International Journal of Approximate Reasoning 73 (2016) 108–122

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Enhancing evolutionary fuzzy systems for multi-class problems: Distance-based relative competence weighting with truncated confidences (DRCW-TC) ☆

Alberto Fernández [a,*], Mikel Elkano [b], Mikel Galar [b], José Antonio Sanz [b], Saleh Alshomrani [d,f], Humberto Bustince [b,c], Francisco Herrera [e,f]

[a] Department of Computer Science, University of Jaén, 23071, Jaén, Spain
[b] Departamento de Automática y Computación, Universidad Pública de Navarra, 31006, Pamplona, Spain
[c] Institute of Smart Cities (ISC), Universidad Pública de Navarra, 31006, Pamplona, Spain
[d] Faculty of Computing and Information Technology, University of Jeddah, 21589, Jeddah, Saudi Arabia
[e] Department of Computer Science and Artificial Intelligence, University of Granada, 18071, Granada, Spain
[f] Faculty of Computing and Information Technology, King Abdulaziz University, 21589, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 28 August 2015
Received in revised form 21 January 2016
Accepted 9 February 2016
Available online 23 February 2016

Keywords:
Multi-class classification
Pairwise learning
One-vs-One
Classifier selection
Evolutionary fuzzy systems
Fuzzy rule based classification systems

ABSTRACT

Classification problems with multiple classes suppose a challenge in Data Mining tasks. There is a difficulty inherent to the learning process when trying to find the most adequate discrimination functions among the different concepts within the dataset. Using Fuzzy Rule Based Classification Systems in general, and Evolutionary Fuzzy Systems in particular, provide the advantage of describing smoother borderline areas, thanks to the linguistic label-based representation.

In multi-classification, the pairwise learning approach (One-vs-One) has gained a notorious attention. However, there is certain dependence between the goodness of the confidence degrees or scores of binary classifiers, and the final performance shown by the global model. Regarding this fact, the problem of non-competent classifiers is of special relevance. It occurs when a binary classifier outputs a positive score for a couple of classes unrelated with the input example, which may degrade the final accuracy. Precisely, the previously exposed properties of fuzzy classifiers make them more prone to the former condition.

In this paper, we propose an extension of the distance-based combination strategy to overcome this non-competence problem. It is based on the truncation of the confidence degrees of the classes prior to the distance-based tuning. This allows taking advantage of the good classification abilities of Evolutionary Fuzzy Systems, while diminishing the adverse effect of the aforementioned non-competence. Experimental results, using FARC-HD with overlap functions as the fuzzy learning algorithm, show that this new adaptation of the Distance-based Relative Competence Weighting model outperforms both the OVO and standard distance-based approaches, and it is competitive with robust classifiers such as Support Vector Machines.

☆ This paper is part of the virtual special issue on IFSA – EUSFLAT 2015, edited by Oscar Cordón, Luis Magdalena and José M. Alonso.

* Corresponding author. Tel.: +34 953 213016; fax: +34 953 212472.
E-mail addresses: alberto.fernandez@ujaen.es (A. Fernández), mikel.elkano@unavarra.es (M. Elkano), mikel.galar@unavarra.es (M. Galar), joseantonio.sanz@unavarra.es (J.A. Sanz), sshomrani@kau.edu.sa (S. Alshomrani), bustince@unavarra.es (H. Bustince), herrera@decsai.ugr.es (F. Herrera).

http://dx.doi.org/10.1016/j.ijar.2016.02.005
0888-613X/

1. Introduction

Among all Data Mining tasks, classification can be regarded as one of the most studied problems [1]. It consists in learning a mapping function from the attributes of a set of examples whose class is a priori known. The final goal is to automatically determine the class for every new pattern arriving to the system. In a standard framework, the classifier is devoted to learn a dichotomization function. However, when a multi-class problem is considered, the complexity inherent to the higher number of decision boundaries, and the overlapping among these classes, makes this task harder than the binary case [2].

In accordance with the former, a successful way to address this problem is using a divide and conquer methodology. The main idea is to apply a decomposition strategy [3], i.e. to simplify the initial multi-class problems by generating several binary sub-problems. There are several strategies that consider a divide-and-conquer scheme, and most of them can be included within Error Correcting Output Codes framework [4,5]. One of the most common approaches is the One-vs-One (OVO) or pairwise learning scheme [6,7]. This approach divides the original problem in as many pairs of classes as possible, ignoring the examples that do not belong to the related classes. Then, these are learnt in an independent way by the so-called base learners or base classifiers of the ensemble [8]. In fact, even when the base classifiers are able to cope with multi-class problems, it has been shown that the use of binarization techniques allows for the achievement of a more robust system, and thus the improvement of the performance with respect to the standard case [9,10].

The simplification of the learning task implies that more attention must be focused on the combination stage, where the classifiers learned to solve each binary problem should be aggregated to output the final decision over the class label [11,12]. Specifically, every new example to be classified is queried by all the classifiers. Each classifier outputs a pair of "scores" (confidence degrees) for the two classes that were considered for the training stage. All these score values are collected to build a decision matrix which will be used to determine the output class.[1]

When the confidence degrees generated by the baseline classifiers are good enough, then any aggregation method will be adequate to determine the correct class. In particular, the simplest and most widely used method in pairwise learning the "Weighted Voting" (WV) [13], which considers the maximum vote among the summation of the scores for the binary classifiers associated with the same class.

If we analyze this procedure in depth, we find an inherent problem: all base classifiers will be fired for a given instance, even when they have not been trained to recognize the real class of this particular instance. Therefore, these classifiers could submit an erroneous score that might alter the decision process degrading the accuracy of the final system. This case is better known as the "non-competent classifiers problem" [13], which can mislead the correct labeling of the query example.

It is straightforward to acknowledge that the competence of a classifier in the OVO approach cannot be established a priori. Therefore, the techniques that have been proposed in the specialized literature to cope with this problem are based on dynamic post-processing methodology. They are aimed at adapting the score matrix values prior to the decision step. In [14], authors compute the closest classes to the query instance in order to remove those classifiers that were not related with this query example. A similar approach was developed in [15], known as Distance-based Relative Competence Weighting (DRCW), where confidence degrees of the classifiers were weighted according to the distance computed from the example to the nearest neighbors of each class, instead of directly removing the classifiers.

Among different classification algorithms, Fuzzy Rule Based Classification Systems (FRBCSs) [16] have shown a good behavior when modeling complex problems due to the proper management of the uncertainty achieved by fuzzy sets, as well as their interpretability based on the linguistic variables [17]. Furthermore, FRBCSs can be further enhanced towards more accurate systems by including the learning and adaptation capabilities of evolutionary optimization, leading to Evolutionary Fuzzy Systems (EFSs) [18]. Their success among other Soft Computing techniques, is related to their smoothness when defining the borderline areas in complex problems [19,20].

The previous fuzzy properties stressed for this type of classifiers make them more prone to the non-competence problem. Therefore, our hypothesis is that the confidence degrees computed by the EFS classification algorithms within the OVO approach may have a higher benefit in cooperation with the distance weighting scheme. In accordance with the former, in this research our objective is three-fold:

1. First, we want to analyze experimentally whether the combination of DRCW and EFSs is able to obtain a significant improvement in the performance of the classifier.
2. Then, we will propose a modification of the DRCW-OVO model by truncating the confidence degrees in the score matrix to 0.0 and 1.0, as in "Simple Voting". The objective is to combine the good behavior shown by both schemes. This new aggregation will be noted as DRCW-TC.
3. Finally, we will contrast the performance of this new approach over different paradigms of classifiers, including Decision Trees, Support Vector Machines (SVMs), and of course EFSs. We suggest that our model will achieve a better synergy with fuzzy based classifiers, allowing it to improve the results of the crisp approaches.

[1] Throughout this document, we will refer interchangeably to "confidence degrees" or "scores" or as those values assigned to the pair of classes addressed by each single classifier. In the case of fuzzy classifiers, these values are computed as the pattern classification soundness degree for each class.