Mathematical and Computer Modelling 54 (2011) 1093–1099

Contents lists available at ScienceDirect

Mathematical and Computer Modelling

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

Learning SVM with weighted maximum margin criterion for classification of imbalanced data

Zhuangyuan Zhao, Ping Zhong *, Yaohong Zhao

College of Science, China Agricultural University, Beijing, 100083, PR China

ARTICLE INFO

Keywords:
Support vector machine
Imbalanced data learning
Kernel optimization
Weighted maximum margin criterion

ABSTRACT

As a kernel-based method, whether the selected kernel matches the data determines the performance of support vector machine. Conventional support vector classifiers are not suitable to the imbalanced learning tasks since they tend to classify the instances to the majority class which is the less important class. In this paper, we propose a weighted maximum margin criterion to optimize the data-dependent kernel, which makes the minority class more clustered in the induced feature space. We train support vector classification with the optimal kernel. The experimental results on nine benchmark data sets indicate the effectiveness of the proposed algorithm for imbalanced data classification problems.

1. Introduction

Learning from imbalanced data sets is an important and on-going issue in machine learning research. The class imbalance problem corresponds to domains for which one class is represented by a large number of instances while the other is represented by only a few. There are many class imbalanced problems in real-world applications [1–6]. Conventional classifiers that seek accuracy over a full range of instances are not suitable to deal with imbalanced learning tasks, since they tend to be overwhelmed by the majority class which is usually the less important class.

There are roughly two types of approaches to deal with imbalanced data classification problems. One is to sample data, either randomly or intelligently, to obtain an altered class distribution. These approaches consist of under-sampling the majority class or over-sampling the minority class, such as randomly under-sampling, randomly over-sampling, one-sided selection, cluster-based over-sampling, Wilson's editing, SMOTE, and borderline-SMOTE [7–11]. The other is to modify the standard learning algorithms. These approaches include cost-sensitive methods [12,13], margin calibration method [14], unsupervised self-organizing method [15], minimax probability machines [16,17], and one-class support vector machine [18].

Support vector machine (SVM) is an excellent kernel-based tool for classification and regression [19]. Within a few years after its introduction, SVM has already outperformed most other systems in a wide variety of applications, which include a wide spectrum of research areas ranging from pattern recognition, text categorization, biomedicine, brain–computer interface, and financial regression. However, the conventional SVM performs poorly on imbalanced learning because they pay less attention to the minority class. Classification rules for predicting the minority class tend to be fewer and weaker than those for the majority class. Consequently, testing instances in the minority class are misclassified more often than those in the majority class. However, the minority class is often the more important class in applications.

As we know, kernel plays an important role in SVM. Whether the kernel matches the data determines its performance. A kernel optimization algorithm [20] was proposed to maximize the class separability of the data in the empirical feature

* Corresponding author. Tel.: +86 10 62736511.
E-mail addresses: zping@cau.edu.cn, pingsunshine@yahoo.com.cn (P. Zhong).

0895-7177/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mcm.2010.11.040