

Guest editorial: model selection and optimization in machine learning

Süreyya Özögür-Akyüz · Devrim Ünay · Alex Smola

Received: 28 June 2011 / Accepted: 1 July 2011 / Published online: 13 August 2011
© The Author(s) 2011

Machine learning methods largely benefit from optimization techniques in order to find an optimal model for future predictions and decisions. The interplay of machine learning and optimization methods is much like operations research (OR). Optimization, also called mathematical programming, is a subfield of OR. Both machine learning and OR are concerned with modeling of systems related to real-world problems. In machine learning (ML), the common practice is to use classical optimization techniques. However, due to massive and large-scale data sets faced in real world problems, optimization becomes a challenging task and traditional approaches cannot keep up with expectations. Accordingly, optimization methods adapted or integrated for machine learning tasks are needed to make ML more feasible for real world data sets. Another important challenging task is model selection. Because of the mathematical structure of the optimization model, there are parameters to be searched offline for the training data. Statistical model selection methods such as cross-validation can be very time consuming when the size and the dimension of the training data is large. Thus, developing powerful model selection methods is an important factor for the feasibility of the algorithm solving the optimization problem.

This special issue on “Model Selection and Optimization in ML” was inspired from the stream “Model Selection and Optimization Techniques in Machine Learning” organized by Kristiaan Pelckmans and Süreyya Özögür-Akyüz at 23rd European Conference on Operational Research held in Bonn, Germany, July 5–8, 2009.

S. Özögür-Akyüz (✉)

Department of Mathematics and Computer Science, Bahçeşehir University, Istanbul, Turkey
e-mail: sureyya.akyuz@bahcesehir.edu.tr

D. Ünay

Department of Electrical and Electronics Engineering, Bahçeşehir University, Istanbul, Turkey
e-mail: devrim.unay@bahcesehir.edu.tr

A. Smola

Yahoo! Research, Santa Clara, USA
e-mail: alex@smola.org

There are seven papers in this special issue which are divided into two categories. Five of them introduced novel optimization approaches for ML and the rest proposed new model selection methods for ML.

The first group of papers mainly focuses on improving machine learning methods by using optimization algorithms. The paper “A Majorization- Minimization Approach to the Sparse Generalized Eigenvalue Problem” by Bharath et al. proposes a general framework for generalized eigenvalue problem (GEV) resulting in sparse solutions for a GEV. They consider a tighter approximation instead of using l_1 -norm approximation by introducing log-likelihood of a Student’s t-distribution that is formulated as a difference of convex function programming. In the paper “Dual Coordinate Descent Methods for Logistic Regression and Maximum Entropy Models” by H. Yu, a coordinate descent method is presented to solve the dual problem of logistic regression and maximum entropy. Furthermore, theoretical convergence and numerical issues are discussed and compared with the state-of-the-art models for primal problems. “SpicyMKL—A Fast Algorithm for MKL with Thousands of Kernels” by T. Suzuki et al. introduces a new optimization algorithm for Multiple Kernel Learning, which iteratively solves smooth minimization problems which does not require to solve Support Vector Machines (SVM), Linear Programming, or Quadratic Programming internally. SpicyMKL can be viewed as a proximal minimization method and converges super-linearly. In the paper “Sparse Conjugate Directions Pursuit with Application to Fixed-Size Kernel Models” by P. Karsmakers et al. an optimization scheme for computing sparse approximate solutions for over-determined linear systems is introduced, where the main idea is to build up a conjugate set of vectors of increasing cardinality at each iteration by solving a small linear subsystem. The paper entitled “Boosted Multi-Task Learning” by O. Chapelle et al. proposes a novel algorithm for multi-task learning based on gradient boosted decision trees that applies it to web search ranking problems. The algorithm is derived using the relationship between l_1 regularization and boosting, and the study mainly focuses on the problem of learning the ranking functions for various countries in a multi-task learning framework.

The second group of papers are on model selection for machine learning. The paper entitled “Model Selection for Primal SVM” by G. Moore, introduces two types of nonsmooth optimization methods for selecting model hyperparameters in primal SVM models based on cross-validation. These approaches are scalable both in the number of hyperparameters and the number of data points, which is achieved by directly working with the primal variables without introducing any dual variables. In their study, all the work is done in the primal space, and therefore inversion of the kernel matrix is not required, and the hyperparameters and the parameters are optimized simultaneously by explicit penalized bilevel programming. The last paper in this group “Resampling Approach for Clustering Model Selection” by Z. Volkovich proposes a re-sampling method for cluster stability assessment, where clustered sample elements are considered as instances of random variables such that the partition quality is measured by probability metrics between the instances.

As demonstrated above, this special issue is dedicated to novel approaches on model selection and optimization algorithms developed for improving ML methods. Our sincere hope is that the reader will share our interest and find this special issue as inspiring as we do. Finally, we would like to thank all authors and reviewers for their careful and timely work, which assisted us throughout the paper selection process and inevitably improved the quality and readability of the papers presented here. We also wish to extend our thanks to the Editor-in-Chief, the editorial board members, and the copyeditors for their insightful comments and help in producing this issue.