Neurocomputing xx (xxxx) xxxx–xxxx

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

# An improved multiple birth support vector machine for pattern classification

Xiekai Zhang[a,b], Shifei Ding[a,b,*], Yu Xue[c]

[a] School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China
[b] Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100090, China
[c] School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China

ARTICLE INFO

Communicated by Dr. J. Kwok

*Keywords:*
Support vector machine
Twin support vector machine
Multiple birth support vector machine
Multi-class classification
Smoothing technique

ABSTRACT

Multiple birth support vector machine is a novel machine learning algorithm for multi-class classification, which is considered as an extension of twin support vector machine. Compared with training speeds of other multi-class classifiers based on twin support vector machine, the training speed of multiple birth support vector machine is faster, especially when the number of class is large. However, one of the disadvantages of multiple birth support vector machine is that when used to deal with some datasets such as “Cross planes” datasets, multiple birth support vector machine is likely to get bad results. In order to deal with this, we propose an improved multiple birth support vector machine. We add a modified item into multiple birth support vector machine to make the variance of the distances from each samples of a given class to their hyperplanes as small as possible. To predict a new sample, our method first determines an interval for each class depending on the distances between training samples and their hyperplanes, and then classifies the new sample depending on the distances between hyperplanes and the new sample which are in the corresponding intervals. In addition, smoothing technique is applied on our model, the first time it was used in multi-class twin support vector machine. The experimental results on artificial datasets and UCI datasets show that the proposed algorithm is efficient and has good classification performance.

## 1. Introduction

Support vector machine (SVM), which was originally proposed by Vapnik and his coworkers [1,2], is a novel machine learning algorithm based on the statistical learning theory [3] and the Vapnik-Chervonenkis (VC) dimension. Since SVM follows the Structural Risk Minimization (SRM) principle, it reduces the occurrence of risk during the training phase as well as enhances its generalization capability. Due to its high performance, it has since been greatly generalized [4–7], and applied to a variety of practical problems [8–12]. Among numerous improved SVMs, twin support vector machine (TWSVM) has drawn an extensive attention of scholars since it is almost four times faster than standard SVM classifier [13]. Different from traditional SVM, TWSVM generates two nonparallel hyperplanes by solving two quadratic programming problems (QPPs) in small size such that each plane is close to one of the two classes and is as far as possible from the other class [14]. Currently, the research of TWSVM has also made a great progress [15–28]. One of the most important extensions is multiple birth support vector machine (MBSVM) for multi-class classification [29]. Before multiple birth support vector machine was proposed, several multi-class TWSVM approaches have been presented [30–37], Chu et al. [30] proposed a multi-class TWSVM by extending TWSVM straightly from binary classification to multi-class classification. This approach handles the samples from one class as positive samples while considering all other samples that do not belong to the class as negative samples, and then trains them in turn to construct one binary TWSVM classifier for each class. Xu et al. [31] proposed a new multi-class classification algorithm, called twin support vector classification machine for k-class classification (Twin-KSVC), which evaluates all the training points into a “1-versus-1-versus-rest” structure. However, they both have not kept the advantage of TWSVM that has lower computational complexity than that of the standard SVM. For K-class classification problem, MBSVM with much lower computational complexity constructs K hyperplanes, one for each class, by solving K smaller size of QPPs. Compared with “one-versus-all” multi-class TWSVM, MBSVM has fewer constraints, leading to a lower complexity. The “one-versus-one” multi-class TWSVM classifier (1-v-1 TWSVM) has to solve K(K-1) QPPs while MBSVM only needs to solve K QPPs.

However, MBSVM cannot well deal with the kind of dataset in which the samples from different classes have some intersections, such

* Corresponding author at: School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China.

http://dx.doi.org/10.1016/j.neucom.2016.11.006
Received 8 November 2015; Received in revised form 28 August 2016; Accepted 14 November 2016
Available online xxxx
0925-2312/ © 2016 Elsevier B.V. All rights reserved.

Please cite this article as: Zhang, X., Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.11.006

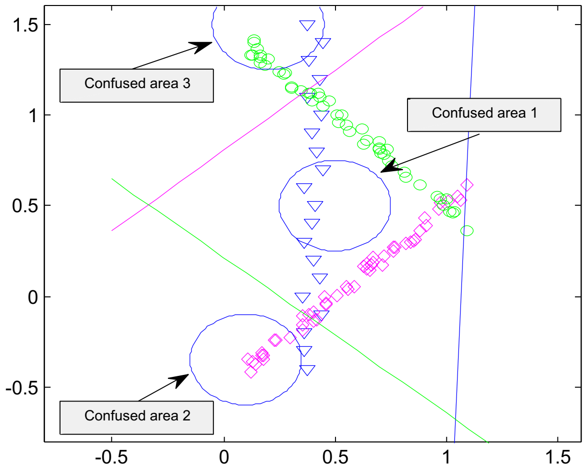


**Fig. 1.** The disadvantage of MBSVM.

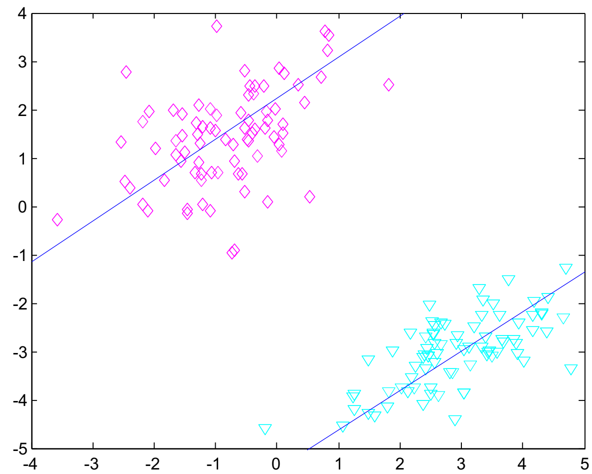

**Fig. 2.** The basic thought of TWSVM.

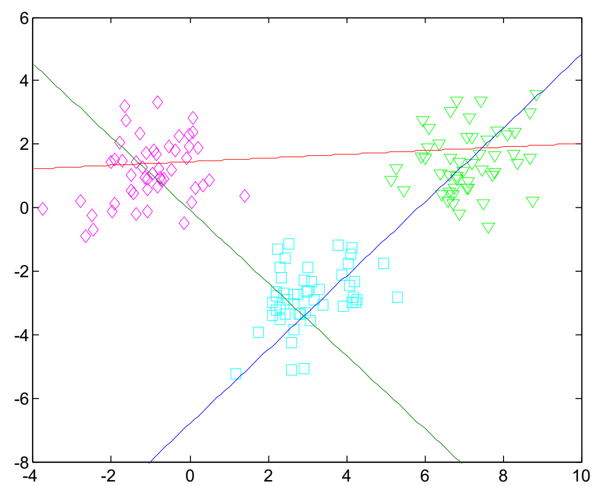

**Fig. 3.** The basic thought of MBSVM.

as the dataset showed in Fig. 1. MBSVM assigns the new sample to class label according to which hyperplane the new sample is farthest to. Fig. 1 shows that MBSVM may misclassify not only samples in cross regions and "confused area1" that other multi-class TWSVM approaches can hardly handle too, but also those samples in "confused area 2" and "confused area 3" that can be classified correctly by other multi-class TWSVM approaches. Take the blue "▽" sample in "confused area 3" as an example. The distance between this point and the blue line which is corresponding to its class is smaller than the distance between it and the green line, and this sample will been misclassified. After careful observation, we find that the distances between a hyperplane of MBSVM and samples of corresponding class are similar which conforms to that samples of one class always get together according to certain law in some way. In this paper, we improve the decision criterion of MBSVM. Our decision criterion first determines an interval for each class depending on the distances between training samples and the hyperplanes. When the distance between a hyperplane and a new sample is in the corresponding interval, the corresponding class will be prioritized. Then our method classifies the new sample depending on the distances between prioritized hyperplanes and the new sample. In order to make the distances between training samples of a class and their hyperplane are in an appropriate range, we add a modified item into multiple birth support vector machine to make the variance of the distances from each samples of a given class to their hyperplanes as small as possible. In addition, we convert the QPPs of the improved MBSVM into unconstrained minimization problems and apply smoothing technique, which has already been successfully applied to TWSVM, to solve the unconstrained minimization problems. Compared with existing extensions of TWSVM for multi-category classification, there are several advantages in the proposed approach. (1) The proposed approach follows the thought of "All-versus-One", which helps to reduce the complexity. (2) A modified item is introduced in the proposed approach. The modified item which is determined by a simple unconstrained minimization problem can make the improved MBSVM take into consideration more distribution information of the dataset. The proposed approach is suitable to more datasets. (3) Smoothing technique promotes the training process. It is the first time the technique is used in multi-class twin support vector machine.

The rest of this article is organized as follows. In the next section, a brief review to TWSVM and MBSVM is provided. In Section 3, we introduce an improved MBSVM classifier method and use smoothing technique to solve our model. Experimental results are depicted in Section 4. In the last section, concluding remarks and further research to be developed are presented.

## 2. Preliminaries

In this section, we give a brief review to TWSVM and MBSVM. Fig. 2 shows the basic thought of TWSVM and Fig. 3 shows the basic thought of MBSVM.

### 2.1. Twin support vector machine

Assume a binary classification problem in the $n$-dimensional real space $R^n$. The set of training data points is represented by $T=\{(x_i, y_i) \mid i=1,2...,m\}$, where $x_i$ is input and $y_i \in\{+1, -1\}$ is the corresponding output. Let matrix $A$ represent the samples belonging to class +1 and matrix $B$ represent the samples belonging to class −1. Linear TWSVM solves the binary classification problem by seeking two nonparallel separating hyperplanes.

$$x^T w_1 + b_1 = 0,\ \ x^T w_2 + b_2 = 0 \tag{1}$$

TWSVM generates the hyperplanes by solving the following QPPs:

$$\begin{aligned} &\min_{w_1,b_1,\zeta} \ \frac{1}{2}\|Aw_1 + e_1 b_1\|^2 + c_1 e_2^T \zeta \\ &\text{s.t} \quad -(Bw_1 + e_2 b_1) + \zeta \ge e_2,\ \zeta \ge 0 \\ &\min_{w_2,b_2,\eta} \ \frac{1}{2}\|Bw_2 + e_2 b_2\|^2 + c_2 e_1^T \eta \\ &\text{s.t} \quad (Aw_2 + e_1 b_2) + \eta \ge e_1,\ \eta \ge 0 \end{aligned} \tag{2}$$

Where $e_1$ and $e_2$ are column vectors of ones of appropriate dimensions, $\zeta$ and $\eta$ are slack variables, and $c_1$ and $c_2$ are penalty parameters with

Download English Version:

https://daneshyari.com/en/article/4948069

Download Persian Version:

https://daneshyari.com/article/4948069

Daneshyari.com