



The Dynamic Selection of Combination Methods in Classifier Ensembles by Region of Competence

Jesáias Carvalho Pereira Silva¹(✉), Anne Magaly de Paula Canuto¹,
and Araken de Medeiros Santos²

¹ Federal University of Rio Grande do Norte, Natal, RN 59.078-970, Brazil
jesayassilva@gmail.com, anne@dimap.ufrn.br

² Federal University of Rural do Semi-Árido, Angicos, RN 59.515-000, Brazil
araken@ufersa.edu.br

Abstract. Over the years, research has advanced in the field of classifier ensembles. Several ways to improve its efficiency have emerged, for homogeneous and heterogeneous ensembles. One challenge when using classifier ensembles is the definition of its structure. Basically, the ensemble structure selection can be done in two different ways, static and dynamic selection. Different static selection, dynamic selection defines the ensemble structure is selected for each testing instance (dynamic selection). Dynamic selection methods have been proposed in the literature, mainly for ensemble members and features, but very little effort has been done to propose dynamic selection methods for combination methods. In this paper, a dynamic combination selection is proposed in which the combination method is selected to each testing instance. The main aim of the proposed dynamic combination selection is to adapt the ensemble structure to the characteristics of each testing instance. In order to assess the feasibility of the proposed method, an empirical analysis is conducted. In this analysis, the proposed method is used along with a dynamic ensemble member selection (KNORA-Eliminate and META-DES) in order to promote more dynamism in the ensemble structure. In this analysis, the proposed methods are compared to classifier ensemble with static combination methods and it improved the performance of all analyzed methods, for almost all analyzed scenarios.

Keywords: Classifier ensembles · Dynamic structure selection · Combination methods

1 Introduction

A classifier ensemble can be defined as a collection of individual classifiers (ensemble members), working in a parallel way, which receives the same pattern input and produces its output. A combination method receives the members outputs and provides the global output of the system [13]. In machine learning, classifier ensembles have been emerged as an efficient technique in different classification problems. In the literature, several studies

have been investigated efficient ways to combine classifiers aiming at improving the classification performance in different applications. In these cases, we can find different algorithms, methods and/or practices for combining classifiers [1–5, 7, 8].

In this context, one important aspect is the definition of the ensemble structure. Several studies have proposed different ways to define the ensemble structure such as: Optimization techniques, meta-learning, among others [21, 22]. Basically, the ensemble structure definition can be done statically or dynamically. In the static selection, the ensemble structure is selected in the beginning of the training phase, and it is used throughout the whole ensemble processing. On the other hand, the dynamic selection defines the ensemble structure dynamically, in which each testing instance has its own ensemble structure. Several studies have shown that the dynamic selection tends to increase the predictive capacity of an ensemble [1, 9].

There are several studies that investigate the dynamic selection of ensemble structure, mainly for ensemble members [7, 8] and features [23] and both of them [24]. As mentioned previously, in an ensemble structure, the combination method aims at combining the outputs of all classifiers in order to provide the final output of an ensemble. Although this component plays an important role in the performance of an ensemble, very little has been done in order to define an efficient dynamic selection of this module.

Aiming at proposing an automatic decision process to select the best classification structure to a testing instance, this paper proposes a dynamic selection method for combination methods in classifier ensembles. Also known as dynamic fusion, the proposed method defines the region of competence of each candidate (combination method) for each testing instance and the most competent combination method is selected. The definition of the region of competence is made based on a pool of classifiers and it can be statically or dynamically formed. In this paper, we will apply the dynamic selection since we aim to promote dynamicity in two important parameters of an ensemble (ensemble members and combination method).

In order to assess the feasibility of the proposed method, an empirical analysis will be conducted. In this analysis, the proposed method will be used along with two well-known ensemble members dynamic selection methods, KNORA-Eliminate (KNORA-E) and META-DES. In the proposed method, a set of eight combination methods are used as candidates to be selected. Additionally, an analysis of the selection distribution of the possible candidates will be performed in order to investigate whether this selection is distributed over all possible candidates or there are one or two candidates that dominates the selection process. Finally, the proposed method will be compared to 12 ensembles in which the combination methods are selected in a static way. All the analyzed methods will be evaluated using 15 classification datasets.

2 Theoretical Concepts and Related Work

2.1 State of the Art

There are several studies that investigate the dynamic selection of ensemble structure, mainly for ensemble members [7, 8] and features [23] and both of them [10, 24].

In relation to ensemble members, in [7], for instance, a new method for dynamic ensemble member selection is presented and it uses the confidence of the base classifiers during the classification and its general credibility as selection criterion. Thus, an ensemble member is selected to compose the ensemble if its selection criterion is higher than an established threshold x . Another interesting way is to use region of competence as selection criterion, making it possible to improve the combination of classifiers, in which the most competent ones in a certain region are selected. The use of region of competence as selection criterion helps to maximize results [11, 12] by focusing only on the most competent classifiers, and examples can be found in KNORA-E [1] and META-DES [3].

In terms of dynamic feature selection, in [23], a dynamic feature select approach was proposed. The main aim of this approach is to select a different subset of features for one instance or a group of instances. The main goal of this approach is to explore the full potential of all instances in a classification problem. In [24], an initial study on how to combine these two dynamic selection techniques was performed. According to the authors, an improvement in performance was detected with the use of this integrated dynamic selection technique.

Although there are several studies to propose dynamic selection of ensemble members and feature selection, very little has been done in order to propose efficient dynamic selection of combination methods. This paper tries to bridge this gap and it proposes a dynamic selection method based on region of competence.

2.2 Classifier Ensembles

It is well-known that there is not a single classifier which can be considered optimal for all problem domains [13]. Therefore, it is difficult to select a good single classifier which provides the best performance in practical pattern classification tasks [14]. Therefore, classifier ensembles have emerged as an efficient classification structure since it combines the advantages and overcomes the limitations of the individual classifiers. Providing better generalization and performance ability, when compared to the individual classifiers [14]. In a classifier ensemble, an input pattern is presented to all individual classifiers [15, 16], and a combination method combines their outputs to produce the overall output of the system [13, 17]. The Machine Learning literature has ensured that diversity plays an important role in the design of ensembles, contributing to their accuracy and generalization [13].

One important issue regarding the design of classifier ensembles involves the appropriate selection of its structure (individual classifiers and combination methods) [18]. As previously mentioned, there are basically two main selection approaches, static and dynamic. In this paper, we will focus on the dynamic approach. The next subsection will describe some existing dynamic selection methods that will be used in this paper.

2.3 Dynamic Ensemble Member Selection

The Dynamic Ensemble Selection (DES) methods perform the dynamic ensemble member selection. These methods select a subset of classifiers to classify each test instance. The selection of the classifier subset is done through the use of a selection procedure

and each DES method has its own procedure. There are several DES methods proposed in the literature. In this paper, we will use two well-known DES methods, KNORA-E and META-DES.

KNORA-E: Knora [1] is a well-known DES method and it seeks to find the best subset of classifiers for a given test instance. It applies a k -Nearest Neighbors methods and the neighbors of a testing instance are selected from the validation set and the competence of each classifier is calculated. Based on a certain selection criterion, the classifier subset is selected.

KNORA-E is a knora-based method, and the selection criterion is to select a set of classifiers formed only by the classifiers that correctly classify all k neighbors of a testing instance. In the case where no classifier can correctly classify all k neighbors, the k value is decremented by one and this is done until at least one classifier can be selected [1].

META-DES: The META-DES [3] is a DES method that uses the idea of selection using meta-learning. In this method, a meta-problem is created to determine whether a classifier is competent for a given test instance. According to [10], the META-DES method uses five criteria for extracting meta-features in order to establish the new region of a meta-problem.

After that, a meta-classifier is trained, based on the defined meta-features. This meta-classifier is then used to identify whether a classifier is competent or not to classify a testing instance. Classifiers that are labeled as competent will be selected to compose the ensemble to classify the test instance.

3 The Proposed Method

The proposed method aims at selecting the combination method of a classifier ensemble dynamically. Algorithm 1 presents the main steps of the proposed method. As it can be observed, the dynamic selection of combination is performed in the testing phase. In this phase, when a testing instance is presented to the classifier ensemble, the competence of each combination method is calculated.

Algorithm 1: The proposed method.

```

01: Procedure Dynamic Fusion
02:   Input: Testing instance ( $T_i$ ), Validation set ( $V$ ),
      pool of classifiers ( $M$ ), combination methods ( $C$ )
03:   Output: The selected combination method ( $C_m$ )
04:    $N = k\text{-NN}(T_i, V)$  % Find the neighbors of  $T_i$ 
05:   FOR  $j=0$  until  $j= \text{size}(C)$  DO
06:      $\text{Acc}(C_j) = \text{Accuracy}(C_j, N)$ 
07:   END FOR
08:    $C_m = \text{Max}(C_j, j=1, 2, \dots, \text{size}(C))$ 
09:   Return  $C_m$ 
10: END procedure

```

This competence can be calculated in several different ways. As the use of local competence has been widely used in dynamic member selection methods, in the proposed method, the k-NN method is used to select the neighbors of the testing instance. Then, the local accuracy of each combination method is calculated in this neighborhood, and the most accurate combination method is selected.

When computing the competence region, a draw in accuracy may occur. In this case, the number of neighbours is increased by 1 ($k = k + 1$) until a winner is detected. If all instances of the Validation set is used and there is still a draw, the winner method is randomly selected.

Additionally, there are two main parameters in the proposed method, the number of neighbors of a testing instance (line 4 of Algorithm 1) and the size of the pool of classifiers. Finally, the proposed method can be applied to a pool of classifiers selected statically or dynamically. In order to provide more dynamicity for the classifier ensemble, in this paper, the proposed method will be applied to two well-known dynamic member selection methods, META-DES and KNORA-E (described in Sect. 2.3).

4 Experimental Methodology

In order to assess the feasibility of the proposed method, an empirical analysis will be conducted. The next subsections will describe the main aspects of this analysis, mainly the used datasets as well as its methods and materials.

4.1 Datasets

The datasets used in this paper are extracted from the UCI Machine Learning Repository. Table 1 describes some characteristics of the used datasets, focusing in the number of instances (Inst), number of attributes (Att) and number of classes (Class) of each dataset.

Table 1. Description of the used datasets

Dataset	Name	Inst	Att	Class
D1	Cardiac insufficiency	368	53	2
D2	Car	1728	6	4
D3	Seismic-bumps	2584	18	2
D4	Zoo	101	16	7
D5	Ionosphere	351	34	2
D6	Prognostic	198	33	2
D7	Wine	178	13	3
D8	Dermatology	366	34	6

(continued)

Table 1. (continued)

Dataset	Name	Inst	Att	Class
D9	Heart	303	13	2
D10	Bone marrow	187	36	2
D11	Algerian Forest Fires	244	13	2
D12	Congres Voting Records	435	16	2
D13	Maternal Health Risk	1014	6	3
D14	Risk Factors Cervical Cancer	855	28	2
D15	Phishing Website	2456	30	2

Each dataset is divided into training, Validation1, Validation2 and Testing sets, in a proportion of 50%, 16.7%, 16.7%, and 16.6%, respectively. The training set is used to generate the pool of classifiers. The testing set is used to assess the performance of the classifier ensembles. The Validation2 set is used to train the trainable combination methods (Neural Networks and Naive Bayes) while the Validation1 set is used to obtain the competence region of the proposed dynamic fusion method. This division is performed 30 times and the presented results of each ensemble configuration represent the average values over these 30 values.

4.2 Methods and Materials

In this paper, we evaluated 6 different pool sizes, 5, 10, 15, 20, 25, and 30 classifiers, in which all of them are, generated through the Bagging method. In addition, 3 different number of neighbors are assessed, 3, 7, and 11 neighbors. It is important to emphasize that the proposed method as well as both member selection techniques (KNORA-E and META-DES) use the idea of competence region. In this sense, the same number of neighbors are used for both cases, the selection of the combination method (proposed method) and the ensemble members (KNORA-E and META-DES). Finally, all ensemble configurations use Decision Trees as ensemble members.

The proposed method used a pool of 8 different combination methods, which are: Majority Vote, Sum, Geometric Mean, Naive Bayes, Edge and three Neural Networks (MLP) versions, Hard, Soft and Soft-Class. The three NN versions differ on the input information received by the ensemble members. In the Hard version, the ensemble member provides only the winner class for the testing instance. In other words, this MLP version is trained and tested using only the winner class of each ensemble member. In the other two MLP versions, the prediction probability for each class is used. In this sense, the prediction probability for each class is provided, for both MLP versions.

As the MLP input must have a fixed size and the number of selected members might vary, a strategy to define the input size must be done. In this paper, we decided to use the maximum possible size for a combination method (pool of classifier times the number of classes). In doing this, it is important to define how to handle the outputs of the unselected classifiers. The way to handle the outputs of the unselected classifiers is the

main difference between Soft and Soft-Class versions. While the Soft MLP version uses -1 to all classes of an unselected classifier, the Soft-Class version uses a fixed value ($1/\text{number of total classes}$) as the value to all classes of the unselected classifiers.

The MLP algorithms were implemented using the Scikit-Learn [6] library with the Multi-layer Perceptron model, using 200 neurons in the hidden layer. All three neural networks followed the same configurations since this configuration provided promising results for all three NNs in a grid search method. For comparison purposes, the performance of the proposed method is compared to 12 classifier ensembles using the following combination methods: Majority Vote, Sum (Sum), Maximum (MAX), Minimum (MIN), Geometric mean, Hard MLP, Soft MLP, Soft-Class MLP, Edge, Naive Bayes, Weighted Sum and Weighted Ensemble Voting. For all analyzed ensembles, the ensemble members are dynamically selected, and the combination method is statistically selected, as originally proposed in both analyzed methods (KNORA-E and META-DES).

Additionally, the Weighted sum and weighted vote methods use weights on their functioning. The used weight is $1/(\text{distance-of-classes})$ and it is applied to the vote procedure in the Weighted Sum as well as the probability of the classifiers in the Weighted sum method.

The obtained results of all analyzed methods will be evaluated using the Friedman statistical test [20]. In cases where a statistically significant difference is detected, the Nemenyi post-hoc test is applied [20]. In order to present the obtained results by the post-hoc test, the critical difference diagram (CD) [20] is used. This diagram was selected in order to have a visual illustration of the statistical test, making it easier to interpret the obtained results. Additionally, all implemented methods are included in the DESLIB [3] library that contains both methods, KNORA-E and META-DES.

5 The Obtained Results

In this section, the obtained results are presented and analyzed. This analysis will be done in three main parts. In the first part, the accuracy of all 13 analyzed methods are assessed. The second part presents the distribution of the selected combination methods while the third part describes the results of the statistical analysis.

5.1 The Accuracy of the Analyzed Methods

Tables 2 and 3 present the accuracy results of all 13 analyzed methods for KNORA-E and META-DES, respectively. As previously mentioned, 18 different ensemble configurations (6 pool sizes and 3 different number of neighbors) and each configuration was performed 30 times. Therefore, values in Tables 2 and 3 represent the average over 540 results. Additionally, the last line in both tables represents the overall accuracy over all 15 datasets. Finally, the bold numbers represent the highest accuracy for each dataset.

For KNORA-E (Table 2), it can be observed that the proposed method (Dynamic Fusion) delivered the highest accuracy in 8 datasets, out of 15, followed by Vote and Sum (6 datasets). Furthermore, the overall accuracy achieved by the proposed method is the highest one of all analyzed methods. For META-DES (Table 3), the proposed method did not deliver the highest accuracy levels in many datasets (4 out of 15). Nevertheless, it presents the highest overall accuracy, followed by Vote and Sum. It shows that the proposed method is the best overall classifier ensemble.

The results obtained in Tables 2 and 3 show that the use of the dynamic selection of the combination methods proved to be efficient since it improves the performance of the classifier ensembles, when compared to the static selection methods. We believe that this improvement in performance is due to the fact that the dynamic fusion technique maximizes, even more, the characteristics of the ensemble members and, as a consequence, to improve the performance of the classifier ensemble. The original KNORA-E and META-DES techniques themselves already provide more efficient classifiers since the best ones are selected to classify a testing instance. Furthermore, the proposed method handles even more efficiently the classifier ensembles since it selects the combination methods that suits better to the ensemble members. Finally, we can state that the inclusion of more dynamicity in the ensemble structure can lead to an improvement in its performance.

Table 2. Results of the classifier ensembles using the KNORA-E method

Dataset	Majority Vote	SUM	MAX	MIN	Geometric average	Weighted Sum	Weighted Vote	MLP-HARD	MLP-SOFT	MLP-SOFT CLASS	Edge	Naive Bayes	Dynamic Fusion
D1	94.58	94.58	95.17	95.17	95.17	88.77	88.77	95.05	94.17	95.02	9458	9293	95.66
D2	97.30	97.30	9335	9181	91.81	84.90	84.90	94.29	96.12	97.15	9639	9385	97.93
D3	92.07	92.06	9306	9306	93.06	91.63	91.64	92.39	91.70	92.11	9207	8543	92.57
D4	93.63	93.63	9310	8997	89.97	76.47	76.47	90.59	88.37	92.45	9082	7621	93.73
D5	90.97	90.97	7984	7984	79.84	77.02	77.02	88.21	89.20	88.96	9097	8431	89.91
D6	69.48	69.48	4475	4475	44.75	62.09	62.09	69.53	68.86	69.28	6948	7180	72.62
D7	93.68	93.68	7835	7289	72.89	64.27	64.27	85.34	91.25	92.99	9261	5716	93.10
D8	95.84	95.84	9094	8475	84.75	78.84	78.84	90.79	93.79	95.82	9403	7801	95.74
D9	76.90	76.90	6344	6344	63.44	67.71	67.71	74.41	73.68	72.61	7690	7301	76.04
D10	91.70	91.70	8839	8839	88.39	90.97	90.97	90.13	88.48	89.78	9170	8918	91.95
D11	97.65	97.65	9613	9613	96.13	97.28	97.28	96.57	96.72	96.90	9765	9688	97.63
D12	94.45	94.45	9170	9170	91.70	92.46	92.46	93.11	93.51	93.36	9445	9280	93.94
D13	77.48	77.45	7812	7798	77.98	64.80	64.80	75.61	74.56	75.73	7762	7340	78.69
D14	91.47	91.47	9321	9321	93.21	90.97	91.00	92.57	92.17	92.36	9147	9105	93.37
D15	95.48	95.48	9248	9248	92.48	92.50	92.50	94.85	94.92	94.77	9548	9099	95.49
Acc Ave	90.18	90.18	8480	8370	83.70	81.38	81.38	88.23	88.50	89.29	8975	8313	90.56

5.2 The Selection Distribution

Once we have analyzed the accuracy of the different classifier ensembles, now we will evaluate the selection distribution made by the proposed methods. In other words, what is the proportion of selection for each combination method which was made by the proposed method in the testing phase.

Table 3. Results of the classifier ensembles using the META-DES method

Datasets	Majority Vote	SUM	MAX	MIN	Geometric average	Weighted Sum	Weighted Voting	MLP-HARD	MLP-SOFT	MLP-SOFT-CLASS	Edge	Naive Bayes	Dynamic Fusion
D1	94.23	94.23	94.52	94.52	94.52	90.43	90.43	94.07	95.09	95.22	9423	9130	95.01
D2	96.80	96.80	89.89	87.12	87.12	83.86	83.86	93.72	96.77	96.99	9568	9108	96.85
D3	92.01	92.01	93.25	93.25	93.25	92.37	92.37	92.37	92.17	92.62	9201	8259	92.66
D4	94.54	94.54	89.54	85.26	85.26	75.92	75.92	90.88	95.13	95.26	8980	75.13	94.48
D5	90.46	90.46	80.22	80.22	80.22	76.99	76.99	88.31	89.42	88.96	9046	8357	89.67
D6	74.26	74.26	41.30	41.30	41.30	65.80	65.80	69.21	69.65	70.28	7426	5861	73.16
D7	92.84	92.84	74.37	67.87	67.87	64.39	64.39	84.39	91.97	92.45	9197	5895	92.36
D8	96.22	96.22	88.23	80.62	80.62	78.96	78.96	90.49	96.10	96.17	9385	7853	96.32
D9	78.54	78.54	63.02	63.02	63.02	68.59	68.59	74.18	75.70	76.46	7854	6694	78.85
D10	92.01	92.01	89.43	89.43	89.43	91.49	91.49	89.77	90.68	90.82	9201	8887	91.79
D11	97.75	97.75	95.86	95.86	95.86	97.42	97.42	96.58	96.82	96.78	9775	9654	97.71
D12	94.54	94.54	92.17	92.17	92.17	92.72	92.72	93.09	93.70	93.50	9454	9272	94.08
D13	77.54	77.62	77.10	76.29	76.28	64.49	64.51	75.55	76.67	77.32	7774	6805	78.93
D14	91.85	91.84	93.66	93.66	93.66	92.16	92.19	92.54	92.19	92.32	9185	8902	93.40
D15	95.50	95.51	92.17	92.17	92.17	92.63	92.62	94.92	95.18	95.12	9550	9036	95.40
Acc Avg	90.61	90.61	83.65	82.18	82.18	81.88	81.88	88.01	89.82	90.02	9001	80.82	90.64

Tables 4 and 5 present the selection distribution for KNORA-E and META-DES methods, respectively. As it can be observed in both tables, all eight combination methods were selected by the proposed method in the testing phase. The most selected combination method is Vote, for both methods (KNORA-E and META-DES). It is an expected result since this method provided the second best accuracy level. On the other hand, the Naïve Bayes was rarely selected as the best combination method. Once again, this combination method provided one of the worst accuracy levels, for both methods (KNORA-E and META-DES). It is important to emphasize that the Sum combination method was also rarely selected since its performance is usually very similar to the Vote method and, in these datasets, the latter method was slightly better, and it was then selected. Finally, although the edge combination method obtained good accuracy levels, it is possible to observe that it was rarely selected by the proposed method.

Still in the analysis of the selection distribution, it can be seen that there is a certain equal selection distribution among five combination methods (Vote, Geometric mean, MLP-Hard, MLP-Soft and MLP-Soft-Class) which shows that there is no best combination method and that an efficient selection can improve even further the performance of the classifier ensemble. Among the NN versions, it can be observed similar selection distribution among themselves, showing a certain similarity among all three version, which together surpasses the majority vote proportion.

Based on the results of Tables 4 and 5, one can conclude that the proposed method generally biases towards the method with the highest accuracy level. This is an expected result since if a combination method is the most successful one, it is usually one of the most successful one in the dynamic fusion competence region.

5.3 Statistical Analysis

In order to evaluate the obtained results from a statistical point of view, the Friedman test [19] was applied to verify if there are statistical differences among all ensemble

Table 4. Selection distribution in the dynamic fusion when using KNORA-E

x	Sum	Majority Vote	Geometric average	MLP-HARD	MLP-SOFT	MLP – SOFT CLASS	Edge	Naive Bayes
Cardiac insufficiency	0.00	27.13	23.05	21.26	9.50	13.20	0.00	5.87
Car	0.00	30.81	1.82	12.22	11.61	26.53	7.06	9.95
Seismic-bumps	0.03	11.27	39.29	15.59	10.63	11.31	0.00	11.88
Zoo	0.69	79.94	0.80	4.57	7.43	6.20	0.19	0.19
Ionosphere	0.00	35.94	3.07	20.54	19.24	12.45	0.00	8.75
Prognostic	0.00	19.69	5.58	18.05	18.83	11.95	0.00	25.90
Wine	0.00	56.62	6.29	6.59	13.80	16.08	0.24	0.37
Dermatology	0.05	64.87	2.54	7.23	11.53	13.45	0.29	0.05
Heart	0.00	22.50	11.58	19.93	19.59	12.94	0.00	13.45
Bone marrow	0.00	46.83	18.20	15.27	6.40	6.35	0.00	6.95
Algerian Forest Fires	0.46	58.11	18.64	9.62	4.49	3.92	0.49	4.26
Congressional Voting Records	0.00	31.45	14.60	22.99	13.02	7.52	0.00	10.42
Maternal Health Risk	1.31	6.98	28.23	22.25	11.46	15.76	5.99	8.00
Risk Factors Cervical Cancer	0.00	5.50	41.65	22.76	10.29	6.90	0.00	12.89
Phishing Website	2.64	27.28	10.39	23.30	16.98	17.23	0.00	2.18
OVERALL AVERAGE	0.35	34.99	15.05	16.15	12.32	12.12	0.95	8.07

classifiers. The Friedman test is used to be able to state the hypothesis that the k-related observations derive from the same population (similar performance) or not (superiority in performance). In this test, the significance level used was set to 0.05. Hence, if the p-value is less than the established value, the null hypothesis is rejected, with a confidence level greater than 95%.

Table 6 presents the results of the Friedman test. As it can be observed in both cases (KNORA-E and META-DES) that the statistical test detected statistical differences among all analyzed methods. In this sense, the post-hoc test was applied the results are presented in the Critical Difference Diagram [20].

In the Critical Difference Diagram, the performance of a method is statistically different from another method if the difference between their average rankings is higher

Table 5. Selection distribution in the dynamic fusion when using META-DES

x	Sum	Majority Vote	Geometric average	MLP-HARD	MLP-SOFT	MLP – SOFTCLASS	Edge	Naive Bayes
Cardiac insufficiency	0.00	24.71	26.66	19.82	13.38	12.42	0.00	3.01
Car	0.00	27.53	1.87	6.19	21.72	26.59	10.34	5.74
Seismic-bumps	0.21	14.42	43.71	14.85	4.64	8.92	0.00	13.26
Zoo	0.05	86.87	0.60	3.22	8.49	0.69	0.05	0.05
Ionosphere	0.00	38.44	4.97	23.05	16.68	11.93	0.00	4.93
Prognostic	0.00	31.72	6.07	17.69	12.37	13.71	0.00	18.45
Wine	0.00	50.36	6.39	8.83	20.03	13.21	1.06	0.12
Dermatology	0.00	66.19	1.74	5.83	22.73	1.15	2.36	0.00
Heart	0.00	27.24	16.11	19.53	14.87	14.39	0.00	7.87
Bone marrow	0.00	46.72	24.43	10.99	8.14	3.52	0.00	6.20
Algerian Forest Fires	0.85	60.28	16.87	9.58	2.72	1.85	1.65	6.20
Congressional Voting Records	0.00	33.12	19.66	22.12	12.64	4.71	0.00	7.75
Maternal Health Risk	5.68	10.18	24.26	16.91	14.66	17.74	3.67	6.91
Risk Factors Cervical Cancer	0.00	12.90	50.08	14.87	8.44	2.33	0.00	11.37
Phishing Website	5.13	28.92	10.77	14.84	18.49	17.32	0.00	4.52
OVERALL AVERAGE	0.79	37.31	16.95	13.89	13.33	10.03	1.28	6.42

Table 6. Friedman test for KNORA-E and META-DES

KNORA-E				META-DES			
Chi-Squared	df	P	Kendall's W	Chi-Squared	df	P	Kendall's W
1.161254	12	<0.05	0.358	960.454	12	<0.05	0.296

than the critical difference calculated by the Critical Difference Diagram (CD). In this case, when two methods are similar, there is a horizontal line linking these two methods.

Figure 1(a) shows the CD diagram for KNORA-E. As it can be observed in this figure, the superiority of the proposed method was detected by the statistical test. It shows that the improvement in performance was strong enough to be detected by the statistical test, when compared to all other analyzed methods.

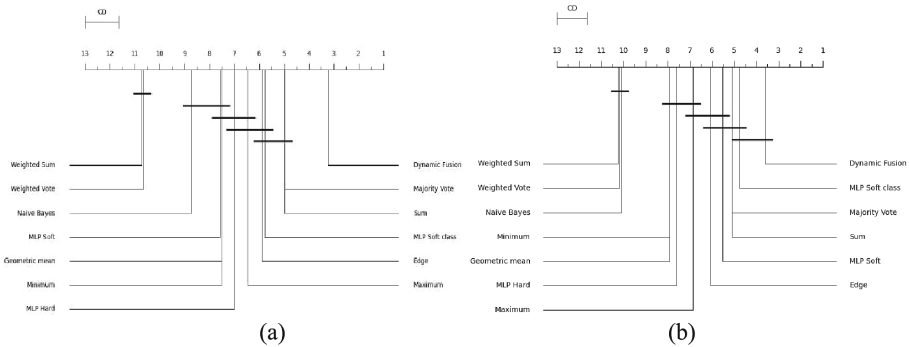


Fig. 1. Critical Difference Diagram for KNORA-E (a) and META-DES (b)

Figure 1(b) presents the CD diagram for META-DES. In this method, unlike KNORA-E, the superiority in performance delivered by the proposed method was not detected by the statistical test. As it can be observed, the CD diagram detected that the accuracy of dynamic fusion, sum, majority vote, and soft class MLP provide similar performance. For the remaining methods, the proposed method provided higher accuracy levels, detected by the statistical test.

6 Final Remarks

This paper proposed a method to dynamically select combination methods for a classifier ensemble. The combination methods are selected for each testing instance based on the competence region of the analyzed combination methods. The dynamic combination method selection was used along the dynamic ensemble member methods in order to include more dynamicity in the ensemble structure selection and, as a consequence, leading to more efficient ensembles.

In order to assess the feasibility of the proposed method, an empirical analysis was conducted. In this analysis, the dynamic combination method selection was made among eight combination methods: Majority Vote, Sum, Geometric Average, Edge, Naive Bayes, and three types of MLP. Additionally, the proposed method was used in two different dynamic ensemble member methods, KNORA-E and META-DES.

Through this analysis, it can be observed that the proposed method provided the highest overall accuracy levels among all analyzed methods, for both methods (KNORA-E and in META-DES). Additionally, we could observe that the proposed method selected almost equally five combination methods (out of 8), showing that it is indeed important to apply different combination methods in the classifier ensemble structures. Finally, the statistical test proved that the superiority in performance of the proposed method was detected by the statistical test, for KNORA-E.

In general, the proposed technique showed promising results and, indeed, improvements in performance, when compared to the static selection of combination methods. We believe that this improvement in performance is due to the fact that the dynamic fusion technique maximizes, even more, the characteristics of the ensemble members and, as a consequence, to improve the performance of the classifier ensemble.

This empirical was limited to 15 classification databases. As future analysis, it is necessary to expand this analysis using another selection approaches and to perform a more detailed analysis with more datasets and different ensemble configurations. It will also be investigated the best number of neighbors to define the dynamic fusion competence region and to expand the tests for KNORA-Union [1] and Overall Local Accuracy (OLA) [9], among others. Finally, different approaches to define the competence of a method will be analyzed.

References

1. Ko, A.H.R., Sabourin, R., Britto, A.S., Jr.: From dynamic classifier selection to dynamic ensemble selection. *Pattern Recogn.* **41**(5), 1718–1731 (2008)
2. Cavalin, P.R., Sabourin, R., Suen, C.Y.: Dynamic selection approaches for multiple classifier systems. *Neural Comput. Appl.* **22**, 673–688 (2013)
3. Cruz, R.M.O., et al.: META-DES: a dynamic ensemble selection framework using meta-learning. *Pattern Recogn.* **48**(5), 1925–1935 (2015)
4. Oliveira, D.V.R., Cavalcanti, G.D.C., Sabourin, R.: Online pruning of base classifiers for dynamic ensemble selection. *Pattern Recogn.* **72**, 44–58 (2017)
5. Cruz, R.M.O., Sabourin, R., Cavalcanti, G.D.C.: METADES. Oracle: meta-learning and feature selection for dynamic ensemble selection. *Inform. Fusion* **38**, 84–103 (2017)
6. Pedregosa, F., et al.: Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011)
7. Nguyen, T.T., et al.: Ensemble selection based on classifier prediction confidence. *Pattern Recogn.* **100**, 107104 (2020)
8. Cruz, R.M.O., et al.: Deslib: a dynamic ensemble selection library in python. *J. Mach. Learn. Res.* **21**(1), 283–287 (2020)
9. Woods, K., Kegelmeyer, W.P., Bowyer, K.: Combination of multiple classifiers using local accuracy estimates. *IEEE*. 405–410 p (1997)
10. Dantas, C.A.: An integration analysis of dynamic selection techniques for a classification system building. (PhD thesis) Federal University of Rio Grande do Norte (2021)
11. Cruz, R.M.O., et al.: Dynamic ensemble selection vs k-nn: why and when dynamic selection obtains higher classification performance? In: *IEEE 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pp. 1–6 (2017)
12. Cruz, R.M.O., Sabourin, R., Cavalcanti, G.D.C.: Dynamic classifier selection: Recent advances and perspectives. *Inform. Fusion* **41**, 195–216 (2018)
13. Kuncheva, L.I.: *Combining pattern classifiers: methods and algorithms*. Wiley (2004)
14. Dietterich, T.G.: *Ensemble methods in machine learning*. *Multiple Classifier Systems: First International Workshop, MCS 2000 Cagliari, Italy, June 21–23, 2000 Proceedings 1*. Springer, Berlin, Heidelberg (2000)
15. Kuncheva, L.I.: A theoretical study on six classifier fusion strategies. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(2), 281–286 (2002)
16. Brown, G., et al.: Diversity creation methods: a survey and categorisation. *Inform. fusion* **6**(1), 5–20 (2005)
17. Braga, L.P.V.: *Introdução à Mineração de Dados-2a edição: Edição ampliada e revisada*. Editora E-papers (2005)
18. Canuto, A.M.P.: Investigating the influence of the choice of the ensemble members in accuracy and diversity of selection-based and fusion-based methods for ensembles. *Pattern Recogn. Lett.* **28**(4), 472–486 (2007)

19. Friedman, M.: A comparison of alternative tests of significance for the problem of m rankings. *Ann. Math. Stat.* **11**(1), 86–92 (1940)
20. Demšar, J.: Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* **7**, 1–30 (2006)
21. Feitosa, A.A., Canuto, A.M.P.: An exploratory study of mono and multi-objective metaheuristics to ensemble of classifiers. *Appl. Intell.* **48**(2), 416–431 (2018)
22. Kordík, P., Cerný, J., Frýda, T.: Discovering predictive ensembles for transfer learning and meta-learning. *Mach. Learn.* **107**(1), 177–207 (2018)
23. Nunes, R.O., et al.: An unsupervised-based dynamic feature selection for classification tasks. In: 2016 International Joint Conference on Neural Networks (IJCNN). IEEE (2016)
24. Dantas, C., et al.: Instance hardness as a decision criterion on dynamic ensemble structure. In: 2019 8th Brazilian Conference on Intelligent Systems (BRACIS). IEEE (2019)