Neurocomputing 166 (2015) 367–381

Contents lists available at ScienceDirect

Neurocomputing

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

BeeOWA: A novel approach based on ABC algorithm and induced OWA operators for constructing one-class classifier ensembles

Elham Parhizkar, Mahdi Abadi *

Faculty of Electrical and Computer Engineering, Tarbiat Modares University, P.O. Box 14115-194, Tehran, Iran

ARTICLE INFO

Article history:
Received 3 September 2014
Received in revised form
10 January 2015
Accepted 17 March 2015
Communicated by A. Suarez
Available online 10 April 2015

Keywords:
One-class classification
Classifier ensemble
Classifier pruning
Classifier fusion
Binary artificial bee colony
Induced OWA operator

ABSTRACT

In recent years, classifier ensembles have received increasing attention in the machine learning and pattern recognition communities. However, constructing classifier ensembles for one-class classification problems has still remained as a challenging research topic. To pursue this line of research, we need to address issues on how to generate a set of diverse one-class classifiers that are individually accurate and how to combine the outputs of them in an effective way. In this paper, we present BeeOWA, a novel approach to construct highly accurate one-class classifier ensembles. It uses a novel binary artificial bee colony algorithm, called BeePruner, to prune an initial one-class classifier ensemble and find a near-optimal sub-ensemble of base classifiers in a reasonable computational time. To evaluate the fitness of an ensemble solution, BeePruner uses two different measures: an exponential consistency measure and a non-pairwise diversity measure based on the Kappa inter-rater agreement. After one-class classifier pruning, BeeOWA uses a novel exponential induced OWA (ordered weighted averaging) operator, called EIOWA, to combine the outputs of base classifiers in the sub-ensemble. The results of experiments carried out on a number of benchmark datasets show that BeeOWA can outperform several state-of-the-art approaches, both in terms of classification performance and statistical significance.

1. Introduction

Combining multiple classifiers, also known as *classifier ensemble*, is an effective technique for solving classification problems using an ensemble of individual base classifiers. It has been theoretically and empirically demonstrated that classifier ensembles can substantially improve the classification accuracy of their constituent members [1–3].

In addition to the accuracy of base classifiers, the success of a classifier ensemble also relies on the diversity being inherent in the base classifiers [4,5]. Diversity ensures that all the base classifiers are as different from each other as possible and so can make uncorrelated errors. To achieve an initial diversity, a common technique is to train a set of homogeneous or heterogeneous base classifiers on different or same training datasets using techniques such as bagging [6], boosting [7], or random subspacing [8]. Bagging generates several different training datasets with bootstrap sampling from the original training dataset and then trains a base classifier from each of those training datasets. Boosting generates a sequence of base classifiers whose training datasets are different and determined by the accuracy of former base classifiers. Random subspacing trains base classifiers independently on the same training dataset using different random subsets of features.

Using a subset, or sub-ensemble, of base classifiers could provide higher diversity and accuracy than using the whole set, or ensemble. Thus, one of the most important issues in constructing a classifier ensemble is to decide which ones of the base classifiers to choose [9,10]. This process is also known as *classifier pruning* or *ensemble pruning* [11] and can be considered as an optimization problem with two objectives, classification accuracy and diversity, that both need to be maximized. When the size of a classifier ensemble is relatively large, classifier pruning is computationally expensive or even prohibitive. One solution to this problem is to use meta-heuristic algorithms, such as artificial bee colony (ABC) [12]. These algorithms can find near-optimal solutions in a short time. ABC is a new swarm based meta-heuristic algorithm that was initially proposed for solving numerical optimization problems. It is as simple as particle swarm optimization (PSO) [13] and differential evolution (DE) [14], and uses only common control parameters, such as population size and maximum cycle number. ABC has shown promising results in the field of optimization [15,16].

Furthermore, it has been shown that increasing the coverage of a classifier ensemble through classifier pruning is not enough to

* Corresponding author. Tel.: +98 21 82884935; fax: +98 21 82884325.
E-mail addresses: e.parhizkar@modares.ac.ir (E. Parhizkar), abadi@modares.ac.ir (M. Abadi).

http://dx.doi.org/10.1016/j.neucom.2015.03.051
0925-2312/

increase the classification accuracy [17,18]. Hence, another important step in constructing a classifier ensemble is to choose a good strategy for combining the outputs of base classifiers, using a process also known as *classifier fusion*. In the literature, several classifier fusion strategies have been proposed, which can be categorized according to the level of classifier outputs into abstract level, rank level, and measurement level [19,20]. Measurement level outputs provide more information than the other types of outputs and a number of aggregation functions or fusion rules, such as *mean*, *max*, and *product* are employed for combining them [21].

Aggregation of different pieces of information obtained from different sources is a common aspect of any fusion system. A very interesting class of powerful aggregation operators is called the ordered weighted averaging (OWA) [22]. An OWA operator takes multiple values as input and returns a single value that is a weighted sum based on an order ranking of all input values. Classifier fusion using OWA operators seems more robust than the simple weighted averaging, where the coefficients are derived based on the classifier accuracy [23].

In recent years, there has been a substantial amount of research conducted in the field of one-class classification, resulting in different one-class classifiers, including one-class SVM (OCSVM) [24], support vector data description (SVDD) [25], and so on. The goal of one-class classification is to distinguish a set of target objects from all the other possible objects [26]. Since in one-class classification problems we have only the information of the target class, therefore constructing highly accurate one-class classifier ensembles is more challenging than constructing multi-class classifier ensembles.

As mentioned before, to construct classifier ensembles, we need to address issues on how to generate a set of accurate and diverse base classifiers and how to combine the outputs of them using an effective fusion rule. Although these issues have been adequately addressed in multi-class classifier ensembles, relatively little work has been reported in the literature to address them in one-class classifier ensembles. In this paper, we present BeeOWA, a novel approach to construct highly accurate one-class classifier ensembles. BeeOWA uses a novel binary artificial bee colony algorithm, called BeePruner, to prune an initial one-class classifier ensemble and find a near-optimal sub-ensemble of base classifiers. More precisely, the goal of BeePruner is to exclude the non-diverse base classifiers from the initial ensemble and, at the same time, keep the classification accuracy. In the subsequent step, if the fusion rule does not properly utilize the ensemble diversity, then no benefit arises from the classifier fusion. Considering this fact, BeeOWA uses a novel exponential induced OWA operator, called EIOWA, to combine the outputs of base classifiers in the sub-ensemble.

The major contributions of this paper are listed as follows:

- We present a novel artificial bee colony algorithm for one-class classifier pruning that utilizes two measures simultaneously, an exponential consistency measure and a non-pairwise diversity measure based on the Kappa inter-rater agreement.
- To the best of our knowledge, the most widely used fusion rules in one-class classification problems are fixed rules, such as *majority voting*, *mean*, *max*, and *product*. We propose a novel exponential induced OWA operator for one-class classifier fusion and experimentally show it can outperform the fixed rules.
- We conduct extensive experiments on benchmark datasets to evaluate the performance of BeeOWA and show that it performs significantly better than state-of-the-art approaches in the literature.

The rest of this paper is organized as follows. Section 2 is fully dedicated to the background. Short descriptions of classifier ensembles, OWA operators, and ABC algorithm are included in this section. Section 3 presents the main steps of BeeOWA. Section 4 provides an overview of current techniques for constructing one-class classifier ensembles. The experimental results are described in Section 5. Finally, conclusions are given in Section 6.

2. Background

In this section, we give a brief introduction to some basic concepts used throughout this paper.

2.1. Classifier ensemble

To improve the performance of different classifiers, which may differ in complexity or training algorithm, an ensemble of classifiers is a viable solution. This may serve to increase the performance and also the robustness of the classification. The underlying idea of ensemble learning for classification problems is to build a number of base classifiers and then combine their outputs using a fusion rule. It has been shown that classifier ensembles outperform single classifiers for a wide range of classification problems [27–29]. The reason is that a combination of multiple classifiers reduces risks associated with choosing an insufficient single classifier.

In general, the process of constructing an ensemble of base classifiers consists of three main steps: classifier generation, classifier pruning, and classifier fusion. In the following, we briefly describe each of these steps.

2.1.1. Classifier generation

One of the key challenges for the classifier ensemble is to generate a set of diverse base classifiers. For this purpose, we can use different strategies. One effective strategy is to train homogeneous classifiers with different datasets. To do this, we can divide the original dataset into partitions or generate a number of subsets through data splitting, bagging, or boosting [6,7,30], in the hope that different classification models are generated for different distributions of the original dataset. Another strategy, also known as *ensemble feature selection*, is to train homogeneous classifiers with different subsets of features. It has been shown that simple random selection of feature subsets, called the random subspacing (RS) or random subspace method (RSM), is an effective technique for ensemble feature selection [8]. The last and intuitive strategy is to train heterogeneous classifiers with the same dataset. It should be noted that we can combine the above strategies to take the advantages of all of them, for example, we can train heterogeneous classifiers with different datasets or different subsets of features [31,32].

2.1.2. Classifier pruning

Classifier pruning, also known as *ensemble pruning*, is a useful technique for reducing the ensemble size by selecting only a subset of the base classifiers that are both accurate and have diversity in their outputs [4]. It is well-known that combining the same base classifiers does not contribute to anything apart from increasing the computational complexity of the classification. Also, combining the diverse but too weak base classifiers is unlikely to bring any benefits in the classification accuracy. Therefore, classifier pruning can be considered as a search for an optimal subset of base classifiers, making trade-off between the classification accuracy and diversity. For small ensembles, the optimal subset can be found through exhaustive search. For large ensembles, a near-optimal subset can be found using meta-heuristic optimization algorithms like genetic algorithms [33,34].

2.1.3. Classifier fusion

An important step in constructing an effective ensemble is to choose a good strategy for combining the outputs of the base