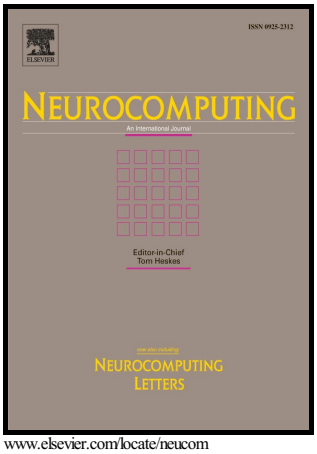

www.elsevier.com/locate/neucom

Group-penalized feature selection and robust twin SVM classification via second-order cone programming

Julio López, Sebastián Maldonado

PII: S0925-2312(17)30008-5
DOI: http://dx.doi.org/10.1016/j.neucom.2017.01.005
Reference: NEUCOM17897

To appear in: *Neurocomputing*

Received date: 30 June 2016
Revised date: 2 November 2016
Accepted date: 1 January 2017

Cite this article as: Julio López and Sebastián Maldonado, Group-penalize feature selection and robust twin SVM classification via second-order con programming, *Neurocomputing,* http://dx.doi.org/10.1016/j.neucom.2017.01.00

Group-Penalized Feature Selection and Robust Twin SVM Classification via Second-order Cone Programming

Julio López[a], Sebastián Maldonado[b,*]

[a] *Facultad de Ingeniería, Universidad Diego Portales, Ejército 441, Santiago, Chile.*
[b] *Facultad de Ingeniería y Ciencias Aplicadas, Universidad de los Andes Monseñor Álvaro del Portillo 12455, Las Condes, Santiago, Chile.*

Abstract

Selecting the relevant factors in a particular domain is of utmost interest in the machine learning community. This paper concerns the feature selection process for twin support vector machine (TWSVM), a powerful classification method that constructs two nonparallel hyperplanes in order to define a classification rule. Besides the Euclidean norm, our proposal includes a second regularizer that aims at eliminating variables in both twin hyperplanes in a synchronized fashion. The baseline classifier is a twin SVM implementation based on second-order cone programming, which confers robustness to the approach and leads to potentially better predictive performance compared to the standard TWSVM formulation. The proposal is studied empirically and compared with well-known feature selection methods using microarray datasets, on which it succeeds at finding low-dimensional solutions with highest average performance among all the other methods studied in this work.

Keywords: Support vector machines, Feature selection, Twin SVM, Second-order cone programming, Group penalty.

1. Introduction

Robustness has been a relevant topic in the SVM literature in recent years [11, 30]. Second-order cone programming (SOCP) [2, 23] is a popular convex optimization approach that has been used to develop robust maximum margin classifiers [5, 29]. In the scheme presented by Nath and Bhattacharyya [29], the worst data distribution is assumed for a given mean and covariance matrix, while each training pattern is classified correctly for predefined false positive and false negative error rates. This strategy has demonstrated superior predictive performance thanks to its robust framework [5, 25, 29].

*Corresponding author. Tel: +56-2-26181874, E-mail: smaldonado@uandes.cl.