our method was able to achieve an average accuracy of 81 % in only about 45 minutes. Similar results were achieved for the Random Forest classifier. Besides a standard genetic algorithm, we also compared the presented method with three state-of-the-art optimization algorithms: Generating Set Search, Dividing Rectangles, and the Covariance Matrix Adaptation Evolution Strategy. Experimental results show that our method achieved the highest average accuracy for both classifiers. Our approach can be particularly useful when training classifiers on large datasets where grid search is not feasible.

Editor: Hendrik Blockeel.

M. Reif (✉) · F. Shafait · A. Dengel
German Research Center for Artificial Intelligence, Trippstadter Straße 122, 67663 Kaiserslautern, Germany
e-mail: matthias.reif@dfki.de

F. Shafait
e-mail: faisal.shafait@dfki.de

A. Dengel
e-mail: andreas.dengel@dfki.de