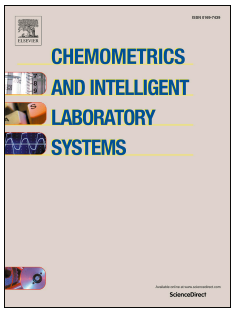
Accepted Manuscript



Support vector machine with truncated pinball loss and its application in pattern recognition

Liming Yang, Hongwei Dong

PII: S0169-7439(17)30282-4

DOI: 10.1016/j.chemolab.2018.04.003

Reference: CHEMOM 3613

To appear in: *Chemometrics and Intelligent Laboratory Systems*

Received Date: 2 May 2017

Revised Date: 15 March 2018

Accepted Date: 3 April 2018

Please cite this article as: L. Yang, H. Dong, Support vector machine with truncated pinball loss and its application in pattern recognition, *Chemometrics and Intelligent Laboratory Systems* (2018), doi: 10.1016/j.chemolab.2018.04.003.

Support vector machine with truncated pinball loss and its application in pattern recognition

Liming Yang, Hongwei Dong

College of Science,China Agricultural University, Beijing, 100083, China.
E-mail: cauyanglm@163.com;donghongwei1994@163.com

Abstract

Support vector machine(SVM) with pinball loss(PINSVM) has been recently proposed and shown its advantages in pattern recognition. In this paper, we present a robust bounded loss function (called L_t-loss) that truncates pinball loss function. Then a novel robust SVM formulation with L_t-loss(called TPINSVM) is proposed to enhance noise robustness. Moreover, we demonstrate that the proposed TPINSVM satisfies Bayes rule and it has a certain sparseness. However, the non-convexity of the proposed TPINSVM makes it difficult to optimize. We develop a continuous optimization method, DC(difference of convex functions) programming method, to solve the proposed TPINSVM. The resulting DC optimization algorithm converges finitely. Furthermore, the proposed TPINSVM is directly applied to recognize the purity of hybrid maize seeds using near-infrared spectral data. Experiments show that the proposed method achieves better performance than the traditional methods in most spectral regions. Meanwhile we simulate the proposed TPINSVM in benchmark datasets in different situations. In noiseless setting, the proposed TPINSVM either improves or shows no significant difference in generalization compared to the traditional approaches. While in noise situations, TPINSVM improves generalization in most cases.

Keywords: Support vector machine, Robustness, Non-convex loss, Pinball loss, DC programming, Quantile, Classification

1. Introduction

The traditional support vector machine(called CSVM)[1]-[2] has been widely applied because of its solid theoretical foundation and good generalization. SVMs have many good properties such as kernel skill, sparsity and