Outlines for a New Approach to Generating Fuzzy Classification Rules through Clustering Techniques

Dmitri A. Viattchenin
United Institute of Informatics Problems of the National Academy of Sciences of Belarus,
Surganov St. 6, 220012 Minsk, Belarus, e-mail: viattchenin@mail.ru

Abstract: *The paper deals with the method of extracting fuzzy classification rules based a heuristic method of possibilistic clustering. The description of basic concepts of the direct possibilistic clustering algorithm based on the concept of allotment among fuzzy cluster is provided. A general plan of the clustering procedure is given. A method of constructing of fuzzy rules based on clustering results is proposed. An illustrative example of the method's application to the Anderson's Iris data is carried out. An analysis of the experimental results is given and preliminary conclusions are formulated.*

Keywords**:** possibilistic clustering, fuzzy cluster, typical point, tolerance threshold, fuzzy rule

1. PRELIMINARIES

Any system can be described through the existing relations between its input variables and its output variables. To identify such relations, a functional input-output description may be not available in the case of complex processes. The use of fuzzy models has been shown to be successful. So, the problem of generation of fuzzy rules is one of more than important problems in the development of fuzzy models.

There are a number of approaches to learning fuzzy rules from data based on methods neural or evolutionary computation. Moreover, fuzzy classification rules can be obtained from fuzzy clustering results. In general, a fuzzy clustering algorithm aims at minimizing the objective function [1]

$$Q(P,\overline{\mathrm{T}}) = \sum_{l=1}^{c}\sum_{i=1}^{n} \upsilon_{li}^{\gamma} d(x_i,\overline{\tau}^l) \tag{1}$$

under the constraints

$$\sum_{i=1}^{n} \upsilon_{li} > 1,\ \forall i \in \{1,\ldots,n\}, \tag{2}$$

and

$$\sum_{l=1}^{c} \upsilon_{li} = 1,\ \forall l \in \{1,\ldots,c\}, \tag{3}$$

where $X = \{x_1,\ldots,x_n\} \subseteq \Re^m$ is the data set, c is the number of fuzzy clusters A^l, $l = 1,\ldots,c$ in the fuzzy c-partition P, $\upsilon_{li} \in [0,1]$ is the membership degree of object x_i to fuzzy cluster A^l, $\overline{\tau}^l \subseteq \Re^m$ is the prototype for fuzzy cluster A^l, $d(x_i,\overline{\tau}^l)$ is the distance between prototype $\overline{\tau}^l$ and object x_i, and the parameter $\gamma > 1$ is the fuzziness index. The selection of the value of γ determines whether the cluster tend to be more crisp or fuzzy. Membership degrees can be calculated as following

$$\upsilon_{li} = \frac{1}{\sum_{a=1}^{c}\left(\frac{d(x_i,\overline{\tau}^l)}{d(x_i,\overline{\tau}^a)}\right)^{1/(\gamma-1)}}, \tag{4}$$

and prototypes can be obtained from the formula

$$\overline{\tau}^l = \frac{\sum_{i=1}^{n}\upsilon_{li}^{\gamma}\cdot x_i}{\sum_{i=1}^{n}\upsilon_{li}^{\gamma}}. \tag{5}$$

Equations (4) and (5) are necessary conditions for (1) to have a local minimum. However, the condition (3) is hard from essential positions. So, a possibilistic approach to clustering was proposed in [2]. In particular, the objective function (1) is replaced by

$$Q(\mathrm{Y},\overline{\mathrm{T}}) = \sum_{l=1}^{c}\sum_{i=1}^{n}\left(\mu_{li}^{\psi} d(x_i,\overline{\tau}^l) + \eta_l (1-\mu_{li})^{\psi}\right) \tag{6}$$

under the constraint of possibilistic partition

$$\sum_{l=1}^{c} \mu_{li} > 1,\ \forall l \in \{1,\ldots,c\}. \tag{7}$$

where c is the number of fuzzy clusters A^l, $l = 1,\ldots,c$ in the possibilistic partition Y, $\mu_{li} \in [0,1]$ is the possibilistic memberships which are typicality degrees, $\overline{\tau}^l \subseteq \Re^m$ is the prototype for fuzzy cluster A^l, $d(x_i,\overline{\tau}^l)$ is the distance between prototype $\overline{\tau}^l$ and object x_i, and the parameter $\psi > 1$ is the analog of the fuzziness index.

Typicality degrees can be calculated as following

$$\mu_{li} = \frac{1}{1+\left(d(x_i,\overline{\tau}^l)/\eta_l\right)^{1/(\psi-1)}}, \tag{8}$$

and the parameters are estimated by

$$\eta_l = \frac{K}{\sum_{i=1}^{n}\upsilon_{li}^{\psi}}\sum_{i=1}^{n}\upsilon_{li}^{\psi} d(x_i,\overline{\tau}^l). \tag{9}$$

where $K = 1$.

The principal idea of extracting fuzzy classification rules based on fuzzy clustering is the following [3]. Each fuzzy cluster is assumed to be assigned to one class for classification and the membership grades of the data to the clusters determine the degree to which they can be classified as a member of the corresponding class. So, with a fuzzy cluster that is assigned to the some class we can associate a linguistic rule. The fuzzy cluster is projected into each single dimension leading to a fuzzy set on the real numbers. From a mathematical position the membership degree of the value $\hat{x}^t$ to the t th projection $\gamma_{B_l^t}(\hat{x}^t)$ of the fuzzy cluster A^l, $l \in \{1,\ldots,c\}$ is the supremum over the membership degrees of all vectors with $\hat{x}^t$ as t th component to the fuzzy cluster, i.e.

$$\gamma_{B_l^t}(\hat{x}^t) = \sup\left\{ \frac{1}{\sum_{a=1}^{c}\left(\dfrac{d(x_i,\bar{\tau}^l)}{d(x_i,\bar{\tau}^a)}\right)^{1/(\gamma-1)}} \middle| x_i = \right. , \qquad (10)$$
$$\left. = (\hat{x}_i^1,\ldots,\hat{x}_i^{t-1},\hat{x}_i^t,\hat{x}_i^{t+1},\ldots,\hat{x}_i^m) \in \Re^m \right\}$$

or

$$\gamma_{B_l^t}(\hat{x}^t) = \sup\left\{ \frac{1}{1+\left(d(x_i,\bar{\tau}^l)/\eta_l\right)^{1/(\psi-1)}} \middle| x_i = \right. \qquad (11)$$
$$\left. = (\hat{x}_i^1,\ldots,\hat{x}_i^{t-1},\hat{x}_i^t,\hat{x}_i^{t+1},\ldots,\hat{x}_i^m) \in \Re^m \right\}$$

in the possibilistic case [3]. An approximation of the fuzzy set by projecting only the data set and computing the convex hull of this projected fuzzy set or approximating it by a trapezoidal or triangular membership function is used for the rules obtaining [3].

Objective function-based fuzzy clustering algorithms are the most widespread methods in fuzzy clustering [1]. Objective function-based fuzzy clustering algorithms are sensitive to initial partition selection and fuzzy rules depend on the selection of the fuzzy clustering method. In particular, the GG-algorithm and the GK-algorithm of fuzzy clustering are recommended in [3] for fuzzy rules generation.

Heuristic algorithms of fuzzy clustering display low level of a complexity. An outline for a new heuristic method of possibilistic clustering was presented in [4], where a basic version of direct possibilistic clustering algorithm was described and the version of the algorithm is called the D-AFC(c)-algorithm [5]. The D-AFC(c)-algorithm can be considered as a direct algorithm of possibilistic clustering. The fact was demonstrated in [5].

The main goal of the paper is a preliminary consideration of the method of fuzzy rules extraction based on the clustering results obtained from the D-AFC(c)-algorithm. The contents of this paper is as follows: in the second section basic concepts of the possibilistic clustering method based on the concept of allotment among fuzzy clusters are outlined, in the third section a method of constructing of fuzzy rules from clustering results is proposed, in the fourth section an illustrative example of deriving of fuzzy rules from the Anderson's Iris data are given, in the fifth section the experimental results are discussed and preliminary conclusions are stated.

2. THE D-AFC(c)-ALGORITHM

Let us remind the basic concepts and the plan of the D-AFC(c)-algorithm. The structure of the set of objects can be described by some fuzzy tolerance, that is – a fuzzy binary symmetric reflexive intransitive relation. The notions of powerful fuzzy tolerance, feeble fuzzy tolerance and strict feeble fuzzy tolerance were considered in [3], as well. However, the essence of the method here considered does not depend on the kind of fuzzy tolerance. That is why the method herein is described for any fuzzy tolerance T.

Let $X = \{x_1,\ldots,x_n\}$ be the initial set of objects. Let T be a fuzzy tolerance on X with $\mu_{T_1}(x_i,x_j) \in [0,1], \forall x_i,x_j \in X$ being its membership function and α be the α-level value of T, $\alpha \in (0,1]$. Columns or lines of the fuzzy tolerance matrix are fuzzy sets $\{A^1,\ldots,A^n\}$. Let $\{A^1,\ldots,A^n\}$ be fuzzy sets on X, which are generated by a fuzzy tolerance T. The α-level fuzzy set $A_{(\alpha)}^l = \{(x_i,\mu_{A^l}(x_i)) \mid \mu_{A^l}(x_i) \geq \alpha,\ x_i \in X\}$ is fuzzy α-cluster or, simply, fuzzy cluster. So, $A_{(\alpha)}^l \subseteq A^l$, $\alpha \in (0,1]$, $A^l \in \{A^1,\ldots,A^n\}$ and μ_{li} is the membership degree of the element $x_i \in X$ for some fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1,n]$. Value of α is the tolerance threshold of fuzzy clusters elements.

The membership degree of the element $x_i \in X$ for some fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1,n]$ can be defined as a

$$\mu_{li} = \begin{cases} \mu_{A^l}(x_i), & x_i \in A_\alpha^l \\ 0, & otherwise \end{cases}, \qquad (12)$$

where an α-level $A_\alpha^l = \{x_i \in X \mid \mu_{A^l}(x_i) \geq \alpha\}$, $\alpha \in (0,1]$ of a fuzzy set A^l is the support of the fuzzy cluster $A_{(\alpha)}^l$.

The value of a membership function of each element of the fuzzy cluster is the degree of similarity of the object to some typical object of fuzzy cluster. Moreover, membership degree defines a possibility distribution function for some fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, and the possibility distribution function is denoted by $\pi_l(x_i)$. Notable that the number c of fuzzy clusters can be equal the number of objects, n.

Let T is a fuzzy tolerance on X, where X is the set of elements, and $\{A_{(\alpha)}^1,\ldots,A_{(\alpha)}^n\}$ is the family of fuzzy clusters for some α. The point $\tau_e^l \in A_\alpha^l$, for which

$$\tau_e^l = \arg\max_{x_i} \mu_{li},\ \forall x_i \in A_\alpha^l \qquad (13)$$

is called a typical point of the fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1,n]$. Obviously, a fuzzy cluster can have several typical points. That is why symbol e is the index of the typical point. A set $K(A_{(\alpha)}^l) = \{\tau_1^l,\ldots,\tau_{|l|}^l\}$ of typical points of the fuzzy cluster $A_{(\alpha)}^l$ is a kernel of the fuzzy cluster and $card\left(K(A_{(\alpha)}^l)\right) = |l|$ is a cardinality of the kernel. Obviously, if the fuzzy cluster have an unique typical point, then $\|l\| = 1$.

Let $R_z^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1,c},\, 2 \le c \le n\}$ be a family of fuzzy clusters for some value of tolerance threshold α, which are generated by some fuzzy tolerance T on the initial set of elements $X = \{x_1,\ldots,x_n\}$. If condition

$$\sum_{l=1}^{c} \mu_{li} > 0\,,\ \forall x_i \in X \qquad (14)$$

is met for all $A_{(\alpha)}^l$, $l = \overline{1,c}$, $c \le n$, then the family is the allotment of elements of the set $X = \{x_1,\ldots,x_n\}$ among fuzzy clusters $\{A_{(\alpha)}^l,\, l = \overline{1,c},\, 2 \le c \le n\}$ for some value of the tolerance threshold α. It should be noted that several allotments $R_z^\alpha(X)$ can exist for some tolerance threshold α. That is why symbol z is the index of an allotment.

The allotment among fuzzy clusters can be considered as the possibilistic partition and fuzzy clusters in the sense of the expression (12) are elements of the possibilistic partition. However, the concept of allotment will be used in further considerations. The next concept introduced should be paid attention to, as well.

Allotment $R_I^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1,n},\, \alpha \in (0,1]\}$ of the set of objects among n fuzzy clusters for some threshold α is the initial allotment of the set $X = \{x_1,\ldots,x_n\}$. In other words, if initial data are represented by a matrix of some fuzzy T then lines or columns of the matrix are fuzzy sets $A^l \subseteq X$, $l = \overline{1,n}$ and level fuzzy sets $A_{(\alpha)}^l$, $l = \overline{1,n}$, $\alpha \in (0,1]$ are fuzzy clusters. These fuzzy clusters constitute an initial allotment for some tolerance threshold and they can be considered as clustering components.

If some allotment $R_z^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1,c},\, c \le n\}$ corresponds to the formulation of a concrete problem, then this allotment is an adequate allotment. In particular, if condition

$$\sum_{l=1}^{c} card(A_\alpha^l) \ge card(X)\,,\ \forall A_{(\alpha)}^l \in R_z^\alpha(X)\,,$$
$$\alpha \in (0,1]\,,\ card(R_z^\alpha(X)) = c\,, \qquad (15)$$

and condition

$$card(A_\alpha^l \cap A_\alpha^m) = 0\,,\ \forall A_{(\alpha)}^l, A_{(\alpha)}^m\,,\ l \ne m\,, \qquad (16)$$

are met for all fuzzy clusters $A_{(\alpha)}^l$, $l = \overline{1,c}$ of some allotment $R_z^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1,c},\, c \le n\}$ then the allotment is the allotment among particularly separate fuzzy clusters and $w = \{0,\ldots,n\}$ is the maximum number of elements in the intersection area of different fuzzy clusters. If $w = 0$ in conditions (15) and (16) then the allotment is the allotment among fully separate fuzzy clusters.

The adequate allotment $R_z^\alpha(X)$ for some value of tolerance threshold $\alpha \in (0,1]$ is a family of fuzzy clusters which are elements of the initial allotment $R_I^\alpha(X)$ for the value of α and the family of fuzzy clusters should satisfy the conditions (15) and (16). Several adequate allotments can exist. Thus, the problem consists in the selection of the unique adequate allotment $R^*(X)$ from the set B of adequate allotments, $B = \{R_z^\alpha(X)\}$, which is the class of possible solutions of the concrete classification problem and $B = \{R_z^\alpha(X)\}$ depends on the parameters the classification problem. The selection of the unique adequate allotment $R^*(X)$ from the set $B = \{R_z^\alpha(X)\}$ of adequate allotments must be made on the basis of evaluation of allotments. The criterion

$$F(R_z^\alpha(X),\alpha) = \sum_{l=1}^{c} \frac{1}{n_l} \sum_{i=1}^{n_l} \mu_{li} - \alpha \cdot c\,, \qquad (17)$$

where c is the number of fuzzy clusters in the allotment $R_z^\alpha(X)$ and $n_l = card(A_\alpha^l), A_{(\alpha)}^l \in R_z^\alpha(X)$ is the number of elements in the support of the fuzzy cluster $A_{(\alpha)}^l$, can be used for evaluation of allotments.

Maximum of criterion (17) corresponds to the best allotment of objects among c fuzzy clusters. So, the classification problem can be characterized formally as determination of the solution $R^*(X)$ satisfying

$$R^*(X) = \arg \max_{R_z^\alpha(X) \in B} F(R_z^\alpha(X),\alpha)\,, \qquad (18)$$

where $B = \{R_z^\alpha(X)\}$ is the set of adequate allotments corresponding to the formulation of a concrete classification problem.

Thus, the problem of cluster analysis can be defined in general as the problem of discovering the unique allotment $R^*(X)$, resulting from the classification process and detection of fixed c number of fuzzy clusters can be considered as the aim of classification. So, the adequate allotment $R_z^\alpha(X)$ is any allotment among c fuzzy clusters in the case. There is a seven-step procedure of classification:

1. Calculate α-level values of the fuzzy tolerance T and construct the sequence $0 < \alpha_0 < \ldots < \alpha_\ell < \ldots < \alpha_Z \le 1$ of α-levels; let $\ell := 1$;
2. Construct the initial allotment $R_I^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1,n}\}$ for every value α_ℓ from the sequence $0 < \alpha_0 < \ldots < \alpha_\ell < \ldots < \alpha_Z \le 1$;
3. Let $w := 0$;
4. Construct allotments $R_z^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1,c}, c \le n\}$, $\alpha = \alpha_\ell$ which satisfy conditions (15) and (16) for every value α_ℓ from the sequence $0 < \alpha_0 < \ldots < \alpha_\ell < \ldots < \alpha_Z \le 1$;
5. Construct the class of possible solutions of the

classification problem $B(c)=\{R_z^\alpha(X)\}$, $\alpha \in \{\alpha_1,\dots,\alpha_Z\}$ for the given number of fuzzy clusters c and different values of the tolerance threshold α as follows:

if for some allotment $R_z^\alpha(X), \alpha \in \{\alpha_1,\dots,\alpha_Z\}$ the condition $card(R_z^\alpha(X))=c$ is met

then $R_z^\alpha(X)\in B(c)$

else let $w:=w+1$ and go to step 4;

6. Calculate the value of the criterion (17) for every allotment $R_z^\alpha(X)\in B(c)$;
7. The result $R^*(X)$ of classification is formed as follows:

   if for some unique allotment $R_z^\alpha(X)$ from the set $B(c)$ the condition (18) is met

   then the allotment is the result of classification; let $\ell := \ell + 1$ and go to step 2

   else the number c of classes is suboptimal.

The allotment $R^*(X)=\{A_{(\alpha)}^l \mid l=\overline{1,c}\}$ among the given number of fuzzy clusters and the corresponding value of tolerance threshold α are the results of classification.

3. A TECHNIQUE OF RULES GENERATION

Mamdani's [6] rule l within the fuzzy inference system is written as follows:

$$\begin{aligned}&\textit{If } \hat{x}^1 \textit{ is } B_l^1 \textit{ and } \dots \textit{and } \hat{x}^m \textit{ is } B_l^m \\ &\textit{then } y_1 \textit{ is } C_1^l \textit{ and} \dots \textit{and } y_c \textit{ is } C_c^l\end{aligned}, \qquad (19)$$

where $l=1,\dots,c$ is the number of rules and classes, m is the number of attributes, and B_l^t, $t\in\{1,\dots,m\}$ and C_l^l, $l\in\{1,\dots,c\}$ are fuzzy sets that define an input and output space partitioning. Let B_l^t be characterized by the membership function $\gamma_{B_l^t}(\hat{x}^t)$. The membership function can be triangular, Gaussian, trapezoidal, or any other type.

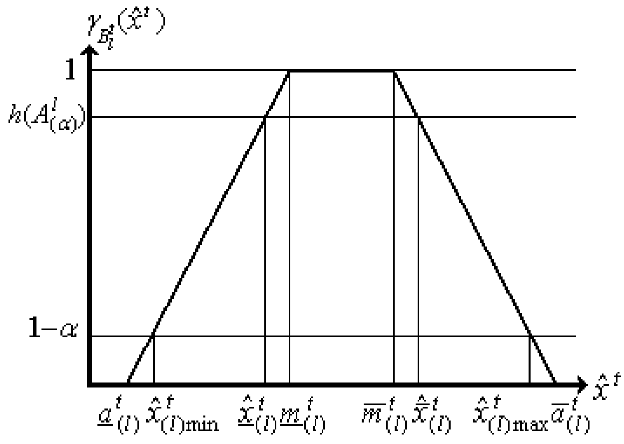


Fig. 1 – Trapezoidal fuzzy set defined by four parameters

In this paper, we consider trapezoidal and triangular membership functions. A trapezoidal fuzzy set is presented in Fig. 1 and the fuzzy set can be defined by four parameters, $B_l^t=(\underline{a}_{(l)}^t, \underline{m}_{(l)}^t, \overline{m}_{(l)}^t, \overline{a}_{(l)}^t)$. A triangular fuzzy set $B_l^t=(\underline{a}_{(l)}^t, m_{(l)}^t, \overline{a}_{(l)}^t)$ can be considered as a particular case of the trapezoidal fuzzy set where $\underline{m}_{(l)}^t=\overline{m}_{(l)}^t$.

The idea of deriving fuzzy rules from fuzzy clusters is the following. We apply the D-AFC(c)-algorithm to the given data and then obtain for each fuzzy cluster $A_{(\alpha)}^l$, $l\in\{1,\dots,c\}$ a kernel $K(A_{(\alpha)}^l)$ and a support A_α^l. The value of tolerance threshold $\alpha\in(0,1]$, which corresponds to the allotment $R^*(X)=\{A_{(\alpha)}^1,\dots,A_{(\alpha)}^c\}$, is the additional result of classification. We calculate the interval $[\hat{x}_{(l)\min}^t, \hat{x}_{(l)\max}^t]$ of values of every attribute $\hat{x}^t$, $t\in\{1,\dots,m\}$ for the support A_α^l. The value $\hat{x}_{(l)\min}^t$, $t\in\{1,\dots,m\}$ can be obtained as follows

$$\hat{x}_{(l)\min}^t=\min_{x_i\in A_\alpha^l}\hat{x}^t,\ \forall t\in\{1,\dots,m\},\ \forall l\in\{1,\dots,c\}, \qquad (20)$$

and the value $\hat{x}_{(l)\max}^t$, $t\in\{1,\dots,m\}$ can be calculated using a formula

$$\hat{x}_{(l)\max}^t=\max_{x_i\in A_\alpha^l}\hat{x}^t,\ \forall t\in\{1,\dots,m\},\ \forall l\in\{1,\dots,c\}. \qquad (21)$$

The parameter $\underline{a}_{(l)}^t$ can be obtained as following

$$\gamma_{B_l^t}(\hat{x}_{(l)\min}^t)=(1-\alpha),\ \gamma_{B_l^t}(\underline{a}_{(l)}^t)=0, \qquad (22)$$

and the parameter $\overline{a}_{(l)}^t$ can be obtained from the conditions

$$\gamma_{B_l^t}(\hat{x}_{(l)\max}^t)=(1-\alpha),\ \gamma_{B_l^t}(\overline{a}_{(l)}^t)=0. \qquad (23)$$

We calculate the value $\underline{\hat{x}}_{(l)}^t$ for all typical points $\tau_e^l\in K(A_{(\alpha)}^l)$ of the fuzzy cluster $A_{(\alpha)}^l$, $l\in\{1,\dots,c\}$ as follows:

$$\underline{\hat{x}}_{(l)}^t=\min_{\tau_e^l\in K(A_{(\alpha)}^l)}\hat{x}^t,\ \forall e\in\{1,\dots,|l|\}, \qquad (24)$$

and the value $\hat{\hat{x}}_{(l)}^t$ can be obtained from the equation

$$\hat{\hat{x}}_{(l)}^t=\max_{\tau_e^l\in K(A_{(\alpha)}^l)}\hat{x}^t,\ \forall e\in\{1,\dots,|l|\}, \qquad (25)$$

Thus, the parameter $\underline{m}_{(l)}^t$ can be calculated from the conditions

$$\gamma_{B_l^t}(\underline{\hat{x}}_{(l)}^t)=h(A_{(\alpha)}^l),\ \gamma_{B_l^t}(\underline{m}_{(l)}^t)=1, \qquad (26)$$

and the parameter $\overline{m}_{(l)}^t$ can be obtained as following

$$\gamma_{B_l^t}(\hat{\hat{x}}_{(l)}^t)=h(A_{(\alpha)}^l),\ \gamma_{B_l^t}(\overline{m}_{(l)}^t)=1, \qquad (27)$$

where $h(A^l_{(\alpha)}) = \sup_{x_i \in A^l_\alpha} \mu_{A^l_{(\alpha)}}(x_i)$ is the height of the $A^l_{(\alpha)}$.

Fuzzy sets C^l_l, $l = 1,\ldots,c$ can be defined on the interval of memberships $[0,1]$ and these fuzzy sets can be ordered as follows: $C^l_1 \prec C^l_2 \prec \ldots \prec C^l_c$. Fuzzy sets C^l_1 and C^l_c can be labeled as a low membership and a high membership, and $\gamma_{C^l_1}(y_1) = 1$ for $y_1 \in [0,(1-\alpha)]$ and $\gamma_{C^l_c}(y_c) = 1$ for $y_c \in [\alpha,1]$. Membership functions of other $(c-2)$ fuzzy sets $C^l_2,\ldots,C^l_{c-1}$ can be constructed as follows: the interval $[(1-\alpha),\alpha]$ must be divided into $(c-1)$ equal subintervals and parameters $\underline{a}^l_{(l)}, m^l_{(l)}, \overline{a}^l_{(l)}$, $l \in \{2,\ldots,c-1\}$ of triangular fuzzy sets $C^l_2,\ldots,C^l_{c-1}$ can be determined.

4. AN ILLUSTRATIVE EXAMPLE

The Anderson's Iris data set consists of the sepal length, sepal width, petal length and petal width for 150 irises [7]. The Anderson's Iris data forms the matrix of attributes $X_{4\times150} = [\hat{x}^t_i]$, $i = 1,\ldots,150$, $t = 1,\ldots,4$, where the sepal length is denoted by $\hat{x}^1$, sepal width is denoted by $\hat{x}^2$, petal length is denoted by $\hat{x}^3$ and petal width is denoted by $\hat{x}^4$.

Results of applications of the D-AFC(c)-algorithm to the Anderson's Iris data for different distance are summarized in Table 1.

Table 1. Results of applications of the D-AFC(c)-algorithm to the Anderson's Iris data for different distances

A distance	Main characteristics of the result of classification	
	The value α	Misclassifications
The normalized Hamming distance	0.8192	14
The normalized Euclidean distance	0.8104	6
The squared normalized Euclidean distance	0.9642	6

The squared normalized Euclidean distance was selected for the data preprocessing. By executing the D-AFC(c)-algorithm for three classes, we obtain that the allotment $R^*(X)$, which is corresponds to the result, was obtained for the tolerance threshold $\alpha = 0.9642$. Six misclassified objects were obtained. The ninety-fifth object is the typical point τ^1 of the fuzzy cluster which corresponds to the first class, the ninety-eighth object is the typical point τ^2 of the second fuzzy cluster, and the seventy-third object is the typical point τ^3 of the third fuzzy cluster. The height of each fuzzy cluster $A^l_{(\alpha)} \in R^*(X)$ is equal one. So, membership functions for fuzzy sets B^t_l and C^l_l, $t = 1,\ldots,4$, $l = 1,\ldots,3$, can be constructed immediately. Antecedents of fuzzy rules are presented on Fig. 2 and membership functions of fuzzy sets corresponding to outputs are presented on Fig. 3.

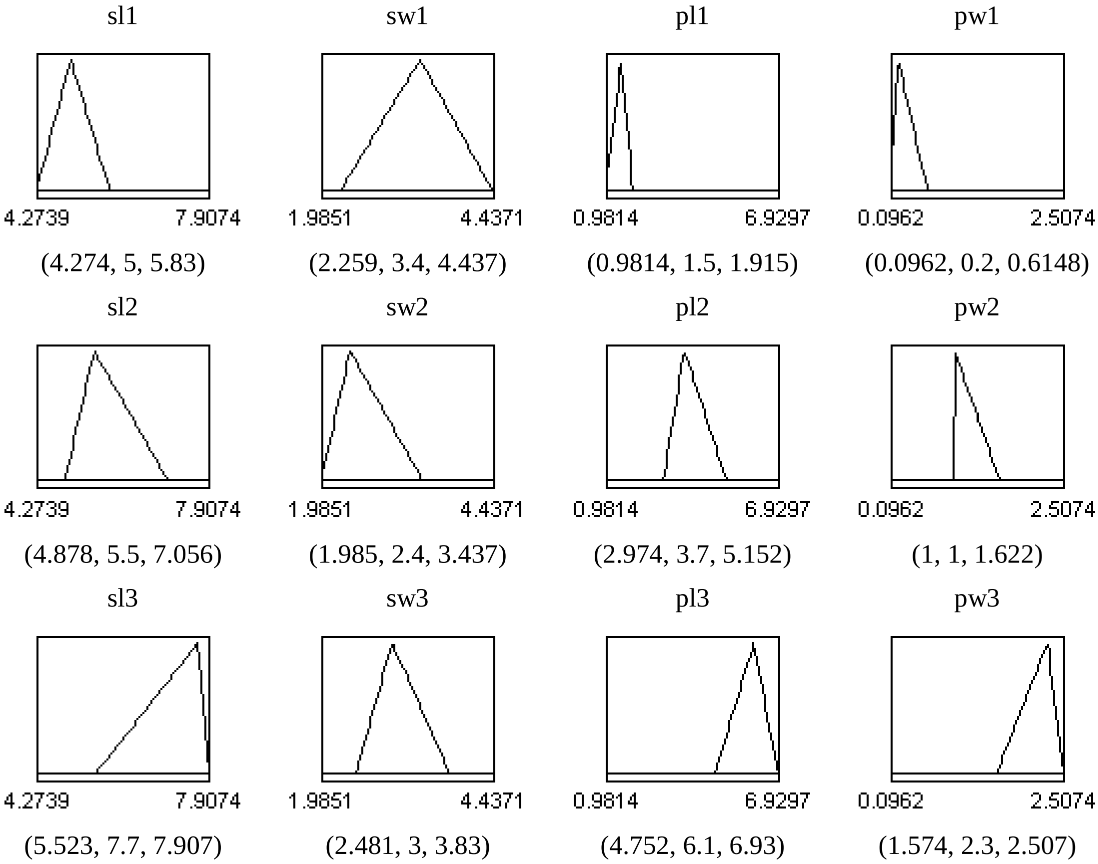
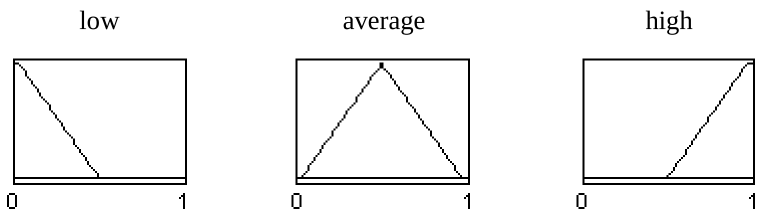


Fig. 2 – Fuzzy sets corresponding to input variables of fuzzy rules



(0, 0, 0.0358, 0.5) (0.0358, 0.5, 0.9642) (0.5, 0.9642, 1, 1)

Fig. 3 – Fuzzy sets corresponding to output variables of fuzzy rules

So, three fuzzy rules can be constructed. These fuzzy rules are presented in Table 2. The allotment is the allotment among fully separate fuzzy clusters. That is why a fuzzy set ‘high’ is used for output variables of fuzzy rules for corresponding classes.

Table 2. Fuzzy rules obtained from fuzzy clusters

Rule	Antecedents				Consequents		
	$\hat{x}^1$	$\hat{x}^2$	$\hat{x}^3$	$\hat{x}^4$	1	2	3
1	sl1	sw1	pl1	pw2	high		
2	sl2	sw2	pl2	pw2		high	
3	sl3	sw3	pl3	pw3			high

The Anderson’s data were classified using the Mamdani’s fuzzy inference system. The rules classify four objects incorrectly and two objects are rejected.

Notable that the fuzzy rules obtained using the D-AFC(c)-algorithm can be interpreted very simply, because membership functions of fuzzy sets which correspond to input variables of fuzzy rules has natural interpretations. So, the results obtained from the fuzzy inference system are corresponding to previous profound interpretation. Evidently, that the results are correlated with the results, obtained from the D-AFC(c)-algorithm.

From other hand, the results of classification are presented in [1] where eight fuzzy rules were generated from fuzzy clusters of the simplified version of the GG-algorithm of fuzzy clustering. The rules classify three objects incorrectly and three more were not classified at all.

5. CONCLUSIONS

The D-AFC(c)-algorithm is a precise and effective numerical procedure for solving classification problems. The results of application of the clustering method based on the allotment concept can be very well interpreted and the clustering results are stable because the D-AFC(c)-algorithm depending on the set $B(c)=\{R_z^{\alpha}(X)\}$ of possible solutions of the classification problem.

An approach to deriving fuzzy rules from fuzzy clusters obtained from the D-AFC(c)-algorithm is outlined in the paper. The approach is a way of a rapid prototyping of a fuzzy model using fuzzy clustering. In the context, the note rapid prototyping means being able to obtain an appropriate collection of fuzzy classification rules that can be considered as a first approximation to the fuzzy inference system.

The test of the fuzzy inference system based on the clustering of the Anderson’s Iris data using D-AFC(c)-algorithm shows an effectiveness of the proposed approach in comparison with the traditional approach based on the fuzzy clustering algorithms application. Moreover, the D-AFC(c)-algorithm can be applied to the three-way data clustering problem very simply [8]. So, the proposed approach can be extended for a case of the three-way data immediately.

ACKNOWLEDGEMENTS

I am grateful to Dr. Jan W. Owsinski, Prof. Janusz Kacprzyk and Prof. Frank Klawonn for their interest in the investigations and support.

REFERENCES

[1] F. Höppner, F. Klawonn, et al. *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition.* Wiley Intersciences. Chichester, 1999.

[2] R. Krishnapuram, J.M. Keller. A possibilistic approach to clustering, *IEEE Transactions on Fuzzy Systems* 1 (2) (1993) p. 98-110.

[3] F. Klawonn, R. Kruse. Constructing a fuzzy controller from data, *Fuzzy Sets and Systems* 85 (1) (1997). p. 177-193.

[4] D.A. Viattchenin. A new heuristic algorithm of fuzzy clustering, *Control & Cybernetics* 33 (2) (2004). p. 323-340.

[5] D.A. Viattchenin. A direct algorithm of possibilistic clustering with partial supervision, *Journal of Automation, Mobile Robotics and Intelligent Systems* 1 (3) (2007). p. 29-38.

[6] E.H. Mamdani, S. Assilian. An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies* 7 (1) (1975). p. 1-13

[7] E. Anderson. The irises of the Gaspe Peninsula, *Bulletin of the American Iris Society* 59 (1) (1934). p 2-5.

[8] D.A. Viattchenin. An outline for a heuristic approach to possibilistic clustering of the three-way data, *Journal of Uncertain Systems* 3 (1) (2009). p. 64-80.