The probabilistic constraints in the support vector machine

Hadi Sadoghi Yazdi a,*, Sohrab Effati b, Zahra Saberi b

a Engineering Department, Tarbiat Modaless University of Sabzevar, Sabzevar, Iran
b Department of Mathematics, Tarbiat Modaless University of Sabzevar, Sabzevar, Iran

Abstract

In this paper, a new support vector machine classifier with probabilistic constrains is proposed which presence probability of samples in each class is determined based on a distribution function. Noise is caused incorrect calculation of support vectors thereupon margin can not be maximized. In the proposed method, constraints boundaries and constraints occurrence have probability density functions which it help for achieving maximum margin. Experimental results show superiority of the probabilistic constraints support vector machine (PC-SVM) relative to standard SVM.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Probabilistic constraints; Support vector machine; Margin maximization

1. Introduction

Such learning only aims at minimizing the classification error in the training phase, and it cannot guarantee the lowest error rate in the testing phase. In statistical learning theory, the support vector machine (SVM) has been developed for solving this bottleneck. Support vector machines (SVMs) as originally introduced by Vapnik within the area of statistical learning theory and structural risk minimization [1] and create a classifier with minimized VC dimension. It have proven to work successfully on wide range applications of nonlinear classification and function estimation such as optical character recognition [2,3], text categorization [4], face detection in images [5], vehicle tracking in video sequence [6], nonlinear equalization in communication systems [7], and generating of fuzzy rule based system using SVM framework [8,9].

Basically, the support vector machine is a linear machine with some very nice properties. It is not possible for such a set of training data to construct a separating hyperplane without encountering classification error. In this case a set of slack variable are used for samples that reduce confidence interval. In this case, it may be formulated to a dual problem form and so slack variable is not appeared in the dual problem and is converted to separable form. Main motivation of this paper rely on probabilistic constraints and obtained results include asymmetric margin depend on to probability density function of the data classes and importance of each samples in determination of hyperplane parameters.

* Corresponding author.
E-mail address: sadoghi@stu.ac.ir (H.S. Yazdi).

0096-3003/$ - see front matter © 2007 Elsevier Inc. All rights reserved.