Experimental Validation of a Novel Fuzzy Classifier Performance

Cosmin Danut Bocaniala

Severin Bumbaru

Department of Computer Science and Engineering
University "Dunărea de Jos" of Galati, Domneasca 47, Galati 6200,
ROMANIA
Cosmin.Bocaniala@ugal.ro
Severin.Bumbaru@ugal.ro

Vasile Palade

Oxford University, Computing Laboratory,
Wolfson Building, Parks Road, Oxford, OX1 3QD, UK
Vasile.Palade@comlab.ox.ac.uk

Abstract: This paper validates a novel fuzzy classification methodology using the ELENA benchmark (ftp://ftp.dice.ucl.ac.be/pub/neural-net/ELENA/databases). The main advantage of this methodology is the high accuracy with which it learns the topological structure of the features space. The fuzzy subsets built by the classifier approximate with a very small error the areas in the features space corresponding to different categories. Its accuracy also manifests through handling with fine precision the discrimination inside overlapping areas. These two properties of the fuzzy classifier have been confirmed by the testing results obtained on ELENA benchmark datasets.

Keywords: Pattern recognition, Fuzzy logic, Fuzzy subsets, Clustering, Performance analysis.

Cosmin Danut Bocaniala is working as a Lecturer with the Department of Computer Science and Engineering, University of Galati, Romania, currently on leave for a research fellowship with the Instituto Superior Técnico, Lisbon, Portugal. He obtained his PhD at the University of Galati, in Jan. 2005, with a thesis on applications of Artificial Intelligence techniques to Fault Diagnosis. His current research interests include Distributed Causal Complex Systems, Computational Intelligence, Fault Diagnosis, Fault-Tolerant Control.

Vasile Palade is currently working as a Lecturer with the Computing Laboratory, Oxford University since 2001. He has previously worked with the University of Hull, UK, and University of Galati, Romania. His research interests include Hybrid Intelligent Systems, Neural and Neuro-fuzzy networks, Rule Extraction, Genetic Algorithms with applications mainly to Fault Diagnosis and Bioinformatics.

Severin Bumbaru is currently Advisor Professor with the Department of Computer Science and Engineering, University of Galati, Romania. He has been working in this Department as Professor and Head of Department until 2004. Professor Bumbaru is PhD supervisor in Computer Science field of University of Galati. His main research area is Artificial Intelligence, contributing with 12 books and 136 articles.

1. Introduction

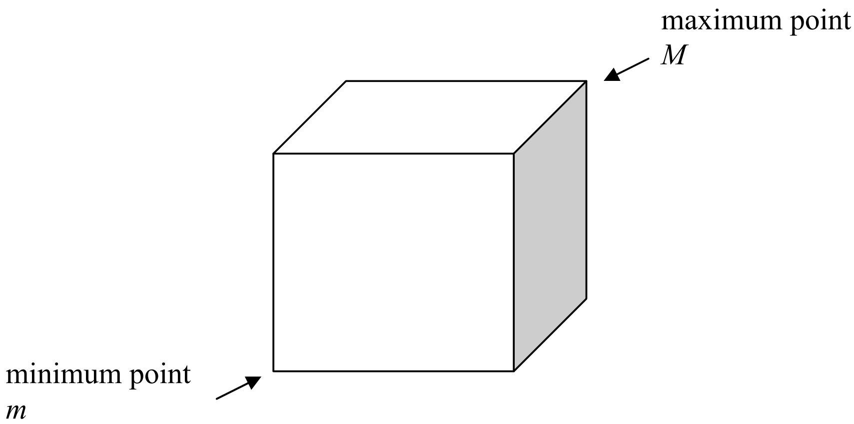
According to Duda *et. al.* [11], classification represents "the assignment of a physical object or event to one of several prespecified categories". There are three main classes of fuzzy classifiers to be found in the literature, and here are some relevant works [14,15,2,3,13].

The most widely used fuzzy classifiers are the if-then rule based classifiers. The main advantage of using sets of fuzzy rules is that they make transparent the relationships between features and categories via the use of linguistic terms. However, notice that if the number of fuzzy rules increases, the number of linguistic terms used to label the fuzzy sets also increases. It follows that the linguistic informational burden of the user may also increase beyond reasonable limits.

Neuro-fuzzy classifiers represent a second class. The main advantage of using neuro-fuzzy classifiers is that the learning, adaptation and parallelism capabilities provided by neural networks may be used to tune the fuzzy rules parameters. The main drawback is represented by a possible too large number of parameters to be tuned, i.e. fuzzy membership functions and neural network weights.

The classifiers in the third class represent the set of categories as fuzzy subsets [10]. The performance of these classifiers is measured by their capability to learn the topological structure of the space. The fuzzy subsets used by Boudaoud and Masson [10] have been proposed within the fuzzy min-max clustering algorithm [18]. The fuzzy subsets defining the set of categories represent hyperboxes *B* defined by a minimum point *m* and a maximum point *M*. Figure 1 shows a hyperbox in R^3. It is important to mention that the hyperboxes are not allowed to overlap [18]. This does not mean that the areas in the features space corresponding to different states do not overlap, but that the hyperboxes delimit the subareas where points have full membership in different categories. Partial membership is achieved inside transition areas between hyperboxes.

It is to be noticed that an important property of the classifier in [10] is that the dimension of each hyperbox depends on only three constraints: its minimum point, its maximal point, and a parameter that controls the decreasing rate of membership to *B* value when the distance between a test point *u* and *B* increases. Thus, the main advantage of the third class of classifiers compared to the previous two classes is the smaller number of parameters to be tuned, i.e. 3 x number of categories considered, which leads to a smaller designing time for the classifier. However, the transparency of relationships between features and categories given by the use of linguistic terms is lost.



A Hyperbox in R^3 Defined by a Minimum Point and a Maximum.

The fuzzy classification methodology evaluated in this paper belongs to the last class mentioned and it is has been proposed by Bocaniala *et. al.* [8,9] in a fault diagnosis application. The main advantage of this methodology is the large accuracy with which it learns the topological structure of the features space. The fuzzy subsets built by the classifier approximate with a very small error the areas in the features space corresponding to different categories. Its accuracy also manifests through handling with fine precision the discrimination inside overlapping areas. It is to be noted that the classifier depends on a very small number of parameters, i.e. equal to the number of categories.

The paper is organized as follows. Section 2 presents the theoretical aspects of the fuzzy classification methodology. Section 3 presents the four tests used for evaluation of the performance of the classifier using data sets proposed by ELENA benchmark [12]. The ELENA benchmark has been built with the purpose of offering a collection of data sets that can be used as a test basis for newly developed classification methodologies. Section 4 gives a brief description of each data set in ELENA benchmark and details the actual results of the four tests proposed. Last section summarizes the paper contributions and mentions possible directions for future work.

1.1. Theoretical Aspects

The fuzzy classification methodology described in this section represents the set of considered categories using fuzzy subsets in the features space and it has been proposed by Bocaniala *et. al.* [8,9]. These fuzzy subsets are induced (built) on the basis of a point-to-set similarity measure between a point and a set of points in the features space [1]. The point-to-set similarity is built at its turn on the basis of a point-to-point similarity measure between points in the features space.

Point-to-point Similarity Measure Based on Distance Functions

The similarity between two points *u* and *v*, *s(u,v)*, may be expressed using a complementary function, *d(u,v)*, expressing dissimilarity. Baker [1] expresses dissimilarity by using the distance function in Eq. 1. Notice that, in this case, the functions *s* and *d* are complementary with regard to unit value, *s(u,v)*=1-*d(u,v)*. The β parameter plays the role of a threshold value for the similarity measure. For a data point *u*, all points *v* residing at a distance $\delta(u,v)$ smaller than β will bear some similarity with *u*. As for the points residing at distances larger or equal than β, the similarity $s(u,v)$ is null.

$$h^{\beta}\left(\delta(u,v)\right)=\begin{cases}\delta(u,v)/\beta, \text{ for } \delta(u,v)\leq\beta\\ 1, \text{ otherwise}\end{cases} \tag{1}$$

Point-to-set Similarity Measure

The similarity measure between two data points may be extended to a similarity measure between a point and a set of points [1]. In this paper, if the point-to-point similarity is given by Eq. 1, the similarity between a given point u and a set of points S is computed as the mean value of the point-to-point similarity values between u and each v in S (Eq. 2, where n denotes the number of elements in S). Notice that the value of $r(u,S)$ remains inside [0,1] interval, as $s(u,v)$ also remains inside [0,1] interval, and the cardinal of S is n.

$$r(u,S)=\frac{\sum_{v\in S}s(u,v)}{n} \tag{2}$$

The effect of using the β parameter is that only those data points from S, whose distance to u is larger than β, contribute to the point-to-set similarity value. The explanation is that only these points have a non-zero similarity with u. It follows that the similarity value between u and S is decided within the neighborhood defined by β. It has been observed in practice that, if different (dedicated) β parameters are used for different categories to express the point-to-point similarity (Eq. 1), the performance of the classifier increases substantially.

Fuzzy Subsets Induced by Single Point-to-set Similarity Measures

Let $C=\{C_i\}_{i=1,\ldots,m}$ be the set of all points in the features space, associated with the problem to solve, where C_i, $i=1,\ldots,m$ represents the set of all points corresponding to the i-th considered category. The membership function μ_i of the fuzzy subset $Fuzz_i$ induced by C_i, computed on the basis of a given point-to-set similarity measure, is given in Eq. 3. The n value represents the cardinal of C, and the n_i value represents the cardinal of C_i.

$$\mu_i(u)=\frac{r(u,C_i)}{r(u,C)} \tag{3}$$

A point u presented at the input of the classifier is assigned to the category C_z whose corresponding degree of assignment $\mu_z(u)$ is the largest (Eq. 4). In case of ties, the assignment to a category cannot be decided and the point is rejected.

$$u\in \text{z-th category} \Leftrightarrow \mu_z(u)=\max_{i=1,\ldots,m}\mu_i(u) \tag{4}$$

Fuzzy Subsets Induced by Multiple Point-to-set Similarity Measures

The practice showed that there are problems for which classifiers designed by using only one point-to-point similarity measure does not provide satisfactory results [7]. When situations like these are met, the advantages brought by two or more similarity measures may be combined in order to improve the performance of the classifier [8], i.e. a hybrid approach is used. This aspect has been also noticed by Baker [1].

In the following, a few possible approaches, when trying to combine the use of two or more similarity measures, are suggested. *Point-to-point similarity measures*: the β parameter may be applied only to one of the similarity measures used; if more than one similarity measure are used, then there is a β parameter for each one of them. *Point-to-set affinity measures*: there may be only one point-to-set affinity measure resulted from the combination of all similarities used; or, there may be one cluster affinity measure for each similarity used. *Fuzzy membership functions*: the fuzzy membership functions represent combinations of cluster affinity measures, if more than one such measure exist.

Designing and Testing the Classifier

Let m be the number of the categories considered for the problem to be solved. The validated methodology first groups the set of all available data C into subsets according to the category they belong to, C_i, $i=1,\ldots,m$. In order to design and test the classifier, each subgroup C_i is split into three representative and distinct subsets, C_i^{ref}, C_i^{param}, and C_i^{test}. On the basis of these subsets, three sets unions, *REF*, *PARAM* and *TEST*, are defined (Eq. 5). They are called the *reference patterns* set, the *parameters tuning* set, and respectively the *test* set. A subset is considered representative for a given set if it covers that set in a satisfactory manner. In the following, the semantic adopted for the expression *satisfactory covering subset* [16] is explained. Then, the role of each one of the three unions is detailed. It is to be noticed that the union of subsets having the satisfactory covering property for a set represents also a satisfactory covering subset of that set.

$$\begin{aligned} REF &= \bigcup_{i=1}^{m} C_i^{ref} \\ PARAM &= \bigcup_{i=1}^{m} C_i^{param} \\ TEST &= \bigcup_{i=1}^{m} C_i^{test} \end{aligned} \tag{5}$$

1.1.1. Satisfactory Covering Subsets

For the work presented in this paper, a *satisfactory covering subset* represents a subset of data that preserves (with a given order of magnitude) the distribution of the data associated with the problem. Selecting the elements that compose a satisfactory covering subset for a given data set can be costly. Therefore, it is more convenient to use selection methods that provide convenient approximations for satisfactory covering subsets. Such a method [16] is presented in the followings.

Let consider a given finite data set A that contains r points in a multidimensional space. First, the maximum distance, *max*, between two elements is computed. During this computation a pair of elements, (a,b), with maximum distance between them is memorized. Then, one of the elements, let it be a, is considered as the center of s hyperspheres, S_i, $i=1,...,s$. The user must provide the s value. Each one of the S_i hyperspheres has the radius equal to

$$r_i = i\frac{max}{s},\ i = 1,...,s \tag{6}$$

The next step is to consider the partition induced by next subsets,

$$\begin{aligned} P_0 &= \{a \in A / a \ inside \ S_1\} \\ P_j &= \{a \in A / a \ inside \ S_{j+1} - S_j\},\ j = 1,...,s-1 \end{aligned} \tag{7}$$

The cardinal of the subset that approximates the satisfactory covering subset is set to a previous given percent t of elements from A. The distribution of elements from A in the partition elements P_0, …, P_{s-1} is not equal. This distribution is taken into account when distributing the percent t among the partition members. Each partition member P_j, $j=1$, …, $s-1$, will be allocated a number of p_j elements. The approximation subset is composed by randomly selecting p_j elements from the P_j subset, $j=1$, …, $s-1$.

1.1.2. Reference Patterns Set (REF)

The point-to-set similarity measures are defined for the representative subsets C_i^{ref}, $i=1,…,m$. Therefore, when using a single point-to-set similarity measure, the fuzzy membership functions are computed as shown in Eq. 8.

$$\mu_i(u) = \frac{r(u, C_i^{ref})}{r(u, C)} \tag{8}$$

1.1.3. Parameters Tuning Set (PARAM)

The shape of the membership functions μ_i, associated to the fuzzy sets $Fuzz_i$, depends not only on the representative subset C_i^{ref}, but also on the value of the β_i parameter, $i=1,…,m$. The algorithm for tuning the parameters β_i, $i=1,…,m$, of the classifier represents a search process in a m-dimensional space for the parameter vector (β_1, β_2,..., β_m) that meets, for each category, the maximal correct classification criterion and the minimal misclassification criterion. In order to perform this search, different methodologies may be used, i.e. genetic algorithms [8], hill-climbing [5] and particle swarm optimization (PSO) [6]. In practice, the PSO methodology proved to be the fastest.

The search for optimal parameters when using genetic algorithms and hill-climbing may be accelerated by using an *optimized initial population* [16]. An optimized initial population can be obtained by performing an iterative search that starts with an individual whose parameters have very small values. Then, at each next step, the values of the parameters will be increased/decreased so that the fitness of the obtained individual, i.e. the classifier performance, increases.

1.1.4. Testing Set (TEST)

The performance of the classifier is measured according to its generalization capabilities when applied on the *TEST* set. It is to be noticed that the *TEST* set contains data that were not presented before at the input of the classifier and that is representative for the whole data set *C*. The practice showed that the performance of the classifier might improve if the testing is performed after adding the data in the *PARAM* set to the *REF* set.

2. Evaluation of the Performance of the Validated Fuzzy Classification Methodology Using ELENA Benchmark

The ELENA benchmark [12] has been built with the purpose of offering a collection of data sets that can be used as a test basis for newly developed classification methodologies. It is important to note that all data sets in ELENA benchmark are static sets, i.e. they do not represent series of measurements that may be ordered in time and whose current values depend on the outcome of past measurements, as it is the case for dynamic sets. For an example of how to apply the fuzzy classifier on a dynamic set see the case study in [8,9].

In order to evaluate the performance of the validated fuzzy classifier, four tests are performed for each data set. The first and the main test is the comparison between the fuzzy classifier performance and the performance of the kNN classifier as it has been recorded by benchmark. The second test is the steadiness of the classification error obtained with different initial conditions for the fuzzy classifier. The third set is checking the generalization capability of the classifier. The last test is the visualization of the fuzzy subsets surfaces for data sets with two dimensions. Each test is described in the following in a separate subsection.

Comparison with the kNN Classifier

The data sets in ELENA benchmark have been used to compare the performances of different classifiers selected by the authors of benchmark. The classifiers have been selected so that they represented well-known techniques in the classification field, i.e. Multilayer Perceptron Neural Networks or kNN (*k*-nearest neighbors) classifiers, and techniques developed by the authors of the benchmark. Out of these classifiers, the kNN classifier is the most similar to the validated fuzzy classification methodology. Indeed, as described in Section 1, the validated methodology performs classification using the points located inside neighborhoods delimited by the β parameter values. The kNN classifiers work in a similar manner, by performing classification using the neighborhood formed by the *k* closest points. Therefore, the performance of the validated methodology is evaluated by comparison with the kNN classifier performance indicated in benchmark. Another reason for choosing the kNN classifier is the fact that it has always either the best performance or the difference between its performance and the best one is minor, on the data sets in the benchmark.

The performance of the kNN classifier on the data sets in the benchmark is measured using the Leave-One-Out test. This test is carried out as follows: if *N* samples are available in the data set, then the classifier is designed using *N-1* samples and it is tested on the sample leaved out. This procedure is repeated *N* times and each time a different sample is leaved out to be tested. Note that the design set has the maximum possible size, *N-1*, and that this fact increases the reliability of the classification result for the sample leaved out. It was proven that, if the theoretical Bayes error may be computed, then the classification error given by Leave-One-Out test represents an upper bound.

Measuring the performance of the validated classifier on the data sets in the benchmark using the Leave-One-Out test represents a very costly operation. The number of calls of the classifier needed may be reduced by modifying the Leave-One-Out test by putting in the *TEST* set more than one sample at a time. The elements in the *TEST* set are selected so that, after a number of calls of the classifier, each sample in the data set to turn into a member of the *TEST* set no more than once. Notice that the difference to the Leave-One-Out test, i.e. the cardinal of *TEST* set is larger than one, may cause the classification error obtained with the proposed modification to be the same or smaller than the classification error obtained with the Leave-One-Out test.

Steadiness of Classification Error Test

The purpose of this test is to show that different initial conditions bring only minor modifications of the classification error obtained with the validated fuzzy methodology. Different initial conditions are obtained by using different distributions of data into the reference set (*REF*), the parameters tuning set (*PARAM*) and the test set (*TEST*) subsets. However, the percentages allocated to each set are kept constant: 30% for the *REF* set, 40% for the *PARAM* set, and 30% for the *TEST* set. The classifier is designed and tested for ten different initial conditions and the standard deviation criterion is used to assess the stability of the classification error value.

Generalization Capability Test

The generalization capability of the fuzzy classifier is assessed by checking the difference between the classification error obtained for the *PARAM* set, i.e. when tuning the parameters, and the classification error obtained for the *TEST* set. The checking is done for ten different initial conditions, i.e. different distributions of samples into the reference set (*REF*), the parameters tuning set (*PARAM*) and the test set (*TEST*) subsets. The percentages allocated to each set are kept constant: 30% for the *REF* set, 40% for the *PARAM* set, and 30% for the *TEST* set. The standard deviation criterion is used to assess the level of generalization achieved. Notice that the results of the ten experiments for the steadiness of classification error test may be also used for this test.

Fuzzy Subsets Visualization Test

The main advantage of the validated classification methodology is the high accuracy with which it learns the topological structure of the space associated with the problem. The fuzzy subsets built by the classifier approximate with a very small error the areas corresponding to different categories. Its accuracy also manifests through handling with fine precision the discrimination inside overlapping areas. In the case of data sets with two dimensions it is possible to assess visually these properties. The areas covered by the fuzzy subsets surfaces may be visually compared with the actual areas occupied by different categories.

3. The Results of Evaluation of the Fuzzy Classifier Performance Using ELENA Benchmark

ELENA benchmark contains three artificial data sets called CLOUDS, GAUSSIAN and CONCENTRIC. They have been generated so that to fulfill the next requirements: (*i*) heavy overlapping of the areas covered by different categories, (*ii*) high degree of non-linearity of category boundaries, and (*iii*) various dimensions of associated problem space

The GAUSSIAN data set contains seven different subsets that differ in the dimension of the problem space. The dimension of the problem space goes from 2 to 8. In this paper, only the subset with dimension 2 is considered.

There also are four real data sets in ELENA benchmark called IRIS, PHONEME, SATIMAGE and TEXTURE. They have been selected according to following requirements: (*i*) classical data sets in the classification field (IRIS), (*ii*) sufficient number of samples to avoid the "*empty space phenomenon*", (*iii*) high number of classes (TEXTURE), and (*iv*) various dimensions of associated problem space

The "*empty space phenomenon*" manifests in the case of data sets that contain a small number of samples. In this case, the areas corresponding to different categories may not be well delimited in the space associated with the problem. For more details, see the discussion on the IRIS data set that confronts this phenomenon as it contains only 150 samples.

It is important to note that all data sets in ELENA benchmark are static sets, i.e. they do not represent series of measurements that may be ordered in time and whose current values depend on the outcome of past measurements as it is the case for dynamic sets. For an example of how to apply the fuzzy classifier on a dynamic set, see the case study in [8,9].

Results of the Tests on the CLOUDS Data Set

The characteristics of this data set are the heavy overlapping between areas covered by different categories and the high degree of non-linearity of the category boundaries. There are 5000 samples equally distributed in 2 categories. The data set has dimension 2 and it is shown in 0.

The classification error given by the kNN classifier with the Leave-One-Out test is 10.94 +/- 1.2 %. The classification error given by the fuzzy classifier with the test proposed in Subsection 3.1 is 10.82 %. The corresponding confusion matrices are shown in 0. The classification error obtained with the fuzzy classifier is practically equal to the one obtained with the kNN classifier.

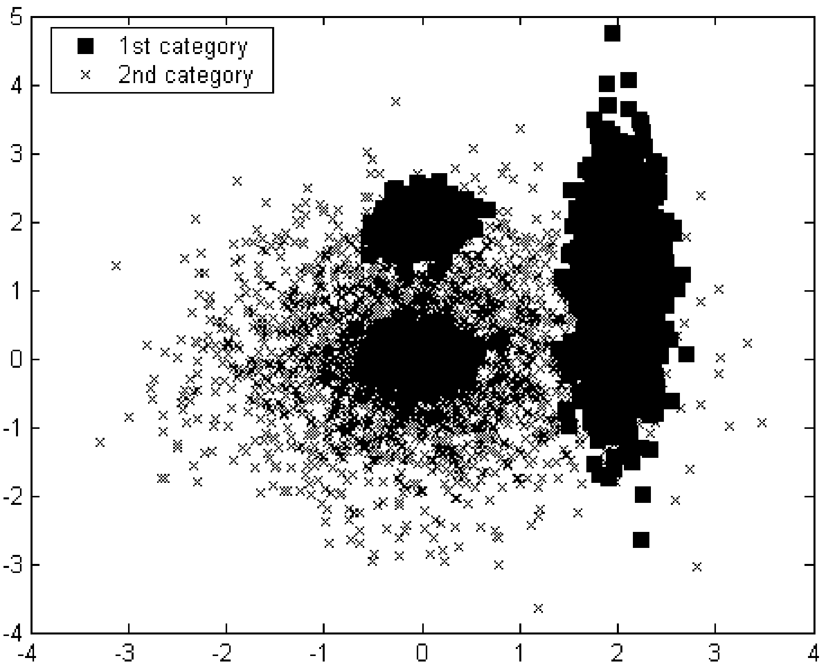
The standard deviation values for the classification error steadiness test and the generalization capability test are 0.49 % and respectively 0.53 %. These values are very small and it may be inferred that the classification error is steady in a small interval around the average value 10.98 %, and that the classifier generalizes very well on the CLOUDS data set.

The fuzzy subsets surfaces corresponding to the two categories in CLOUDS data set are shown in 0. The areas covered by the two surfaces match the areas occupied by the two surfaces in 0. It is also noticeable the fine precision of discrimination inside overlapping areas.

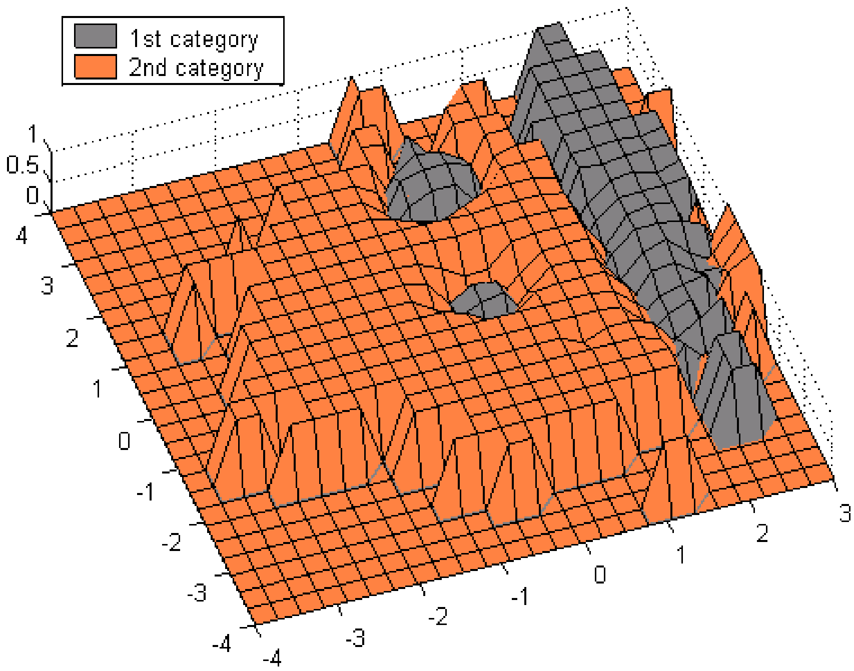
1	2	rejected
95.56	4.36	0.08
17.16	82.80	0.04

	1	2
1	92.8 +/- 1.0	7.2 +/- 1.0
2	14.7 +/- 1.4	85.3 +/- 1.4

Confusion Matrices for CLOUDS: Fuzzy Classifier (left), kNN Classifier (right)



The Two Categories in CLOUDS Data Set



The Fuzzy Subsets Surfaces for the two Categories in CLOUDS Data Set

Results of the Tests on the GAUSSIAN Data Set

The GAUSSIAN data set contains seven different subsets that differ in the dimension of the problem space. The dimension of the problem space goes from 2 to 8. Here, for visualization purposes, only the subset with dimension 2 is considered. The characteristics of this data set are the same as for the CLOUDS data set, i.e. heavy overlapping between areas covered by different categories and the high degree of non-linearity of the category boundaries. There are 5000 samples equally distributed in 2 categories. The data set has dimension 2 and it is shown in 0.

The classification error given by the kNN classifier with the Leave-One-Out test is 27.3 %. The classification error given by the fuzzy classifier with the test proposed in Subsection 3.1 is 27.16 %. The corresponding confusion matrices are shown in **Error! Reference source not found.**. The classification error obtained with the fuzzy classifier is practically equal to the one obtained with the kNN classifier.

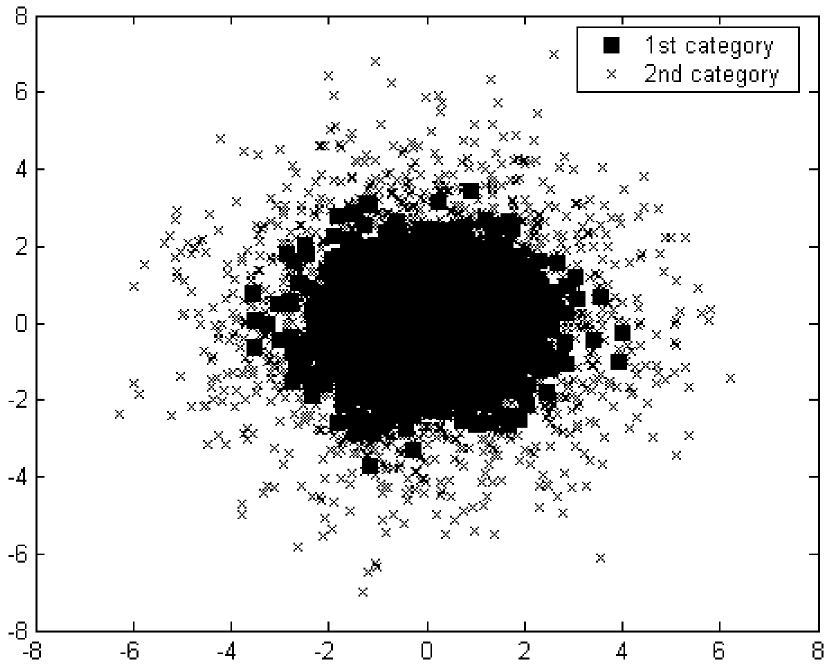
The standard deviation values for the classification error steadiness test and the generalization capability test are 0.63 % and respectively 0.89 %. These values are very small and it may be inferred that the classification error is steady in a small interval around the average value 26.17 %, and that the classifier generalizes very well on GAUSSIAN data set.

The fuzzy subsets surfaces corresponding to the two categories in GAUSSIAN data set are shown in 0. The areas covered by the two surfaces match reasonably the areas occupied by the two surfaces in 0. The large overlapping between the two categories is reflected by the size and the shape of the surface corresponding to 1st category. The area occupied by the 1st category in 0 may be approximated by a circle with radius 2, while in 0 the corresponding surface covers the same area. This implies that all samples belonging to the 2nd category located inside the circle will be misclassified.

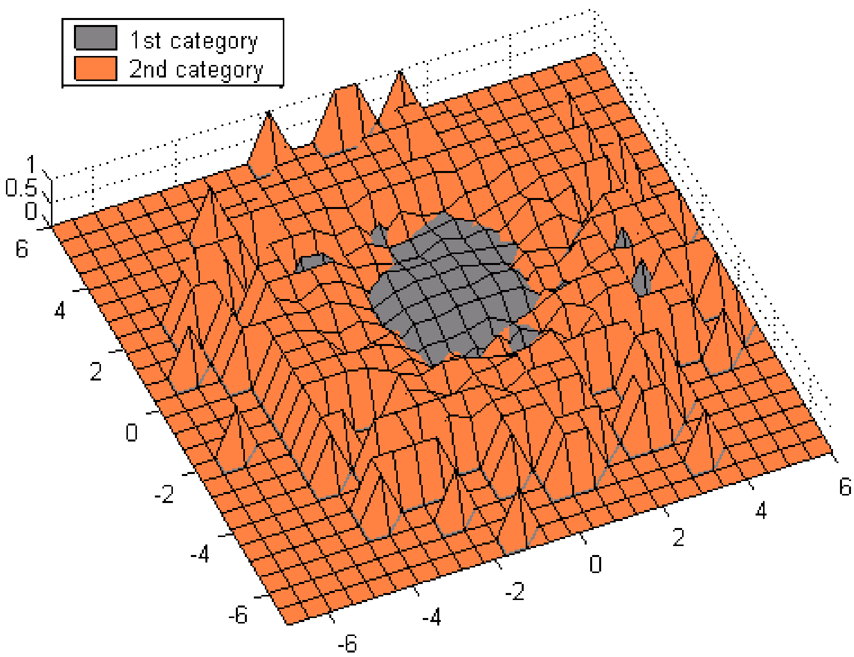
	1	2	rejected
1	88.08	11.92	0.00
2	42.36	57.60	0.04

	1	2
1	84.4	15.6
2	39.1	60.9

Confusion matrices for GAUSSIAN: fuzzy classifier (left), kNN classifier (right)



The two categories in GAUSSIAN data set



The Efuzzy Subsets Surfaces for the two Categories in GAUSSIAN Data Set

Results of the Tests on the CONCENTRIC Data Set

This data set contains two categories, the first one being nested into the second one (0). The main characteristic of this data set is the lack of overlapping between areas covered by the two categories. There are 2500 samples distributed as follows: 921 (36.8 %) in the 1st category and 1579 (63.2 %) in the 2nd category.

The classification error given by the kNN classifier with the Leave-One-Out test is 0.96 +/- 0.5 %. The classification error given by the fuzzy classifier with the test proposed in Subsection 3.1 is 1.12 %. The corresponding confusion matrices are shown in **Error! Reference source not found.**. The difference between the two classification errors is negligible.

The standard deviation values for the classification error steadiness test and the generalization capability test are 0.67 % and respectively 0.56 %. These values are very small and it may be inferred that the classification error is steady in a small interval around the average value 1.91 %, and that the classifier generalizes very well on CONCENTRIC data set.

	1	2
1	98.15	1.85
2	0.70	99.30

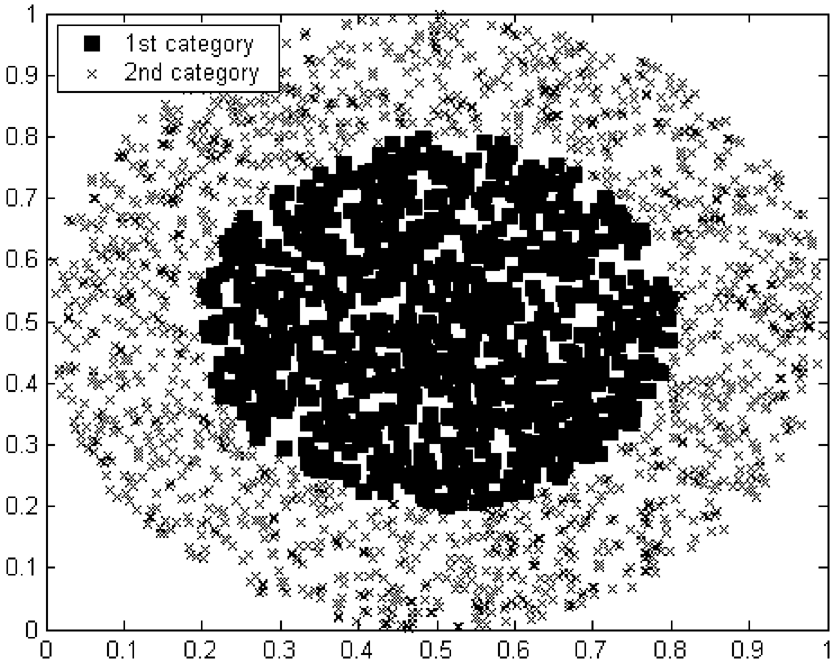
	1	2
1	99.1 +/- 0.6	0.9 +/- 0.6
2	1.0 +/- 0.5	99.0 +/- 0.5

Confusion Matrices for CONCENTRIC: Fuzzy Classifier (left), kNN Classifier (right)

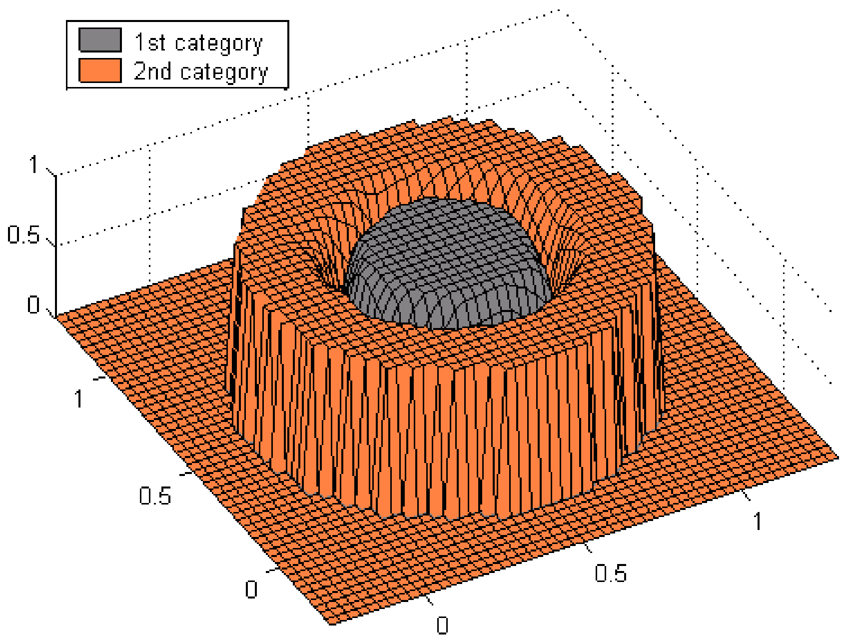
The fuzzy subsets surfaces corresponding to the two categories in CONCENTRIC data set are shown in 0. The areas covered by the two surfaces match the areas occupied by the two surfaces in 0. It is also noticeable the fine precision of discrimination inside overlapping area.

Results of the Tests on the IRIS Data Set

This data set is taken from "UCI Repository of Machine Learning Databases and Domain Theories" (http://www.ics.uci.edu/~mlearn/MLRepository.html). It is a very simple data set. The data set contains three categories corresponding to three types of iris plant. There are only 150 samples equally distributed among the 3 categories. Each plant is described by four measurements: sepal length, sepal width, petal length and petal width.



The two Categories in CONCENTRIC Data Set



The Fuzzy Subsets Surfaces for the two Categories in CONCENTRIC Data Set

	1	2	3
1	100	0.0	0.0
2	0.0	94.0	6.0
3	0.0	6.0	94.0

	1	2	3
1	100	0.0	0.0
2	0.0	96.0	4.0
3	0.0	2.0	98.0

Confusion Matrices for IRIS: Fuzzy Classifier (left), kNN Classifier (right)

The classification error given by the kNN classifier with the Leave-One-Out test is 3.3 %. The classification error given by the fuzzy classifier with the test proposed in Subsection 3.1 is 4 %. The corresponding confusion matrices are shown in **Error! Reference source not found.**. The difference between the classification errors is very small.

The standard deviation values for the classification error steadiness test and the generalization capability test are 3.05 % and respectively 3.70 %. The average values are 9.11 % and respectively 5.08 %. These large variations may be explained by the small number of samples available for this data set. This is called the "*empty space phenomenon*".

Results of the Tests on the PHONEME Data Set

This data set has been built with the purpose of distinguishing between nasal and oral vowels. There are 5404 samples distributed as follows: 3818 (70.65 %) samples for nasals vowels category and 1586 (29.35 %) for orals vowels category. Each sample is described by five measurements.

	1	2	rejected
1	89.47	9.80	0.73
2	18.41	81.59	0.0

	1	2
1	95.0	5.0
2	18.5	81.5

Confusion Matrices for PHONEME: Fuzzy Classifier (left), kNN Classifier (right)

The classification error given by the kNN classifier with the Leave-One-Out test is 8.97 %. The classification error given by the fuzzy classifier with the test proposed in Subsection 3.1 is 12.84 %. The corresponding confusion matrices are shown in **Error! Reference source not found.**. The classification error obtained with kNN classifier is smaller than the one obtained with the fuzzy classifier. The content of the two confusion matrices shows that the kNN classifier has better performance for the 1st category. The difference may appear due to the settings of the test used instead Leave-One-Out test (see Section 3.1).

The standard deviation value for the classification error steadiness test is 1.4 %. It may be inferred that the classification error is steady in a small interval around the average value 13.18 %. The standard deviation value for the generalization capability test is 3.59%, which indicates poor generalization on PHONEME data set.

	No. samples	*Percentage*
1	1533	23.82 %
2	703	10.92 %
3	1358	21.10 %
4	626	9.73 %
5	707	10.99 %
5	1508	23.43 %

Distribution of Samples Into the 6 Categories from SATIMAGE Data Set

Results of the Tests on the SATIMAGE Data Set

This data set is taken from "UCI Repository of Machine Learning Databases and Domain Theories". The data set has been built with the purpose of distinguishing between 6 types of elements that may appear in an image taken by satellite: red soil, cotton crop, grey soil, damp grey soil, soil with vegetable stubble, and very damp grey soil. There are 6435 samples distributed in different categories according to 0. Each sample is described by 65 measurements.

The classification error given by the kNN classifier with the Leave-One-Out test is 8.89 +/- 1.6 %. The classification error given by the fuzzy classifier with the test proposed in Subsection 3.1 is 8.94 %. The corresponding confusion matrices are shown in **Error! Reference source not found.**. The difference between the two classification errors is negligible.

	1	2	3	4	5	6	rejected
1	98.17	0.07	1.17	0.0	0.46	0.0	0.13
2	0.0	97.44	0.28	0.57	1.14	0.28	0.28
3	0.52	0.15	97.05	1.55	0.07	0.66	0.0
4	0.32	0.64	22.04	65.81	0.48	10.70	0.0
5	5.09	0.28	0.0	1.41	87.27	5.52	0.42
6	0.0	0.27	2.25	7.63	1.92	87.73	0.20

	1	2	3	4	5	6
1	98.1	0.2	1.1	0.1	0.5	0.0
2	0.0	96.5	0.1	0.7	2.0	0.7
3	0.5	0.1	93.04	4.6	0.0	1.4
4	0.0	0.8	13.7	70.6	0.8	14.1
5	0.0	0.1	1.9	7.3	2.0	88.7
6	0.0	0.27	2.25	7.63	1.92	87.73

Confusion Matrices for SATIMAGE: Fuzzy Classifier (left), kNN Classifier (right)

The standard deviation value for the classification error test is 0.71 % and it may be inferred that the classification error is steady in a small interval around the average value 11.76 %. The standard deviation for the generalization capability test is small, 1.28 %, around a small average, 0.69 %. This does not affect critically the generalization capability of the classifier on the SATIMAGE data set.

Results of the Tests on the TEXTURE Data Set

The characteristics of this data set are the large number of classes and the large number of measurements. This data set has been built with the purpose of distinguishing between11 types of textures. There are 5500 samples equally distributed among the 11 categories. Each sample is described by 40 measurements.

The classification error given by the kNN classifier with the Leave-One-Out test is 1.0 +/- 0.8 %. The corresponding confusion matrix is shown in **Error! Reference source not found.**. The classification error given by the fuzzy classifier with the test proposed in Subsection 3.1 is 2.82 %. The corresponding confusion matrix is shown in **Error! Reference source not found.**. The difference between the two classification errors is small.

	1	2	3	4	5	6	7	8	9	10	11
1	97.0	1.0	0.4	0.0	0.0	0.0	1.6	0.0	0.0	0.0	0.0
2	0.2	99.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.4
3	1.0	0.0	98.8	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	99.4	0.0	0.0	0.0	0.6	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0
6	0.0	0.0	0.0	0.0	0.0	98.6	0.0	1.4	0.0	0.0	0.0
7	0.4	0.0	0.2	0.0	0.0	0.2	98.8	0.4	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	1.4	0.0	98.6	0.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.8	0.2
11	0.0	0.4	0.0	0.0	0.0	0.4	0.0	0.0	0.2	0.0	99.0

Confusion Matrix for TEXTURES Obtained with the kNN Classifier

	1	2	3	4	5	6	7	8	9	10	11	rejected
1	96.0	1.0	0.1	0.0	0.0	0.0	1.8	0.0	0.0	0.0	0.0	0.2
2	0.2	98.6	0.0	0.0	0.0	0.0	0.2	0.4	0.0	0.0	0.4	0.2
3	1.0	0.0	99.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	99.0	0.0	0.0	0.2	0.8	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	0.2	0.2	0.0	0.0	0.0	90.4	0.6	6.6	0.0	0.0	0.4	1.4
7	0.8	0.0	0.8	0.0	0.0	0.2	96.0	1.8	0.0	0.0	0.0	0.4
8	0.0	0.2	0.0	0.8	0.0	0.8	0.4	95.2	0.0	0.0	0.0	2.6
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	98.8	0.2	1.0
11	0.0	0.2	0.0	0.0	0.0	1.8	0.0	0.8	0.2	0.4	96.0	0.6

Confusion Matrix for TEXTURES Obtained with the Fuzzy Classifier

The standard deviation value for the classification error test is 0.54 % and it may be inferred that the classification error is steady in a small interval around the average value 2.64 %. The standard deviation for the generalization capability test is 0.71 %, around a very small average value, 0.13 %. This fact proves a good generalization on the TEXTURE data set.

4. Conclusions

This paper evaluated the performance of a novel fuzzy classification methodology. The high accuracy with which it learns the topological structure of the classification space and its fine precision the discrimination inside overlapping areas have been validated using ELENA benchmark data sets [12]. The two properties have been confirmed by the results of the four tests proposed, i.e. comparison with kNN classifier performance, steadiness of classification error, generalization capabilities, and surfaces visualization.

Future research on the proposed classification methodology needs to concentrate on obtaining a computational complexity for both the design and the test phase that is small enough to make the classifier suitable for application to fault diagnosis of real systems. The computational complexity of the design phase has already been significantly reduced by using the particle swarm optimization technique, as presented in our recent paper [5]. Also, it has been observed in practice that the classifier generalizes reasonably well even for small dimensions of the *REF* and *PARAM* sets [4]. The computational complexity of both the design phase and the test phase depends heavily on the size of these two sets. This leads to the conclusion that a technique should be found so that the size of these two sets reduces substantially and the performance of the classifier stays at least the same. A solution to this might be obtained by applying kernel methods [17].

Acknowledgements

This work has been partially supported by Fundação para a Ciência e a Tecnologia, Minister of Science, Innovation and Technology, Portugal, grant number SFRH/BD/18651/2004.

REFERENCES

1. BAKER, E., **Cluster analysis by optimal decomposition of induced fuzzy sets**, PhD thesis, DELFTSE UNIVERSITAIRE PRES, Delft, Holland, 1978.
2. BEZDEK, J.C., **Pattern Recognition with Fuzzy Objective Function Algorithms**, KLUWER, 1981.
3. BEZDEK, J.C., PAL S.K., **Fuzzy Models for Pattern Recognition**, IEEE PRESS, Piscataway, NJ, 1992.
4. BOCANIALA, C.D.**, Tehnici de inteligenţă artificială aplicate în diagnoza defectelor: Aplicaţii ale tehnicilor de clasificare**, Technical Research Report within doctoral training, University "Dunarea de Jos" of Galati, Romania. (available in English for download at http://www.gcar.dem.ist.utl.pt/Pessoal/Cosmin/publicat.htm)
5. BOCANIALA, C.D. and SA DA COSTA, J., **Tuning The Parameters Of A Fuzzy Classifier For Fault Diagnosis. Hill-Climbing vs. Genetic Algorithms**, Proceedings of the Sixth Portuguese Conference on Automatic Control CONTROLO 2004, 7-9 June, Faro, Portugal, 2004.
6. BOCANIALA, C.D. and SA DA COSTA, J., **Tuning The Parameters Of A Fuzzy Classifier For Fault Diagnosis. Particle Swarm Optimization vs. Genetic Algorithms**, Proceedings of the 1st International Conference on Informatics in Control, Automation and Robotics ICINCO 2004, 25-28 August. Setubal, Portugal, 2004.
7. BOCANIALA, C.D., SA DA COSTA, J. and LOURO R., **A Fuzzy Classification Solution for Fault Diagnosis of Valve Actuators**, Proceedings of 7th International Conference on Knowledge-Based Intelligent Information & Engineering Systems, Oxford, UK, September 3-5, LNAI Series, Springer-Verlag, Heidelberg, Germany, 2003.
8. BOCANIALA, C. D., SA DA COSTA, J. and PALADE, V., **A Novel Fuzzy Classification Solution for Fault Diagnosis**, J. INTELLIGENT AND FUZZY SYSTEMS 14(3), IOS Press, 2004.
9. BOCANIALA, C. D., SA DA COSTA, J. and PALADE, V., **Refinement of the Diagnosis Process Performed with a Fuzzy Classifier**, Proceedimgs of the 8th International Conference on Knowledge-Based Intelligent Information & Engineering Systems, Wellington, New Zealand, 2004.
10. BOUDAOUD, N. and MASSON, M., **Diagnosis of transient states using pattern recognition approach**, JESA 3, 2000.
11. DUDA, R. O., HART, P. E. and STORK D. G., **Pattern Classification**, JOHN WILEY, 2001.
12. ESPRIT Basic Research Project No. 6891, ELENA (Enhanced Learning for Evolutive Neural Architecture), ftp://ftp.dice.ucl.ac.be/ pub/neural-net/ELENA/databases.
13. HÖPPNER, F., KLAWONN, F., KRUSE, R. and RUNKLER, T., **Fuzzy Cluster Analysis**, JOHN WILEY, 1999.
14. KUNCHEVA, L. I., **Fuzzy Classifier Design**, Studies in Fuzziness and Soft Computing Series, SPRINGER-VERLAG, 2000.
15. PAL, S. K. and MITRA, S., **Neuro-fuzzy Pattern Recognition: Methods in Soft Computing Paradigm**, JOHN WILEY, 1999.
16. SA DA COSTA, J., BOCANIALA, C.D. and LOURO R., **A Fuzzy Classifier for Fault Diagnosis of Valve Actuators**, Proceedings of the IEEE Conference on Control Applications CCA 2003, Istanbul, Turkey, 2003.
17. SHAWE-TAYLOR, J. and CRISTIANINI, N., **Kernel methods for pattern analysis**, CAMBRIDGE UNIVERSITY PRESS, 2004.
18. SIMPSON, P., **Fuzzy Min-Max neural networks, part 1: classification**, IEEE TRANS ON NEURAL NETWORKS 3(5), 1992.