Journal of Machine Learning Research 2 (2001) 95-96 Published 12/01

Introduction to the Special Issue on Kernel Methods

Nello Cristianini NELLO@SUPPORT-VECTOR.NET
BIOwulf Technologies
Berkeley, CA 94704

John Shawe-Taylor JST@CS.RHUL.AC.UK
Department of Computer Science, Royal Holloway, University of London
Egham, Surrey TW20 0EX, UK

Robert C. Williamson BOB.WILLIAMSON@ANU.EDU.AU
Research School of Information Sciences and Engineering,
Australian National University
Canberra ACT 0200 Australia

This special issue arose from a workshop held at NIPS 2000 on New Directions in Kernel Methods, though not all the submissions received were from talks at the workshop. With the great help of around forty referees we selected the following ten papers from some 28 submissions, an acceptance rate of 36%.

The high number of submissions we received illustrates the vitality and popularity of the field of kernel methods in machine learning. We are pleased to be able to support the fledgling *Journal of Machine Learning Research* in this way and to provide a rapid but refereed route to publication for the papers presented at the workshop less than a year ago.

The papers in the special issue cover a wide range of topics in kernel-based learning machines, but mostly reflect three of the main current research directions: exporting the design principles of standard Support Vector Machines to a variety of other algorithms, producing alternative and more efficient implementations, and deepening the theoretical understanding of kernel methods.

The first five papers in the special issue describe extensions of the basic algorithms:

Kernel Partial Least Squares Regression in RKHS by Roman Rosipal and Leonard J. Trejo describes the development of kernel partial least squares regression. This technique is similar to kernel PCA or latent semantic kernels, but the projection is chosen by modeling the relationship between input and output variables. The paper compares performance of a number of different projection methods and obtains encouraging results, particularly in terms of the number of dimensions required to obtain a certain level of performance.

In *Support Vector Clustering*, Asa Ben-Hur, David Horn, Hava T. Siegelmann and Vladimir Vapnik present a novel clustering method using Support Vector Machines. Data points are mapped by means of a Gaussian kernel to a high dimensional feature space, where the minimal enclosing sphere can be calculated. When mapped back to data space, this sphere can separate into several components, each enclosing a separate cluster of points. A simple algorithm for identifying these clusters is discussed and evaluated experimentally.

One-Class SVMs for Document Classification by Larry M. Manevitz and Malik Yousef provides extensive experimentation comparing the SVM approach to one-class classification of text documents with more traditional methods such as nearest neighbour, naive Bayes and one more advanced neural network method based on 'bottleneck' compression. The neural network method gave generally comparable performance to the one-class SVM and in the experiments reported proved more robust.

Uniform Object Generation for Optimizing One-Class Classifiers by David M.J. Tax and Robert P.W. Duin discusses a novelty detection algorithm (one-class classifier) for estimating the support of a data distribution as well as methods to set the tunable parameters of the algorithm.

In *A Generalized Kernel Approach to Dissimilarity Based Classification* by Elzbieta Pekalska, Pavel Paclik and Robert P.W. Duin, the philosophy of kernel based classification is extended to dissimilarity-based algorithms. Two different ways of using generalized dissimilarity kernels are discussed theoretically and evaluated experimentally.

The next four papers focus on alternative implementations of Support Vector Machines:

A New Approximate Maximal Margin Classification Algorithm by Claudio Gentile presents a new incremental algorithm that approaches the Support Vector Machine in the limit, and a mistake-bound style analysis of its convergence rate.

Efficient SVM Training Using Low-Rank Kernel Representation by Shai Fine and Katya Scheinberg presents new results that allow the solution of much larger problems (in terms of data set size) by exploiting the low effective rank of the kernel matrix.

On the Algorithmic Implementation of Multiclass Kernel-based Vector Machines by Koby Crammer and Yoram Singer describes an efficient algorithm to solve multi-class problems with an SVM-type algorithm, and presents an effective decomposition method for solving the associated quadratic programming problem.

Simplifying Support Vector Solutions by T. Downs, K.E. Gates and A. Masters presents a trick to reduce the number of support vectors (and hence increase the speed of the trained classifier) with no change to the statistical performance.

The final contribution considers general theoretical issues relative to kernel functions:

Classes of Kernels for Machine Learning: A Statistics Perspective by Marc G. Genton summarises a number of existing results on the suitability of kernels as well as defining some new classes of kernels.

Overall, we believe that these papers provide a useful snapshot of current trends in kernel methods, an area of research that has already found many practical applications, at the same time that early theoretical results are still being extended. This is certainly an indication of the maturity of a field that started with a paper at COLT 1992 (Boser et al., 1992), has grown through a series of workshops at the neural networks conference NIPS (Schölkopf et al., 1999, Smola et al., 2000) and has produced a range of algorithmic techniques that are now part of the toolbox of many machine learning practitioners.

References

B. E. Boser, I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classifiers. In D. Haussler, editor, *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pages 144–152. ACM Press, 1992.

B. Schölkopf, C. J. C. Burges, and A. J. Smola. *Advances in Kernel Methods – Support Vector Learning*. MIT Press, 1999.

A. J. Smola, P. Bartlett, B. Schölkopf, and D. Schuurmans. *Advances in Large Margin Classifiers*. MIT Press, 2000.