



# A new ensemble learning methodology based on hybridization of classifier ensemble selection approaches

R. Mousavi<sup>a,\*</sup>, M. Eftekhari<sup>b</sup>

<sup>a</sup> Department of Electrical and Computer Engineering, Graduate University of Advanced Technology, Kerman, Iran

<sup>b</sup> Department of Computer Engineering, Shahid Bahonar University of Kerman, Iran



## ARTICLE INFO

### Article history:

Received 12 August 2014

Received in revised form 26 June 2015

Accepted 5 September 2015

Available online 12 September 2015

### Keywords:

Ensemble learning system

Static ensemble selection

Dynamic ensemble selection

Classifier combination

Classifier diversity

Multi-objective optimization

## ABSTRACT

Ensemble learning is a system that improves the performance and robustness of the classification problems. How to combine the outputs of base classifiers is one of the fundamental challenges in ensemble learning systems. In this paper, an optimized Static Ensemble Selection (SES) approach is first proposed on the basis of NSGA-II multi-objective genetic algorithm (called SES-NSGAII), which selects the best classifiers along with their combiner, by simultaneous optimization of error and diversity objectives. In the second phase, the Dynamic Ensemble Selection-Performance (DES-P) is improved by utilizing the first proposed method. The second proposed method is a hybrid methodology that exploits the abilities of both SES and DES approaches and is named Improved DES-P (IDES-P). Accordingly, combining static and dynamic ensemble strategies as well as utilizing NSGA-II are the main contributions of this research. Findings of the present study confirm that the proposed methods outperform the other ensemble approaches over 14 datasets in terms of classification accuracy. Furthermore, the experimental results are described from the view point of Pareto front with the aim of illustrating the relationship between diversity and the over-fitting problem.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

In machine learning, Ensemble Learning Systems (ELs) are inspired from an innate behavior of humans; hence, the opinions of several experts are being collected for making a decision and then, based on these opinions, the final decision is made, especially if these decisions lead to financial, social, and medical consequences [1]. Today, ELs are important on a broad spectrum of real-world applications such as biomedical [2], financial [3], political [4], and medicine [5]. This learning method is useful in several cases, including online learning [6], incremental learning [7], fusion data [8], feature selection [9], and confidence estimation [10]. Ensemble learning is composed of three different main parts: sample selection, training the base classifiers to compose the Base Classifier Pool (BCP), and combining the BCP [11]; moreover, as ensemble learning decreases the risk of selecting a single classifier with a weak performance [1], it improves the classification accuracy in comparison with single classifier. Numerous ensemble-based algorithms have been proposed, that the most common of them are bagging,

boosting, AdaBoost, and random forest. A complete description of these algorithms can be found in [12].

How to effectively combine the classifiers outputs is a key issue in ELs. In the literature, there exist a number of combination approaches which are generally divided into two different categories: Static Classifier Ensemble (SCE) and Dynamic Classifier Ensemble (DCE). SCE determines a unified ensemble scheme for all test samples during the training phase. In general, there are three kinds of SCE strategies: Classifier Fusion (CF), Static Classifier Selection (SCS) [13,14], and Static Ensemble Selection (SES) [15]. While DCE selects one or more classifiers from the BCP on the basis of validation set for each test sample. Presently, DCE is divided into two strategies: Dynamic Classifier Selection (DCS) [16–18] and Dynamic Ensemble Selection (DES) [19–24].

In the CF strategy, there may exist inaccurate and redundant classifiers which have the potential to reduce the diversity and performance of ELs (that will, by themselves, reduce the simplicity of system) [25]. To address this problem, the selection approaches (i.e. SCS, SES, DCS, and DES) have been recently received increasing attention to improve the classification performance. In these approaches, the diversity and accuracy are important criteria for selection of classifiers [20,26]. For designing the classifier ensemble selection, it is also important to obtain a trade-off between error and diversity of the classifiers. Hence, evolutionary multi-objective

\* Corresponding author. Tel.: +98 3433235901.

E-mail addresses: [r.mousavi@student.kgut.ac.ir](mailto:r.mousavi@student.kgut.ac.ir) (R. Mousavi), [m.eftekhari@uk.ac.ir](mailto:m.eftekhari@uk.ac.ir) (M. Eftekhari).

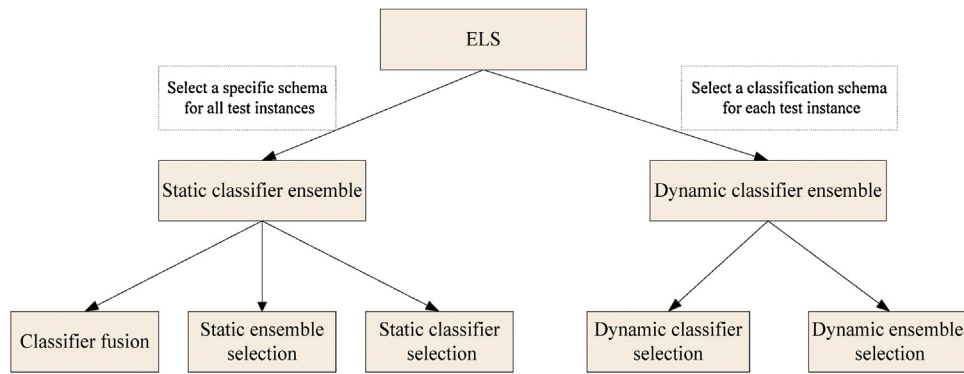


Fig. 1. ELSs schema for the combination of classifiers.

algorithms are deserved for finding the optimal trade-off between these objectives [27]. So far, a number of multi-objective ensemble learning approaches have been proposed that are briefly presented in Section 3. However, after careful analysis we can find that none of these approaches have taken into consideration the effect of the NSGA-II algorithm along with a large set of structurally different classifiers (i.e. heterogeneous classifiers), that a subset of them together with a combiner is encoded in a chromosome.

In the present study, the authors propose a new hybrid approach for ELSs that utilizes the benefits of SES and DES strategies. The approach consists of two main stages that in the first stage, NSGA-II is employed for optimizing SES learning system (named as SES-NSGAI). The aim in the first step is to select the optimal set of classifiers and their combiner for all test samples, by simultaneous optimization of error and diversity objectives. In the second stage, DES-Performance (DES-P) [20] is improved using SES-NSGAI, named as Improved DES-P (IDES-P). IDES-P selects an appropriate classifier subset for a test sample based a competence measure. The key problem for IDES-P is how to classify the test sample if no classifier is selected. To tackle this problem, we utilize the proposed technique in DES-Knora-Eliminate (DES-KE) [19], with the aim of increasing generalization ability of IDES-P. Moreover, the research tries to answer the following questions:

1. Is there one suitable combiner for all datasets, or there should be a special type of combiner for each dataset?
2. What kinds of classifiers are needed for each dataset?
3. What is the relation between the diversity of classifiers on training data and the over-fitting problem?

The major contributions of this paper are listed as follows. First, inspired by the idea of SES, a new method (SES-NSGAI) is developed using NSGA-II. Besides, according to the results achieved by the developed method, we discuss the relationship between diversity and over-fitting problem from the view point of Pareto optimal front. Second, based on SES-NSGAI, DES-KE and DES-P, a new hybrid method is developed to classify a test sample by the selected classifiers and their combiner. Third, we use the heterogeneous ensemble and different combiners to design the ensemble. Also, the two proposed methods (i.e. SES-NSGAI and IDES-P) are compared with other ensemble approaches over 14 UCI Machine Learning Repository [28] and Ludmila Kuncheva Collection (LKC) [29] datasets. The obtained results show the superiority of our proposed methods in comparison with other ones. Finally, we answer the above research questions on the basis of these experimental results.

The rest of this paper is organized as follows: in Section 2, the related works are briefly introduced on the ELSs. Section 3 gives the background knowledge about multi-objective ensemble learning

and describes NSGA-II algorithm. In Section 4, the methods of creating diversity in ensemble classifiers as well as diversity measures are explained, respectively. Section 5 describes the proposed hybrid ELS in detail. The experimental setup, experimental results, and discussion are presented in Section 6. Finally, Section 7 includes the concluding remarks.

## 2. Related works

According to the recent studies, ELSs have been received more and more attentions by researchers from the view point of the fusion strategies with the aim of improving the classification performance. For the combination of classifiers, there are five different types of strategies: CF, SCS, SES, DCS, and DES (as illustrated in Fig. 1). In the following these strategies are briefly described.

### 2.1. Static classifier ensemble (SCE)

There are three different types of SCE schemas as follows.

#### 2.1.1. Classifier fusion (CF)

The CF employs all classifiers in the ensemble for such decision makings as Majority Voting (MV), Weighted Majority Voting, Negative Correlation Learning, Bagging, Boosting, AdaBoost, and Random Forest. A complete description of these approaches can be found in [12].

#### 2.1.2. Static classifier selection (SCS)

This method first selects the best classifier for each region of competence in the feature space during training phase; next, each test sample is classified with a classifier related to its region. In this field, Kuncheva [13] proposed an algorithm for classifier combination based on clustering and selection technique. Also, the Single Best (SB) approach can be considered as one of the SCS methods since it selects the best classifier in the ensemble with the highest training accuracy for all test samples [20].

#### 2.1.3. Static ensemble selection (SES)

The SES selects an optimal classifier ensemble for all test samples. Yang [15] proposed Selection based on Accuracy (SA) and Selection based on Accuracy and Diversity (SAD) approaches in this area. The SA algorithm selects 75% of more accurate classifiers of an ensemble based on training error for all test instances, while by the SAD method, 90% of the most accurate classifiers in the ensemble are selected according to the classification accuracy on the training set and 15% of the remaining classifiers of an ensemble with the lowest diversity are eliminated. Therefore, 75% of the most accurate and diverse classifiers are totally selected; the SAD also calculates the diversity among trained classifiers using Q-statistic method. In

Download English Version:

<https://daneshyari.com/en/article/6905012>

Download Persian Version:

<https://daneshyari.com/article/6905012>

[Daneshyari.com](https://daneshyari.com)