Cumulative Fuzzy Class Membership Criterion Decision-based Classifier

Patrik Sabol[1], Peter Sinčák[1], Ján Buša[2], Pitoyo Hartono[3]
[1]Technical University of Kosice / Department of Artificial Intelligence and Cybernetics, Košice, Slovakia
[2]Technical University of Kosice / Department of Mathematics and Theoretical Informatics, Košice, Slovakia
[3]School of Engineering, Chukyo University, Nagoya, Japan
{patrik.sabol, peter.sincak, jan.busa}@tuke.sk, hartono@ieee.org

***Abstract*— This paper deals with classification algorithms as one of the basic principles of pattern recognition. We analyze their effect to a feature space and compare the type and the shape of the separating and decision surface, respectively. We proposed a novel classification approach based on Cumulative Fuzzy Membership Function that creates a decision surface in a different way as an MF ARTMAP neural network. We call the proposed decision surface Cumulative Fuzzy Class Membership Criterion (CFCMC), which we compared with the decision surface of MF ARTMAP termed as Membership Function. The analysis of both decision surfaces shows that CFCMC has better adaptability and flexibility in forming a decision boundary than Membership Function from MF ARTMAP classifier. Based on the result of this analysis we assumed that classifier built based on CFCMC should achieve higher classification accuracy than the one built based on Membership Function. Furthermore, we identified some issues, solutions and possible future challenges of our proposed novel method, such as the expansion into incremental learning and semantic information extraction.**

Keywords—component; feature space, decision surface, fuzzy membership function of unknown input to class, fuzzy sets, neural networks, classification accuracy, "not classified" pattern

I. INTRODUCTION

Classification is one of the basic principles of pattern recognition, where it plays important roles. The principles are presented in Fig. 1. As it is obvious from the Fig. 1, the determination of the feature space plays a key role in overall pattern recognition process. Feature selection and classification are based on this information, and the ability to approximate the nonlinear discrimination function determines the success of a classification procedure. These principles raise a question of whether it is more useful to invest time for selecting a good set of features or to develop a general and robust classification procedure based on the all given features. Certainly, these two problems are closely related and usually feature selection techniques assume a specific type of classifier. The first approach is based on the consideration that the selection of relevant features help the classifying process while the second one is based on the consideration that features selection itself is inherent in the learning process of the classifier without the necessity for features selection. Often, a combination of them could be the best approach for real-world applications [1][2].

In this paper, we focus mainly on the formation of the decision boundary for classification. We compare classifiers with respect to their ability to form relevant decision boundaries. Here, we present a novel fuzzy modeling approach of Cumulative Fuzzy Class Membership Criterion (CFCMC) that creates decision hyper-surface over feature space for classification problems. This method is compared with a similar classifier, namely neural network Membership Function Adaptive Resonance Theory MAP (MF ARTMAP), which is based on fuzzy membership decision. The advantage of both methods is that they are not "black-box" in a strictly way, in contrast to, e.g. Multi-Layer Perceptron because they created decision surfaces that can offer semantic information about one particular class as well as relations among classes. The advantage of CFCMC over MF ARTMAP is that it is able to form a decision surface with better adaptability and flexibility. Therefore, it provides higher classification accuracy and better generalization ability.

In recent years, another different fuzzy modeling approaches were built and used for different tasks. In [3], an adaptive Evolving Fuzzy System-based approach for creating dynamic and evolving model of Activities of Daily Living, which are described by one or more fuzzy rules, were proposed. New framework for the symbolic representation of data, referred to as signatures is presented in [4]. Authors explain signatures as convenient hierarchical symbolic representation of data structuring into vectors of fuzzy values. This data structure includes values and also the significance of these values by the nested vector construction. Fuzzy signatures in conjunction with fuzzy S-trees were also used for fuzzy medical image retrieval (FMIR) using vector quantization in the medical area, concretely in mammography [5].

The paper is organized as follows: Section II explains about the difference of the decision boundaries in the feature space. In Section III explains about CFCMC and its comparison with Membership Function. Section IV presents the learning algorithm of the proposed method. Section V describes experiments and results and Section VI attempts to raise issues of the proposed method. Conclusion and future work are given in Section VII.

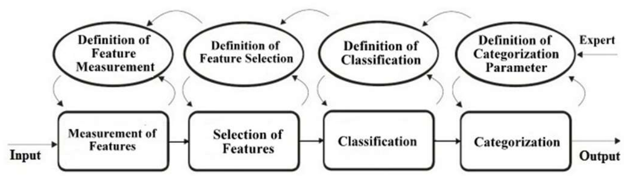


Fig. 1 Basic principles of pattern recognition

II. Classifiers and Feature Space

A. Definition of feature space

Let's consider data where the i-th pattern is described with n features as follows:

$$\overline{x_i} = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})^T$$

Hence, pattern $\overline{x_i}$ is a point in an abstract space $\mathbb{R}^n$.

B. Learning algorithms and feature space

Classifiers differ from one another in the separating (decision) boundary they create in the feature space. Given a problem with n features, then we have two types of approaches.

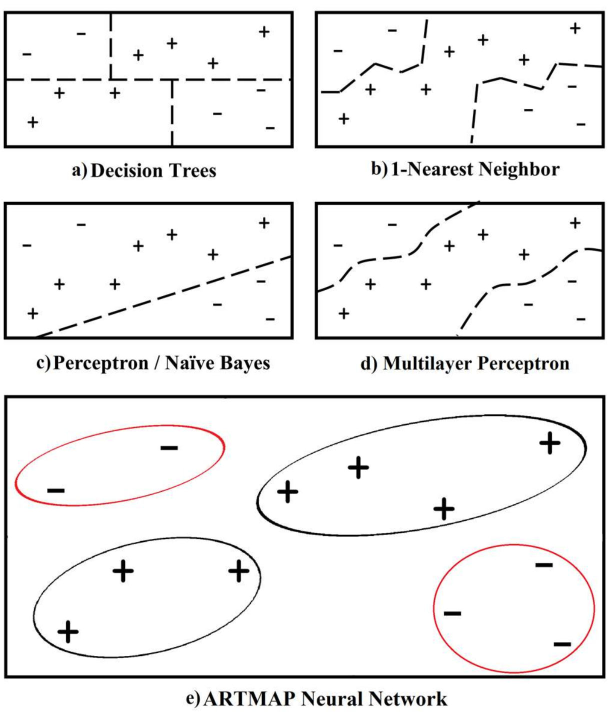


Fig. 2 Different learning models and their effect on feature space [13]

The first group of classifiers calculates $n-1$ dimensional classification boundaries in the n dimensional feature space. For instance, Decision trees [6] algorithm recursively splits up patterns based on one feature at a time and draw axis parallel boundaries to separate two classes (Fig. 2 a)). The k-nearest neighbors [7] algorithm draws a Voronoi diagram, which construct convex polygons around the patterns for a more complex partitioning (Fig. 2 b)). Perceptron [8] is a linear classifier. Thus, it creates hyperplane in the feature space just as a naïve Bayes classifier [9] does (Fig. 2 c)). Multi-Layered Perceptron (MLP) [10] is the extension of linear Perceptron being able to run non-linear classification (Fig. 2 d)). The ARTMAP neural networks [11] belong to the model of neural networks based on Adaptive Resonance Theory (ART) [12]. The fundamental idea of basic ARTMAP neural networks is the clustering of feature space and mapping created clusters to classes. That leads to constructing of ellipsoids around the examples (Fig. 2 e)).

The second group of classifiers calculates decision surface over the n dimensional feature space by adding one more dimension and create dimension $n+1$. A good example of such a classifier is Membership Function (MF) ARTMAP approach introduced in [1] and reviewed in [14] (Fig. 3).

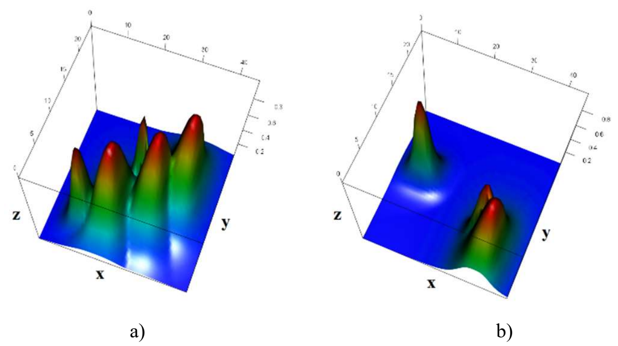


Fig. 3 Decision surface of dimension $n+1$ is referring to Fig. 2, which represents $n = 2$ dimensional feature space created by MF ARTMAP. Axis x and y represent dimensions of the feature space; axis z denotes the value of membership. Surface for class "+" is visualized on a), for class "-" on b)

MF ARTMAP assumes that data in feature space are organized in fuzzy clusters. The fuzzy cluster is considered as a fuzzy set A in multidimensional feature space representing a set of n_p ordered couples, e.g.:

$$A \in \left\{ [\overline{x}_1, \mu_A(\overline{x}_1)], \dots, \left[\overline{x}_{n_p}, \mu_A\left(\overline{x}_{n_p}\right)\right] \right\}, \qquad (1)$$

where A is a fuzzy set and $\left\{[\overline{x}_1, \mu_A(\overline{x}_1)]\right\}$ are ordered couples, $\overline{x} = [x_1, \dots, x_n]^T$ being a point in n dimensional feature space and $\mu_A(\overline{x})$ a value of the membership function of $\overline{x}$ to the fuzzy cluster (set) A, while n_p is the number of pattern. There are many fuzzy clusters in the feature space and a certain set of fuzzy cluster creates fuzzy class C_i where $i = 1 \dots n_c$ in which n_c is the number of classes defined in the training set. Fuzzy class is the union of fuzzy clusters belonging to a considered class, e.g.:

$$C_i = \left\{ \bigcup_{j=1}^{m_i} A_j^i \right\}, \qquad (2)$$

where A_j^i is the j-th fuzzy set belonging to i-th class. Therefore, the relation between $\mu_{C_i}(\overline{x})$ and $\mu_A(\overline{x})$ is as follows:

$$\mu_{C_i}(\overline{x}) = \max_j \left(\mu_{A_j^i}(\overline{x})\right), \quad (3)$$

where A_j^i is a fuzzy cluster belonging to class C_i and m_i is the number of fuzzy clusters creating a class C_i. Membership function is considered as:

$$\mu_A(\overline{x}) = \frac{1}{1 + \sum_{i=1}^{n} \left|\frac{x_i - x\ _i}{E_i}\right|^{F_i}}, \quad (4)$$

where n is the number of dimensions, x_i is a value of the i-th feature of the input vector $\overline{x}$ and xs_i, E_i and F_i are parameters of the fuzzy cluster for the i-th dimension. The output from MF ARTMAP for an unknown pattern $\overline{x}$ is a vector of memberships $\overline{\mu}_C$. Then the decision rule for a winner class CL of the unknown pattern $\overline{x}$ is as follows:

$$\mathrm{CL}(\overline{x}) = C_{\underset{i}{\operatorname{argmax}}\left(\mu_{C_i}(\overline{x})\right)} \quad (5)$$

III. Proposed Method

A. Cumulative Fuzzy Class Membership Criterion

The method is based on the assumption that data in feature space are split into n_c classes C_i where $i = 1 \ldots n_c$. Each class C_i is defined with training patterns $\tilde{p}_j^i$ where $j = 1 \ldots N_i$ and N_i is the number of training patterns of the i-th class. Each training pattern defines a fuzzy class membership criterion $\kappa_{\tilde{p}}(\overline{x})$, which is considered as Cauchy-like bell shaped function:

$$\kappa_{\tilde{p}_j^i}(\overline{x}) = \frac{1}{1 + \left(\frac{\|\tilde{p}_j^i - \overline{x}\|}{E_i}\right)^{F_i}}, \quad (6)$$

where $\overline{x}$ is an input vector, $\|\tilde{p}_j^i - \overline{x}\|$ is Euclidian distance between j-th training pattern of the i-th class and input vector. E_i and F_i are set of parameters for the i-th class.

Then Cumulative Fuzzy Class Membership Criterion for class C_i is defined as follows:

$$\chi_{C_i}(\overline{x}) = \sum_{j=1}^{N_i} \kappa_{\tilde{p}_j^i}(\overline{x}), \quad (7)$$

where $\chi_{C_i}(\overline{x})$ is the value of CFCMC for an unknown pattern $\overline{x}$ to the class C_i. The output from proposed method for an unknown pattern $\overline{x}$ is a vector of values of CFCMC $\overline{\chi}$. Then the decision rule for winner class CL of the unknown pattern $\overline{x}$ is as follows:

$$\mathrm{CL}(\overline{x}) = C_{\underset{i}{\operatorname{argmax}}\left(\chi_{C_i}(\overline{x})\right)} \quad (8)$$

B. Comparison of Cumulative Fuzzy Class Membership Criterion and Membership Function surface

The main difference between CFCMC and Membership Function can be seen by comparing equations (3) and (7) for the computation of membership to one particular class. In MF ARTMAP, at first, membership of unknown pattern to each cluster of the same class is computed. The final membership value to class C_i is the highest one. In CFCMC, fuzzy class membership criterion $\kappa_{\tilde{p}}(\overline{x})$ of unknown pattern $\overline{x}$ for all training patterns are computed. However, in contrast with MF ARTMAP, the final value of criterion membership will be the sum of them. In Fig. 4 and Fig. 5, decision surfaces of Membership Function and CFCMC is shown, respectively. For this purpose, simple one dimensional data with two classes (red and blue color) with a small overlap between them was used (displayed in Fig. 4 and Fig. 5 on dotted line). Each class consists of four samples, which are displayed using a star symbol. As can be seen in Fig. 4, MF ARTMAP created two clusters, one for each class. Two samples were incorrectly classified (marked with arrows); one sample from the blue class was classified to the red class, and one sample belonging to the red class were classified to the blue class. Whereas in Fig. 5, CFCMC was able to adapt more precisely to given data, which caused a higher classification accuracy, because all samples were correctly classified despite of an overlap between classes. Thus, we assume that classifier built based on CFCMC should achieve higher classification accuracy than the one built based on Membership Function, because of its better adaptability and flexibility in forming the decision boundary.

If we consider a two-dimensional classification situation according to the feature space in Fig. 2, we can visualize the CFCMC approach in three-dimensional space in the Fig. 6. We can see a comparison with Fig. 3 which represents an MF ARTMAP approach. It can be seen that the proposed CFCMC is also able to generate decision boundary with better generalization ability than MF ARTMAP.

Another difference between Membership Function and CFCMC is based on the domain of criterion membership values. The membership function is based on fuzzy theory. Therefore range of membership values of unknown vector $\overline{x}$ to class are from interval $< 0; 1 >$. The range of CFCMC does not have to fall on this interval because the value of criterion membership to class depends on all train patterns of the same class.

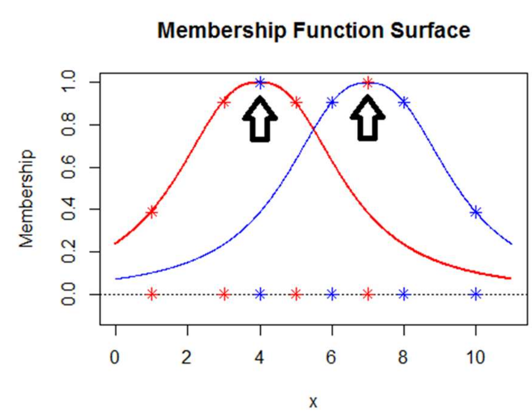


Fig. 4 Decision surface created by MF ARTMAP

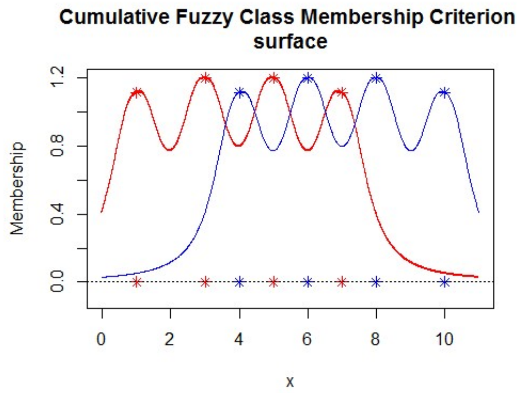
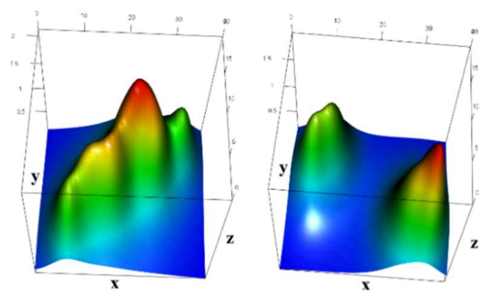


Fig. 5 Decision surface created by CFCMC



a) b)

Fig. 6 Decision surface of dimension *n+1* is referring to Fig. 2, which represents *n=2* dimensional feature space created by CFCMC. Axis *x* and *y* represent dimensions of the feature space; axis *z* denotes the value of cumulative fuzzy membership criterion of a point in the feature space determined by (x,y) to the class. Surface for class "+" is visualized on a), for class "-" on b)

IV. Algorithm Description

In this section, we describe the algorithm for finding the best shape of CFCMC surface for input data. The algorithm consists of two phases:

- Initialization phase – creating the CFCMC surface
- Learning phase – adjusting the shape of the CFCMC surface

A. Initialization phase

Initialization phase starts with splitting input data into three sets: a training set, a validation set, and a test set. Each set plays a specific role in one of each phase. The role of the training set is to create CFCMC decision surface. By validation set, during the learning phase, this surface is stretched (extended) to cover all validation pattern. Testing set is used for the final evaluation of the created decision surface.

Afterward, the values of parameters E and F for each class for equation (6) are initialized. These parameters affect the shape of the fuzzy class membership criterion $\kappa_{\bar{p}}$. Every fuzzy class membership criterion $\kappa_{\bar{p}^i}$ of the i-th class has the same value of parameters E and F. Later, the threshold for the "not classified" patterns is initialized. If the value of criterion membership of an unknown pattern $\bar{x}$ is below the threshold, pattern is "not classified". Finally, the CFCMC surface is created using training set and initialized parameters. First, using each training pattern, fuzzy class membership criterion $\kappa_{\bar{p}}$ is computed using equation (6). The shape of CFCMC surface for each class is created applying equation (7).

B. Learning phase

Assume that training patterns of a given training set belong to n_c classes. This generates vector $\bar{p}$ of parameters for each class C_i, where $i = 1 \ldots n_c$,

$$\bar{p} = [E_1, F_1; E_2, F_2; \ldots; E_i, F_i; \ldots; E_{n_c}, F_{n_c}]. \tag{9}$$

The goal is to find optimal values of E and F parameters for each class. To reach this goal, it is possible to use any optimization algorithm, for instance, evolutionary or genetic algorithm, hill climbing methods or gradient methods, we decided to employ a well-known simulated annealing [15] because of its simplicity and implementation clarity.

Simulated annealing is suitable for utilization in the wide range of optimization problems; the only requirements is a definition of the cost function. We defined cost function as the classification accuracy of training and validation set and as a ratio of "not classified" pattern. The classification accuracy was computed as a Kappa coefficient [16] from the contingency table. The ratio of "not classified" patterns is the ratio of a number of "not classified" patterns to a number of all patterns from the corresponding set.

V. Experiments and Results

For experiments were used freely available benchmark data from UCI Machine Learning Repository. The aim was a comparative analysis of method developed during this research with a similar method, improved MF ARTMAP proposed in [14].

A. Experiments

In the experiments, 20% selections of the patterns of each data set are allocated for testing set for both methods. For MF ARTMAP, remaining 80% of selections form the training set and for CFCMC, 60% form training set and remaining 20% patterns form validation set. Parameters for MF ARTMAP was initialized as follows: E = 10; F = 6; recognition layer threshold ("not classified" pattern threshold) – 0.15; comparison layer threshold – 5.

Parameters for CFCMC was initialized in a different way. The value of E parameter was initialized as the average distance between each pattern and the closest one in the training set. F parameter was set to value 2.5. The threshold for "not classified" patterns was set to 15% of the highest criterion membership value χ.

During learning phase, for CFCMC approach, vector $\bar{p}$ from Eq. 9 for each data set was generated with initialized values and optimized with Simulated Annealing using cost function defined in section IV.B. MF ARTMAP neural network was trained by learning algorithm described in [11].

B. Results

Table I. presents the result from the experiments from testing data of selected benchmarks. Results show percentage of the correctly classified, incorrectly classified and "not classified" patterns.

The results show that proposed method achieve higher classification accuracy than MF ARTMAP with every used benchmark data set. We suppose that higher classification accuracy is caused by the ability of CFCMC creating much more complex decision surface and thus, it has a better adaptability and flexibility. Another improvement can be seen in the ratio of patterns, which were "not classified". This ratio is significantly lower in the CFCMC. This result is evidence of the better generalization ability of CFCMC.

TABLE I. COMPARISON OF CLASSIFICATION ACCURACY MF ARTMAP AND CFCMC

Dataset		***MF ARTMAP***	***CFCMC***
IRIS	***correct***	**98%**	**99%**
	incorrect	***2%***	***1%***
	not classified	***0%***	***0%***
BUPA	***correct***	**52%**	**64%**
	incorrect	***32%***	***32%***
	not classified	***16%***	***4%***
PIMA	***correct***	**60%**	**72%**
	incorrect	***11%***	***27%***
	not classified	***29%***	***1%***
CANCER	***correct***	**36%**	**94%**
	incorrect	***2%***	***5%***
	not classified	***62%***	***1%***
WINE	***correct***	**44%**	**82%**
	incorrect	***8%***	***18%***
	not classified	***48%***	***0%***

VI. DESCRIPTION OF IDENTIFIED ISSUES AND SOLUTIONS

In this section, we discuss some issues and their theoretical solutions and a possible usage of the proposed method.

A. *Loss of the surface flexibility*

The proposed method is based on the fact that every fuzzy class membership criterion $\kappa_{\tilde{p}^i}$ of the i-th class has the same set of parameters (values of parameters E and F). This fact can lead to loss of complexity, thus flexibility of the decision surface during optimization procedure. The problem is explained by one dimensional data with five training points shown as a red triangles, one test point shown as a green square and one validation point shown as a blue circle in Fig. 7. The decision surface for a given data before optimization procedure is shown in Fig. 7 also.

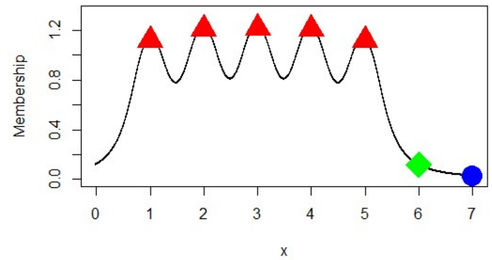


Fig. 7 Decision surface created by CFCMC for a given data before optimization

It can be seen that the criterion membership value of the validation pattern is significantly lower than the value of any training pattern. Therefore, validation pattern will be not classified to any class. Based on the cost function of the optimization algorithm, the goal is to increase the criterion membership value of validation point to prevent to be "not classified". The final shape of decision surface after optimization of the CFCMC with the same parameters for every fuzzy class membership criterion $\kappa_{\tilde{p}}$ is shown in Fig. 8. Although, the validation pattern can be classified to known class because of a sufficient value of the criterion membership, it is evident that the shape of decision surface loses its complexity because of its smooth shape, thus it loses its flexibility and adaptability.

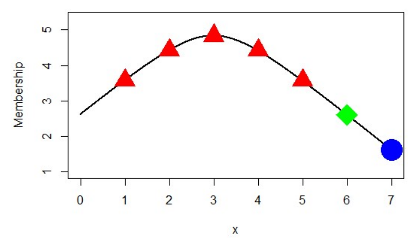


Fig. 8 Decision surface for a given data created by CFCMC with the same set of parameters for each fuzzy class membership criterion of the same class after optimization

One of the possible theoretical solutions is to increase the number of parameters. In contrast with current setting, it means that every fuzzy class membership criterion $\kappa_{\tilde{p}}$would have own sets of parameters. This setting allows to create even more complex decision surface, but for the price of creating a much larger vector of parameters to optimize, which may be difficult task for every optimization algorithm. Therefore, it is necessary to develop a more sophisticated method to find optimal parameters. To show the possibilities of own sets of parameters for every $\kappa_{\tilde{p}}$, we used the same data set as in Fig. 7. To increase the criterion membership value of validation point we changed the parameters of the $\kappa_{\tilde{p}}$ defined by the closest training pattern to the validation point only (training pattern with value of $x = 5$). Final shape of the decision surface is visualized in Fig. 9. It can be seen that the criterion membership value of validation pattern increased to a sufficient level without losing of shape complexity of decision surface in contrast with the case illustrated in Fig. 8. This issue is an open challenge for further improvement of the proposed approach as well as the next described issue.

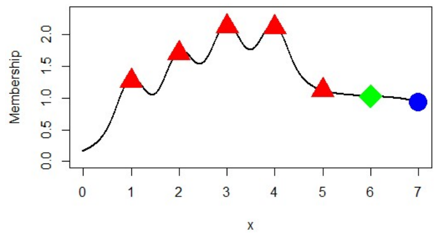


Fig. 9 Decision surface for a given data created by CFCMC with different set of parameters for each fuzzy class membership criterion of the same class after changing parameters of one $\kappa_{\tilde{p}}$ only

B. *Absence of an incremental learning*

The best advantage of MF ARTMAP is incremental learning ability due to the self-organization nature of the network structure. This ability means that if we have a classifier, which is able to classify data to, *e.g.* three classes, and we want to retrain this classifier for additional fourth class, we will not have to modify knowledge about previous three classes. Unfortunately, our proposed method is not capable of incremental learning. A possible solution would be to create set of experts, where each expert is expert to one specific class. If new class came to the training set, the new expert would be added, and only his knowledge would be modified.

C. *Semantic information extraction*

Knowledge of the shape of decision surface over feature space is a base for high-level processing, giving auxiliary output about data structures or similarities and dissimilarities between classes. The post-processing can offer information about one particular class (Intra-class knowledge), which usually express a homogeneity of the class. Contrariwise, the relation between several classes (Inter-class knowledge) extracts relation between classes using a computation of coverage between decision surfaces of classes [17].

VII. Conclusion and Future work

In this paper, we provide an introduction to classification algorithm as one of the principles of pattern recognition. We examined the various type of classifier with a focus on their effect on the feature space and compared the type and the shape of separating and decision surface, respectively. We have proposed a novel classification approach based on Cumulative Fuzzy Membership Function that creates a decision surface in a different way as an MF ARTMAP neural network. We called the proposed decision surface Cumulative Fuzzy Class Membership Criterion (CFCMC). We compared these two classification approaches. The analysis showed that CFCMC is more adaptable and flexible in forming a decision boundary than MF ARTMAP. After that we create the classifier built based on CFCMC that employ a well-known optimizing algorithm, Simulated Annealing, for optimizing the parameters of CFCMC. As experiments show, this approach provides us better classification accuracy than the MF ARTMAP.

For the future work, we raised some issues and their possible solutions, such as cases of losing the flexibility of the CFCMC or absence of an incremental learning. Furthermore, we mentioned possible usage of the proposed method. In fact, a clear advantage of the CFCMC, as well as MF ARTMAP, is that they are not "black-boxes" in a strictly way because they create created a decision surfaces that can offer semantic information about one particular class as well as relations among classes. This challenge should be objective of the future research.

Acknowledgment

This research work was supported by the Slovak Research and Development Agency under the contract No. APVV-015-0731, supported from 07-2016 to 06-2020.

References

[1] P. Sinčák et al., "Membership Function-ARTMAP Neural Networks," *TASK Q.*, vol. 7, no. 1, pp. 43–52, 2003.

[2] K. Fukunaga, "Introduction," in *Introduction to Statistical Pattern Recognition*, 1990, pp. 1–10.

[3] J. A. Iglesias and S. Drive, "Human activity recognition based on Evolving Fuzzy System," *Int. J. Neural Syst.*, vol. 20, no. 5, pp. 355–364, 2010.

[4] C. Pozna, "Signatures : Definitions , operators and applications to fuzzy modelling," *Fuzzy Sets Syst.*, vol. 201, pp. 86–104, 2012.

[5] J. Nowakov and M. Pr, "Medical Image Retrieval Using Vector Quantization and Fuzzy S-tree," *J. Med. Syst.*, 2017.

[6] J. Quinlan, "Induction of Decision Trees," *Expert Syst.*, pp. 81–106, 1987.

[7] D. T. Larose, "k-Nearest Neighbor Algorithm," *Discov. Knowl. Data An Introd. to Data Min.*, pp. 90–106, 2004.

[8] B. Widrob et al., "30 years of adaptive neural networks: perceptron, Madaline, and backpropagation," *Proc. IEEE*, vol. 78, no. 9, pp. 1415–1442, 1990.

[9] K. P. Murphy, "Naive Bayes classifiers Generative classifiers," *Bernoulli*, vol. 4701, no. October, pp. 1–8, 2006.

[10] M. Gardner and S. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," *Atmos. Environ.*, vol. 32, no. 14–15, pp. 2627–2636, 1998.

[11] G. A. Carpenter et al., "ARTMAP: supervised real-time learning and classification of nonstationary data by a self-organizing neural network," *IEEE Conf. Neural Networks Ocean Eng.*, vol. 4, pp. 565–588, 1991.

[12] G. A. Carpenter and S. Grossberg, *The Handbook of Brain Theory and Neural Networks - Adaptive Resonance Theory*, Second edi., no. 617. 2002.

[13] B. H. Settles, "Feature Spaces," *Univ. Wisconsin - Madison*, 2003.

[14] J. Ondo et al., "A Review of MF-ARTMAP Toward an Improvement Classification Accuracy using Simulated Annealing," *IEEE Conf. SMC 2016*, pp. 2038–2043, 2016.

[15] P. J. M. van Laarhoven and E. H. L. Aarts, *Simulated Annealing, Theory with Applications*. 1987.

[16] J. R. Landis et al., "The measurement of observer agreement for categorical data.," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.

[17] P. Smolár et al., "Making the Statements," in *Object Categorization with Artmap Neural Networks*, vol. 1, 2012, p. 189.