Pattern Recognition 72 (2017) 44–58

Contents lists available at ScienceDirect

Pattern Recognition

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

# Online pruning of base classifiers for Dynamic Ensemble Selection

Dayvid V.R. Oliveira[a,*], George D.C. Cavalcanti[a], Robert Sabourin[b]

[a] Centro de Informática, Universidade Federal de Pernambuco, Recife, PE, Brazil
[b] LIVIA, École de Technologie Supérieure, University of Quebec, Montreal, Quebec, Canada

ARTICLE INFO

*Article history:*
Received 4 December 2016
Revised 1 May 2017
Accepted 25 June 2017
Available online 28 June 2017

*Keywords:*
Ensemble of classifiers
Dynamic ensemble selection
Classifier competence

ABSTRACT

Dynamic Ensemble Selection (DES) techniques aim to select only the most competent classifiers for the classification of each test sample. The key issue in DES is how to estimate the competence of classifiers for the classification of each new test sample. Most DES techniques estimate the competence of classifiers using a given criterion over the set of nearest neighbors of the test sample in the validation set, these nearest neighbors compose the region of competence. However, using local accuracy criteria alone on the region of competence is not sufficient to accurately estimate the competence of classifiers for the classification of all test samples. When the test sample is located in a region with borderline samples of different classes (indecision region), DES techniques can select classifiers with decision boundaries that do not cross the region of competence, assigning all samples in the region of competence to the same class. In this paper, we propose a dynamic selection framework for two-class problems that detects if a test sample is located in an indecision region and, if so, prunes the pool of classifiers, pre-selecting classifiers with decision boundaries crossing the region of competence of the test sample (if such classifiers exist). After that, the proposed framework uses a DES technique to select the most competent classifiers from the set of pre-selected classifiers. Experiments are conducted using the proposed framework with 9 different dynamic selection approaches on 40 classification datasets. Experimental results show that for all DES techniques used in the framework, the proposed framework outperforms DES in classification accuracy, demonstrating that our proposal significantly improves the classification performance of DES techniques, achieving statistically equivalent classification performance to the current state-of-the-art DES frameworks.



## 1. Introduction

Multiple Classifier Systems (MCS) [1] combine classifiers expecting that several classifiers outperform any single base classifier in classification accuracy [2], [3]. Several studies present MCS as an alternative to increase classification accuracy in many pattern recognition tasks, such as image labeling [4], handwritten recognition [5], signature verification [6], recommendation systems [7], banking [8], and face recognition [9].

MCS has three general phases [10]: (1) generation, in which the training set is used to generate a pool of classifiers; (2) selection, in which a subset of the pool of classifiers is selected to perform the classification, we refer to this subset of classifiers as ensemble of classifiers; (3) combination (or integration), in which the final decision is made based on the predictions of the classifiers.

The selection phase can be either static or dynamic. In static selection, the selection of classifiers is performed in the training phase. In dynamic selection, the selection of classifiers is performed for each new test sample in the classification phase. Recent works have shown that dynamic selection techniques achieve higher classification accuracy than static selection techniques, especially on ill-defined problems [10], [11], [12]. Dynamic selection can be either Dynamic Classifier Selection (DCS) [13] or Dynamic Ensemble Selection (DES) [11]. DCS selects one single classifier for the classification of each new test sample, and DES selects one or more classifiers for the classification of each new test sample. Since DCS is a specific case of DES, in this paper, we refer both as DES.

DES techniques rely on the assumption that different classifiers are competent ("experts") in different local regions of the feature space. For this reason, given a test sample $x_{query}$ and a pool of classifiers $C$, dynamic selection techniques try to select the most competent classifier $c$, or an ensemble of competent classifiers $C'$, $C' \in C$, for the classification of $x_{query}$. The key issue in DES is how to estimate the competence of a base classifier for the classification of a new test sample. Many DES techniques [11], [14], [15], [16],

* Corresponding author.
*E-mail addresses:* dvro@cin.ufpe.br (D.V.R. Oliveira), gdcc@cin.ufpe.br (G.D.C. Cavalcanti), robert.sabourin@etsmtl.ca (R. Sabourin).

http://dx.doi.org/10.1016/j.patcog.2017.06.030
0031-3203/

[17] estimate the competence of a classifier $c$ for the classification of a test sample $x_{query}$ using the accuracy of the classifier $c$ on a set of labeled samples similar to $x_{query}$, obtained using the K-Nearest Neighbors (KNN) on the validation set $\mathcal{D}_{SEL}$. The set of $K$ nearest neighbors of $x_{query}$ in $\mathcal{D}_{SEL}$ is called the region of competence ($\Psi$) of $x_{query}$.

According to Britto et al. [10], most DES techniques use some criteria on the region of competence of the test samples to estimate the competence of base classifiers. META-DES [18], a recent DES framework published after the survey [10], achieved the highest DES classification performance to this date, and it also uses the region of competence to extract meta-features that are used to predict the competence of base classifiers. So, a crucial issue in the design of DES techniques is the definition of the region of competence. We expect that the better the region of competence, the higher the precision of DES systems.

State-of-the-art DES techniques do not take into account the existence of different scenarios when estimating the competence of a classifier for the classification of a test sample using its region of competence. Given a test sample and a classifier, the test sample can be located in a region where almost all samples belong to the same class (safe region), or in a region where samples belong to more than one class (indecision region), and the classifier can have its decision boundary crossing or not crossing the region of competence of the test sample. Based on that, Fig. 1 presents 4 scenarios, given a test sample $x_{query}$, a region of competence $\Psi$, and a classifier $c$: (I) $x_{query}$ located in a safe region, and $c$'s decision boundary crossing $\Psi$. (II) $x_{query}$ located in a safe region, and $c$'s decision boundary not crossing $\Psi$. (III) $x_{query}$ located in an indecision region, and $c$'s decision boundary crossing $\Psi$. (IV) $x_{query}$ located in an indecision region, and $c$'s decision boundary not crossing $\Psi$. Where ▲ is the test sample ($x_{query}$), the dotted circle delimits the region of competence, the markers "○" and "■" are samples from different classes, the continuous straight line is the decision boundary of the classifier $c$.

Scenarios I and II (Fig. 1(a) and 1(b)) show a test sample ▲ located in a safe region (all samples in the region of competence of the test sample are from the class "○"). In Scenario I, the decision boundary of the classifier $c$ crosses the region of competence, and $c$ correctly classifies 20% of the samples in the region of competence. In Scenario II, the decision boundary of the classifier $c$ does not cross the region of competence, and $c$ correctly classifies 100% of the samples in the region of competence. A scenario with the classifier $c$ not crossing the region of competence of the test sample ▲ and $c$ misclassifying all samples in the region of competence of the test sample was not detailed because any accuracy based DES technique estimates the competence of such classifier as 0.0, and therefore, such classifier is never selected. This shows that, in safe regions, classifiers with decision boundaries not crossing the region of competence have higher competence estimation, meaning accuracy based DES techniques are sufficient to estimate the competence of base classifiers when the test sample is located in a safe region.

Scenarios III and IV (Figs. 1(c) and 1(d)) show a test sample ▲ located in an indecision region (samples from different classes "○" and "■" in the region of competence of the test sample). In Scenario III, the decision boundary of the classifier $c$ crosses the region of competence of the test sample, and $c$ correctly classifies 60% of the samples in the region of competence. In Scenario IV, the decision boundary of the classifier $c$ does not cross the region of competence of the test sample, and $c$ correctly classifies 60% of the samples in the region of competence. Since the classifiers from Scenarios III and IV have the same classification accuracy (60%), accuracy based competence estimation schemes gives the same competence estimative to them, even though the classifier $c$ from Scenario III is more competent because it correctly classifies samples of the different classes in the region of competence, and the classifier $c$ from Scenario IV classifies all samples in the region of competence as being from the same class ("○").

The scenarios from Fig. 1 show that depending on the type of region in which the test sample $x_{query}$ is located and whether the decision boundary of the classifier $c$ crosses the region of competence or not, the criteria used by DES techniques are not enough to decide if $c$ is locally competent for the classification of $x_{query}$. Our hypothesis is that, when the test sample is located in an indecision region, performing the selection of classifiers from a subset of the pool containing only classifiers with decision boundaries that cross the region of competence of the test sample (if such classifiers exist) is a promising dynamic selection approach.

In this paper, we propose a dynamic ensemble selection framework for two-class classification problems. The framework is divided into three phases: (1) Overproduction, where the framework generates the pool of classifiers; (2) Region of Competence Definition, where the framework defines the region of competence of each new test sample; (3) Selection, where the framework selects locally competent classifiers for the classification of each new test sample. The Selection phase is divided in three main steps: Indecision Region Detection, Dynamic Pruning, and Dynamic Selection. The Indecision Region Detection step decides if the test sample is located in an indecision region. The Dynamic Pruning step pre-selects locally competent classifiers by removing (or "pruning") classifiers with decision boundaries that do not cross the region of competence of the test sample when the test sample is located in an indecision region, if no classifier has decision boundaries crossing the region of competence, the Dynamic Pruning step pre-selects all classifiers. The Dynamic Selection step selects the most competent classifiers from the set of pre-selected classifiers for the classification of the test sample using any DES technique.

In the experiments, we evaluated the proposed framework with 9 dynamic selection schemes from the literature using 40 datasets from KEEL [19], and compared the framework with 3 state-of-the-art DES approaches, named: Randomized Reference Classifier (RRC) [20], META-DES [18], and META-DES.Oracle [21]. The results showed that the proposed framework outperforms DES, for all DES techniques used in our experiments, demonstrating that the framework significantly improves the classification accuracy of Multiple Classifier Systems. Also, using simple DES techniques in the dynamic selection step, the proposed framework was able to achieve statistically equivalent performance to the current state-of-the-art DES frameworks from the literature. These results were confirmed by Wilcoxon Signed Rank Test [22], Sign Test [23], Friedman test [24], and Nemenyi post hoc test [25].

This paper is organized as follows: Section 2 presents the problem statement. Section 3 presents the proposed framework. Section 4 presents the experimental study. Finally, Section 5 presents the conclusion.

## 2. Problem statement

In this section, we define indecision regions (Section 2.1), and show that DCS (Section 2.2) and DES (Section 2.3) techniques have problems evaluating the competence of classifiers and selecting competent classifiers for the classification of test samples located in indecision regions.

### 2.1. Indecision regions

According to García et al. [26], there are three types of test samples (Fig. 2): safe samples, borderline samples, and noisy samples. Safe samples (labeled as S) are located in a neighborhood of samples with relatively homogeneous class labels. Borderline samples (labeled as B) are located in areas surrounding classes boundaries,