Dynamic and Static Weighting in Classifier Fusion

Rosa M. Valdovinos[1], J. Salvador Sánchez[2], and Ricardo Barandela[1]

[1] Instituto Tecnológico de Toluca, Av. Tecnológico s/n, 52140 Metepec, México
{li_rmvr,rbarandela}@hotmail.com
[2] Dept. Llenguatges i Sistemes Informàtics, Universitat Jaume I, 12071 Castelló, Spain
sanchez@uji.es

Abstract. When a Multiple Classifier System is employed, one of the most popular methods to accomplish the classifier fusion is the simple majority voting. However, when the performance of the ensemble members is not uniform, the efficiency of this type of voting is affected negatively. In this paper, a comparison between simple and weighted voting (both dynamic and static) is presented. New weighting methods, mainly in the direction of the dynamic approach, are also introduced. Experimental results with several real-problem data sets demonstrate the advantages of the weighting strategies over the simple voting scheme. When comparing the dynamic and the static approaches, results show that the dynamic weighting is superior to the static strategy in terms of classification accuracy.

1 Introduction

A multiple classifier system (MCS) is a set of individual classifiers whose decisions are combined when classifying new patterns. There are many different reasons for combining multiple classifiers to solve a given learning problem [6], [12]. First, MCSs try to exploit the local different behavior of the individual classifiers to improve the accuracy of the overall system. Second, in some cases MCS might not be better than the single best classifier but can diminish or eliminate the risk of picking an inadequate single classifier. Another reason for using MCS arises from the limited representational capability of learning algorithms. It is possible that the classifier space considered for the problem does not contain the optimal classifier.

Let D = { $D_1, ..., D_h$ } be a set of classifiers. Each classifier assigns an input feature vector $\mathbf{x} \in \Re^n$ to one of the c problem classes. The output of a MCS is an h-dimensional vector containing the decisions of each of the h individual classifiers:

$$[D_1(\mathbf{x}),..., D_h(\mathbf{x})]^T \tag{1}$$

It is accepted that there are two main strategies in combining classifiers: selection and fusion. In classifier selection, each individual classifier is supposed to be an expert in a part of the feature space and therefore, we select only one classifier to label the input vector **x**. In classifier fusion, each component is supposed to have knowledge of the whole feature space and correspondingly, all individual classifiers decide the label of the input vector.

Focusing on the fusion strategy, the combination can be made in many different ways. The simplest one employs the majority rule in a plain voting system [4]. More

J.S. Marques et al. (Eds.): IbPRIA 2005, LNCS 3523, pp. 59–66, 2005.

elaborated schemes use weighted voting rules, in which each individual component is associated with a different weight [5]. The final decision can be made by majority, average [6], minority, medium [7], product of votes, or using some other more complex methods [8], [9], [10], [19].

In the present work, some methods for weighting the individual components in a MCS are proposed, and their effectiveness is empirically tested over real data sets. Three of these methods correspond to the so-called dynamic weighting, by using the distances to a pattern. The last method, which belongs to the static weighting strategy, estimates the leaving-one-out error produced by each classifier in order to set the weights of each component [21].

From now on, the rest of the paper is organized as follows. Sect. 2 provides a brief review of the main issues related to classifier fusion and makes a very simple categorization of weighting methods, distinguishing between dynamic and static weighting of classifiers. Moreover, seveal weighting procedures are also introduced in Sect. 2. The experimental results are discussed in Sect. 3. Finally, some conclusions and possible further extensions are given in Sect. 4.

2 Classifier Fusion

As pointed out in Sect. 1, classifier fusion assumes that all individual classifiers are competitive, instead of complementary. For this reason, each component takes part in the decision of classifying an input test pattern.

In the simple voting (by majority), the final decision is taken according to the number of votes given by the individual classifiers to each one of the classes, thus assigning the test pattern to the class that has obtained a majority of votes. When working with data sets that contain more than two classes, in the final decision ties among some classes are very frequently obtained. To solve this problem, several criteria can be considered. For instance, to randomly take the decision, or to implement an additional classifier whose ultimate goal is to bias the decision toward a certain class [15].

An important issue that has strongly called the attention of many researchers is the error rate associated to the simple voting method and to the individual components of a MCS. Hansen and Salomon [17] show that if each one of the classifiers being combined has an error rate less than 50%, it may be expected that the accuracy of the ensemble improve when more components are added to the system. However, this assumption not always is fulfilled. In this context, Matan [18] asserts that in some cases, the simple voting might perform even worse than any of the members of the MCS. Thus some weighting method can be employed in order to partially overcome these difficulties.

A weighted voting method has the potential to make the MCS more robust to the choice of the number of individual classifiers. Two general approaches to weighting can be remarked: dynamic weighting and static weighting of classifiers. In the dynamic strategy, the weights assigned to the individual classifiers can change for each test pattern. On the contrary, in the static weighting, the weights are computed for each classifier in the training phase, and they are maintained constant during the classification of the test patterns.