

Peleg, Dori; Meir, Ron

A sparsity driven kernel machine based on minimizing a generalization error bound. (English) [Zbl 1175.68385](#)

Pattern Recognition 42, No. 11, 2607-2614 (2009).

Summary: A new sparsity driven kernel classifier is presented based on the minimization of a recently derived data-dependent generalization error bound. The objective function consists of the usual hinge loss function penalizing training errors and a concave penalty function of the expansion coefficients. The problem of minimizing the non-convex bound is addressed by a successive linearization approach, whereby the problem is transformed into a sequence of linear programs. The algorithm produced comparable error rates to the standard support vector machine but significantly reduced the number of support vectors and the concomitant classification time.

MSC:

68T10 Pattern recognition, speech recognition

68T05 Learning and adaptive systems in artificial intelligence

Keywords:

sparsity; classification; generalization error bounds; statistical learning theory

Full Text: [DOI](#)

References:

- [1] Schölkopf, B.; Smola, A.J., *Learning with kernels: support vector machines, regularization, optimization, and beyond*, (2002), MIT Press Cambridge, MA
- [2] Herbrich, R., *Learning kernel classifiers, theory and algorithms*, (2002), MIT Press Cambridge
- [3] Steinwart, I., Sparseness of support vector machines, *The journal of machine learning research*, 4, 1071-1105, (2003) · [Zbl 1094.68082](#)
- [4] Gat, Y., Adaptive sparseness for supervised learning, *Machine learning*, 42, 233-239, (2001)
- [5] Jaeger, S.A., Generalization bounds and complexities based on sparsity and clustering for convex combinations of functions from random classes, *The journal of machine learning research*, 6, 307-340, (2005) · [Zbl 1222.62072](#)
- [6] Mangasarian, O.L.; Musicant, D.R., *Complementarity: applications, algorithms and extensions*, (2001), Kluwer Academic Publishers Dordrecht
- [7] Drezet, P.M.; Harrison, R.F., A new method for sparsity control in support vector classification and regression, *Pattern recognition*, 34, 111-125, (2001) · [Zbl 0970.68746](#)
- [8] Fung, G.M.; Mangasarian, O.L.; Smola, A.J., Minimal kernel classifiers, *The journal of machine learning research*, 3, 303-321, (2002) · [Zbl 1130.68321](#)
- [9] Figueiredo, M., Adaptive sparseness for supervised learning, *IEEE transactions on pattern analysis and machine intelligence*, 25, 1150-1159, (2003)
- [10] Nair, P.B.; Choudhury, A.; Keane, A.J., Some greedy learning algorithms for sparse regression and classification with Mercer kernels, *The journal of machine learning research*, 3, 781-801, (2002) · [Zbl 1089.68602](#)
- [11] Wu, M.; Schölkopf, B.; Bakir, G., A direct method for building sparse kernel learning algorithms, *Journal of machine learning research*, 7, 603-624, (2006) · [Zbl 1222.68335](#)
- [12] R. Collobert, F. Sinz, J. Weston, L. Bottou, Trading convexity for scalability, in: *ICML, 2006*, pp. 201-208.
- [13] J. Valyon, G. Horvath, A sparse least squares support vector machine classifier, in: *Proceedings IEEE International Joint Conference on Neural Networks*, vol. 1, 2004. · [Zbl 1146.68439](#)
- [14] Tipping, M.E., Sparse Bayesian learning and the relevance vector machine, *The journal of machine learning research*, 1, 211-244, (2001) · [Zbl 0997.68109](#)
- [15] M. Fazel, H. Hindi, S.P. Boyd, Log – det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices, in: *Proceedings of the American Control Conference*, vol. 3, 2003, pp. 2156-2162.
- [16] Bertsimas, D.; Tsitsiklis, J.N., *Introduction to linear optimization*, (1997), Athena Scientific
- [17] Cristianini, N.; Shawe-Taylor, J., *An introduction to support vector machines*, (2000), Cambridge university Press Cambridge
- [18] Horst, R.; Pardalos, P.M., *Handbook of global optimization*, (1995), Kluwer Academic Publishers Dordrecht · [Zbl 0805.00009](#)

- [19] Bertsekas, D.P., Nonlinear programming, (1999), Athena Scientific · [Zbl 0935.90037](#)
- [20] Vapnik, V., Statistical learning theory, (1998), Wiley Interscience New York · [Zbl 0935.62007](#)
- [21] Weston, J.; Elisseeff, A.; Schölkopf, B.; Tipping, M., Use of the zero norm with linear models and kernel methods, The journal of machine learning research, 3, 1439-1461, (2003) · [Zbl 1102.68605](#)
- [22] Boyd, S.; Vandenberghe, L., Convex optimization, (2004), Cambridge University Press Cambridge · [Zbl 1058.90049](#)
- [23] Duda, R.O.; Hart, P.E.; Stork, D.G., Pattern classification, (2000), Cambridge University Press Cambridge

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.