

Various Hyperplane Classifiers Using Kernel Feature Spaces

Kornél Kovács and András Kocsor

Research Group on Artificial Intelligence
of the Hungarian Academy of Sciences and University of Szeged
H-6720 Szeged, Aradi vértanúk tere 1., Hungary
{kkornel, kocsor}@inf.u-szeged.hu
<http://www.inf.u-szeged.hu/speech>

Extended Abstract

In machine learning the classification approach may be linear or nonlinear, but it seems that by using the so-called kernel idea, linear methods can be readily generalized to the nonlinear ones. The key idea was originally presented in Aizermann's paper [1] and it was successfully renewed in the context of the ubiquitous Support Vector Machines (SVM) [2]. The roots of SV methods can be traced back to the need for the determination of the optimal parameters of a separating hyperplane, which can be formulated both in input space or in kernel induced feature spaces. While the former is a linear method, the latter results in a nonlinear counterpart. Optimality can vary from method to method and SVM is just one of several possible approaches.

In this paper we present a new family of hyperplane classifiers, that make use of various contrast functions for different optimality aspects. However, in contrast to SVM - where a constrained quadratic optimization is used - some of the proposed methods lead to the unconstrained minimization of convex functions while others merely require solving a linear system of equations similar to [3]. Unconstrained convex minimization arises when a convex penalty function is applied to non-separated samples, while the other case occurs for an implicitly defined regression hyperplane. We should add that in our regression-based approach not only the features but the class information is transformed to kernel feature spaces as well.

The main consequence of the paper is that in numerous cases our algorithms proved to be the more beneficial to the classification task out of the methods examined.

References

1. M. AIZERMANN, E. BRAVERMAN, AND L. ROZONOER *Theoretical foundations of the potential function method in pattern recognition learning*, Automation and Remote Control 25:821-837, 1964.
2. VAPNIK, V. N. *Statistical Learning Theory*, John Wiley & Sons Inc., 1998.
3. LEE, Y.-J. AND MANGASARIAN, O. L. *SSVM: A Smooth Support Vector Machine for Classification*, Philadelphia INFORMS, November 7-10, 1999.