

AUTOMATIC GENERATION OF FUZZY INFERENCE SYSTEMS USING HEURISTIC POSSIBILISTIC CLUSTERING

Received 10th February 2010; accepted 26th April 2010.

Dmitri A. Viattchenin

Abstract:

The interpretability and flexibility of fuzzy classification rules make them a popular basis for fuzzy controllers. Fuzzy control methods constitute a part of the areas of automation and robotics. The paper deals with the method of extracting fuzzy classification rules based on a heuristic method of possibilistic clustering. The description of basic concepts of the heuristic method of possibilistic clustering based on the allotment concept is provided. A general plan of the D-AFC(c)-algorithm is also given. A method of constructing and tuning 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. Introduction

Some remarks on fuzzy inference systems are considered in the first subsection. The second subsection includes a brief review of methods of extracting of fuzzy rules based on fuzzy clustering and the aims of the paper.

1.1. Preliminaries

Fuzzy inference systems are one of the most famous applications of fuzzy logic and fuzzy sets theory. They can be helpful to achieve classification tasks, process simulation and diagnosis, online decision support tools and process control. So, the problem of generation of fuzzy rules is one of more than important problems in the development of fuzzy inference systems.

There are a number of approaches to learning fuzzy rules from data based on techniques of evolutionary or neural computation, mostly aiming at optimizing parameters of fuzzy rules. From other hand, fuzzy clustering seems to be a very appealing method for learning fuzzy rules since there is a close and canonical connection between fuzzy clusters and fuzzy rules. The idea of deriving fuzzy classification rules from the data can be formulated as follows: the training data set is divided into homogeneous group and a fuzzy rule is associated to each group.

Fuzzy clustering procedures are exactly pursuing the strategy: a fuzzy cluster is represented by the cluster center and the membership degree of a datum to the cluster is decreasing with increasing distance to the cluster center.

So, each fuzzy rule from a fuzzy inference system can be characterized by a typical point and membership func-

tion that is decreasing with increasing distance to the typical point.

1.2. Fuzzy clustering and fuzzy rules

Let us consider some methods of fuzzy rules extracting from the data using fuzzy clustering algorithms. Some basic definitions must be given in the first place.

The training set contains n data pairs. Each pair is made of a m -dimensional input-vector and a c -dimensional output-vector. We assume that the number of rules in the fuzzy inference system rule base is c . So, Mamdani's [1] rule l within the fuzzy inference system is written as follows:

$$\begin{aligned} \text{If } \hat{x}^1 \text{ is } B_1^l \text{ and } \dots \text{ and } \hat{x}^m \text{ is } B_m^l \\ \text{then } y_1 \text{ is } C_1^l \text{ and } \dots \text{ and } y_c \text{ is } C_c^l \end{aligned} \quad (1)$$

where B_t^l , $t \in \{1, \dots, m\}$ and C_l^l , $l \in \{1, \dots, c\}$ are fuzzy sets that define an input and output space partitioning.

A fuzzy inference system, which is described by a set of fuzzy rules with the form (1) is the multiple inputs, multiple outputs system. Note that any fuzzy rule with the form (1) can be presented by c rules with the form of multiple inputs, single output:

$$\begin{aligned} \text{If } \hat{x}^1 \text{ is } B_1^l \text{ and } \dots \text{ and } \hat{x}^m \text{ is } B_m^l \text{ then } y_1 \text{ is } C_1^l \\ \dots \\ \text{If } \hat{x}^1 \text{ is } B_1^l \text{ and } \dots \text{ and } \hat{x}^m \text{ is } B_m^l \text{ then } y_c \text{ is } C_c^l \end{aligned} \quad (2)$$

Let B_t^l be characterized by the membership function $\gamma_{B_t^l}(\hat{x}^t)$. The membership function can be triangular, Gaussian, trapezoidal, or any other shape. In this paper, we consider trapezoidal and triangular membership functions.

Fuzzy classification rules can be obtained directly from fuzzy clustering results. In general, a fuzzy clustering algorithm aims at minimizing the objective function [2]

$$Q(P, \bar{T}) = \sum_{l=1}^c \sum_{i=1}^n v_{li}^\gamma d(x_i, \bar{t}^l), \quad (3)$$

under the constraints

$$\sum_{i=1}^n v_{li} > 1, \quad \forall i \in \{1, \dots, n\}, \quad (4)$$

and

$$\sum_{l=1}^c v_{li} = 1, \quad \forall l \in \{1, \dots, c\}, \quad (5)$$

where $X = \{x_1, \dots, x_n\} \subseteq \mathfrak{R}^m$ is the data set, c is the number of fuzzy clusters A^l , $l = 1, \dots, c$ in the fuzzy c -partition P , $v_{li} \in [0, 1]$ is the membership degree of object

x_i to fuzzy cluster A^l , $\bar{\tau}^l \subseteq \mathfrak{R}^m$ is the prototype for fuzzy cluster A^l , $d(x_i, \bar{\tau}^l)$ is the distance between prototype $\bar{\tau}^l$ 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

$$v_{li} = \frac{1}{\sum_{a=1}^c \left(\frac{d(x_i, \bar{\tau}^l)}{d(x_i, \bar{\tau}^a)} \right)^{1/(\gamma-1)}}, \quad (6)$$

and prototypes can be obtained from the formula

$$\bar{\tau}^l = \frac{\sum_{i=1}^n v_{li}^\gamma \cdot x_i}{\sum_{i=1}^n v_{li}^\gamma}. \quad (7)$$

Equations (6) and (7) are necessary conditions for (3) to have a local minimum. However, the condition (5) is hard from essential positions. So, a possibilistic approach to clustering was proposed in [3]. In particular, the objective function (3) is replaced by

$$Q(Y, \bar{T}) = \sum_{l=1}^c \sum_{i=1}^n (\mu_{li}^\psi d(x_i, \bar{\tau}^l) + \eta_l (1 - \mu_{li})^\psi), \quad (8)$$

under the constraint of possibilistic partition

$$\sum_{l=1}^c \mu_{li} > 1, \quad \forall l \in \{1, \dots, c\}, \quad (9)$$

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

$$\mu_{li} = \frac{1}{1 + (d(x_i, \bar{\tau}^l) / \eta_l)^{1/(\psi-1)}}, \quad (10)$$

and the parameters η_l , $l = 1, \dots, c$ are estimated by

$$\eta_l = \frac{K}{\sum_{i=1}^n v_{li}^\psi} \sum_{i=1}^n v_{li}^\psi d(x_i, \bar{\tau}^l), \quad (11)$$

where $K = 1$.

The principal idea of extracting fuzzy classification rules based on fuzzy clustering is the following [2]. 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}^l to the l th projection $\gamma_{B_i^l}(\hat{x}^l)$ of the fuzzy cluster A^l , $l \in \{1, \dots, c\}$ is the supremum over the membership degrees of all vectors with \hat{x}^l as l th component to the fuzzy cluster, i.e.

$$\gamma_{B_i^l}(\hat{x}^l) = \sup \left\{ 1 / \sum_{a=1}^c \left(\frac{d(x_i, \bar{\tau}^l)}{d(x_i, \bar{\tau}^a)} \right)^{1/(\gamma-1)} \mid x_i = (\hat{x}_i^1, \dots, \hat{x}_i^{l-1}, \hat{x}_i^l, \hat{x}_i^{l+1}, \dots, \hat{x}_i^m) \in \mathfrak{R}^m \right\}, \quad (12)$$

or

$$\gamma_{B_i^l}(\hat{x}^l) = \sup \left\{ 1 / \left(1 + (d(x_i, \bar{\tau}^l) / \eta_l)^{1/(\psi-1)} \right) \mid x_i = (\hat{x}_i^1, \dots, \hat{x}_i^{l-1}, \hat{x}_i^l, \hat{x}_i^{l+1}, \dots, \hat{x}_i^m) \in \mathfrak{R}^m \right\}, \quad (13)$$

in the possibilistic case. 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 [4].

Objective function-based fuzzy clustering algorithms are the most widespread methods in fuzzy clustering [2]. 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 [2] for fuzzy rules generation. All algorithms of possibilistic clustering are also objective functions-based algorithms.

Heuristic algorithms of clustering display low level of a complexity. An outline for a heuristic method of possibilistic clustering was presented in [5], where a basic version of direct possibilistic clustering algorithm was described and the version of the algorithm is called the D-AFC(c)-algorithm [6].

The main goal of the paper is a detail consideration of the method of the rapid prototyping fuzzy inference systems, which was outlined in [7]. The method is based on deriving fuzzy classification rules from the data on a basis of 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 and a plan of the D-AFC(c)-algorithm is given, in the third section a method of constructing of fuzzy rules is proposed, in the fourth section an illustrative example of deriving fuzzy rules from the Anderson's Iris data are given, in the fifth section preliminary conclusions are stated and some perspectives are outlined.

2. A heuristic method of possibilistic clustering

The basic concepts of the heuristic method of possibilistic clustering are considered in the first subsection. A plan of the direct clustering algorithm is given in the second subsection. The third subsection includes a review of methods of the data preprocessing.

2.1. Basic concepts

The D-AFC(c)-algorithm is based on a concept of an allotment of elements of the set of classified objects among fuzzy α -clusters. The allotment of elements of the set of objects among the fixed number c of fuzzy α -clusters can be considered as a special case of possibilistic partition. The fact was demonstrated in [6] and [8]. That is why the basic version of the algorithm, which is described in [5], can be considered as a direct algorithm of possibilistic clustering and the algorithm was called the D-AFC(c)-algorithm [6].

Let us remind the basic concepts of the heuristic method of possibilistic clustering. The concept of fuzzy

tolerance is the basis for the concept of fuzzy α -cluster. That is why definition of fuzzy tolerance must be considered in the first place.

Let $X = \{x_1, \dots, x_n\}$ be the initial set of elements and $T: X \times X \rightarrow [0,1]$ some binary fuzzy relation on X with $\mu_T(x_i, x_j) \in [0,1], \forall x_i, x_j \in X$ being its membership function. Fuzzy tolerance is the fuzzy binary intransitive relation, which possesses the symmetricity property

$$\mu_T(x_i, x_j) = \mu_T(x_j, x_i), \quad \forall x_i, x_j \in X, \quad (14)$$

and the reflexivity property

$$\mu_T(x_i, x_i) = 1, \quad \forall x_i \in X. \quad (15)$$

Let $X = \{x_1, \dots, x_n\}$ be the initial set of objects. Let T be a fuzzy tolerance on X and α be α -level value of $T, \alpha \in (0,1]$. Columns or lines of the fuzzy tolerance matrix are fuzzy sets $\{A^1, \dots, A^n\}$. Let $\{A^1, \dots, A^n\}$ be fuzzy sets on X , which are generated by a fuzzy tolerance T . The α -level fuzzy set $A^l_{(\alpha)} = \{(x_i, \mu_{A^l}(x_i)) | \mu_{A^l}(x_i) \geq \alpha\}, l \in [1, n]$ is fuzzy α -cluster or, simply, fuzzy cluster. So $A^l_{(\alpha)} \subseteq A^l, \alpha \in (0,1], A^l \in \{A^1, \dots, A^n\}$ and μ_{li} is the membership degree of the element $x_i \in X$ for some fuzzy cluster $A^l_{(\alpha)}, \alpha \in (0,1], l \in [1, n]$. 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 $A^l_{(\alpha)}, \alpha \in (0,1], l \in [1, n]$ can be defined as a

$$\mu_{li} = \begin{cases} \mu_{A^l}(x_i), & x_i \in A^l_{(\alpha)} \\ 0, & \text{otherwise} \end{cases}, \quad (16)$$

where an α -level $A^l_{(\alpha)} = \{x_i \in X | \mu_{A^l}(x_i) \geq \alpha\}, \alpha \in (0,1]$ of a fuzzy set A^l is the support of the fuzzy cluster $A^l_{(\alpha)}$. So, condition $A^l_{(\alpha)} = \text{Supp}(A^l_{(\alpha)})$ is met for each fuzzy cluster $A^l_{(\alpha)}, \alpha \in (0,1], l \in [1, n]$. Membership degree can be interpreted as a degree of typicality of an element to a fuzzy cluster. The value of a membership function of each element of the fuzzy cluster in the sense of (16) is the degree of similarity of the object to some typical object of fuzzy cluster. Membership degree defines a possibility distribution function for some fuzzy cluster $A^l_{(\alpha)}, \alpha \in (0,1]$. The fact was demonstrated in [8] and the possibility distribution function is denoted by $\pi_l(x_i)$.

Let T is a fuzzy tolerance on X , where X is the set of elements, and $\{A^1_{(\alpha)}, \dots, A^n_{(\alpha)}\}$ is the family of fuzzy clusters for some $\alpha \in (0,1]$. The point $\tau^l_e \in A^l_{(\alpha)}$, for which

$$\tau^l_e = \arg \max_{x_i} \mu_{li}, \quad \forall x_i \in A^l_{(\alpha)} \quad (17)$$

is called a typical point of the fuzzy cluster $A^l_{(\alpha)}, \alpha \in (0,1], l \in [1, n]$. A fuzzy cluster $A^l_{(\alpha)}$ can have several typical points. That is why symbol e is the index of the typical point. A set $K(A^l_{(\alpha)}) = \{\tau^l_1, \dots, \tau^l_{|l|}\}$ of typical points of the fuzzy cluster $A^l_{(\alpha)}$ is a kernel of the fuzzy cluster and $\text{card}(K(A^l_{(\alpha)})) = |l|$ is a cardinality of the kernel. Obviously, if the fuzzy cluster have a unique typical point, then $|l| = 1$.

Let $R^\alpha_z(X) = \{A^l_{(\alpha)} | l = 1, c, 2 \leq c \leq n, \alpha \in (0,1]\}$ be a family of fuzzy clusters for some value of tolerance threshold $\alpha, \alpha \in (0,1]$ which are generated by some fuzzy tolerance T on the initial set of elements $X = \{x_1, \dots, x_n\}$.

If condition

$$\sum_{l=1}^c \mu_{li} > 0, \quad \forall x_i \in X \quad (18)$$

is met for all fuzzy clusters $A^l_{(\alpha)} \in R^\alpha_z(X), l = \overline{1, c}, c \leq n$, then the family is the allotment of elements of the set $X = \{x_1, \dots, x_n\}$ among fuzzy clusters $\{A^l_{(\alpha)}, l = \overline{1, c}, 2 \leq c \leq n\}$ for some value of the tolerance threshold α .

It should be noted that several allotments $R^\alpha_z(X)$ could exist for some tolerance threshold α . That is why symbol z is the index of an allotment.

The condition (18) requires that every object $x_i, i = \overline{1, \dots, n}$ must be assigned to at least one fuzzy cluster $A^l_{(\alpha)}, l = \overline{1, c}, c \leq n$ with the membership degree higher than zero. The condition $2 \leq c \leq n$ requires that the number of fuzzy clusters in each allotment $R^\alpha_z(X)$ must be more than two.

Obviously, the definition of the allotment among fuzzy clusters (18) is similar to the definition of the possibilistic partition (9). So, the allotment among fuzzy clusters can be considered as the possibilistic partition and fuzzy clusters in the sense of (16) are elements of the possibilistic partition.

If condition

$$\sum_{l=1}^c \text{card}(A^l_{(\alpha)}) \geq \text{card}(X), \quad \forall A^l_{(\alpha)} \in R^\alpha_z(X), \quad (19)$$

where $\alpha \in (0,1], \text{card}(R^\alpha_z(X)) = c$, and condition

$$\text{card}(A^l_{(\alpha)} \cap A^m_{(\alpha)}) \leq w, \quad \forall A^l_{(\alpha)}, A^m_{(\alpha)}, l \neq m, \alpha \in (0,1], \quad (20)$$

are met for all fuzzy clusters $A^l_{(\alpha)}, l = \overline{1, c}$ of some allotment $R^\alpha_z(X) = \{A^l_{(\alpha)} | l = \overline{1, c}, c \leq n\}$ then the allotment is the allotment among particularly separate fuzzy clusters and $0 \leq w \leq n$ is the maximum number of elements in the intersection area of different fuzzy clusters. Obviously, if $w = 0$ in conditions (19) and (20) then the intersection area of any pair of different fuzzy cluster is an empty set and fuzzy clusters are fully separate fuzzy clusters.

The adequate allotment $R^\alpha_z(X)$ for some value of tolerance threshold $\alpha \in (0,1]$ is a family of fuzzy clusters which are elements of the initial allotment $R^l_\alpha(X)$ for the value of α and the family of fuzzy clusters should satisfy the conditions (19) and (20).

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, $B = \{R^\alpha_z(X)\}$, which is the class of possible solutions of the concrete classification problem. The set of adequate allotments is depending on the number of fuzzy clusters c in the sought allotment. So, $B(c) = \{R^\alpha_z(X)\}$ is the set of adequate allotments corresponding to the formulation of a classification problem.

The selection of the unique adequate allotment $R^*(X)$ from the set $B(c) = \{R^\alpha_z(X)\}$ of adequate allotments must be made on the basis of evaluation of allotments. The criterion

$$F(R^\alpha_z(X), \alpha) = \sum_{l=1}^c \frac{1}{n_l} \sum_{i=1}^{n_l} \mu_{li} - \alpha \cdot c, \quad (21)$$

where c is the number of fuzzy clusters in the allotment $R^\alpha_z(X)$ and $n_l = \text{card}(A^l_{(\alpha)}), A^l_{(\alpha)} \in R^\alpha_z(X)$ 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 (21) 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

$$R^*(X) = \arg \max_{R_z^\alpha(X) \in B(c)} F(R_z^\alpha(X), \alpha). \quad (22)$$

The adequate allotment $R_z^\alpha(X)$ is any allotment among c fuzzy clusters in the case. 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.

2.2. The D-AFC(c)-algorithm

Detection of fixed c number of fuzzy clusters can be considered as the aim of classification. There is a seven-step procedure of classification:

1. Calculate α -level values of the fuzzy tolerance T and construct the sequence $0 < \alpha_0 < \dots < \alpha_\ell < \dots < \alpha_Z \leq 1$ of α -levels; l
2. Construct the initial allotment $R_l^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, n}\}$ for every value α_ℓ from the sequence $0 < \alpha_0 < \dots < \alpha_\ell < \dots < \alpha_Z \leq 1$;
3. Let $w := 0$;
4. Construct allotments $R_z^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}\}$, $c \leq n$, $\alpha = \alpha_\ell$ which satisfy conditions (19) and (20) for every value α_ℓ from the sequence $0 < \alpha_0 < \dots < \alpha_\ell < \dots < \alpha_Z \leq 1$;
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 $\text{card}(R_z^\alpha(X)) = c$ 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 (21) 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 (22) is met,
then the allotment is the result of classification,
else the number c of classes is suboptimal.

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

Some modifications of the D-AFC(c)-algorithm are proposed in [6], [9] and [10].

2.3. Notes on the data preprocessing

Let us consider a method for the data preprocessing. The matrix of fuzzy tolerance $T = [\mu_T(x_i, x_j)]$, $i, j = \overline{1, n}$ is the matrix of initial data for the D-AFC(c)-algorithm of possibilistic clustering. However, the data can be presented as a matrix of attributes $\hat{X}_{n \times m} = [\hat{x}_i^t]$, $i = \overline{1, n}$, $t = \overline{1, m}$, where the value \hat{x}_i^t is the value of the t th attribute for i th object. In the first place, the data can be

normalized as follows:

$$x_i^t = \frac{\hat{x}_i^t}{\max_i \hat{x}_i^t}. \quad (23)$$

In the second place, the data can be normalized using a formula

$$x_i^t = \frac{\hat{x}_i^t - \min_i \hat{x}_i^t}{\max_i \hat{x}_i^t - \min_i \hat{x}_i^t}. \quad (24)$$

So, each object can be considered as a fuzzy set x_i , $i = \overline{1, n}$ and $x_i^t = \mu_{x_i}(x^t) \in [0, 1]$, $i = \overline{1, n}$, $t = \overline{1, m}$ are their membership functions.

The matrix of coefficients of pair wise dissimilarity between objects $I = [\mu_I(x_i, x_j)]$, $i, j = \overline{1, n}$ can be obtained after application of some distance to the matrix of normalized data $X_{n \times m} = [\mu_{x_i}(x^t)]$, $i = \overline{1, n}$, $t = \overline{1, m}$. The most widely used distances for fuzzy sets x_i , x_j , $i, j = \overline{1, n}$ in $X = \{x_1, \dots, x_n\}$ are [11]:

- The normalized Hamming distance:

$$l(x_i, x_j) = \frac{1}{m} \sum_{t=1}^m |\mu_{x_i}(x^t) - \mu_{x_j}(x^t)|, \quad i, j = \overline{1, n}, \quad (25)$$

- The normalized Euclidean distance:

$$e(x_i, x_j) = \sqrt{\frac{1}{m} \sum_{t=1}^m (\mu_{x_i}(x^t) - \mu_{x_j}(x^t))^2}, \quad i, j = \overline{1, n}, \quad (26)$$

- The squared normalized Euclidean distance:

$$\varepsilon(x_i, x_j) = \frac{1}{m} \sum_{t=1}^m (\mu_{x_i}(x^t) - \mu_{x_j}(x^t))^2, \quad i, j = \overline{1, n}. \quad (27)$$

The matrix of fuzzy tolerance $T = [\mu_T(x_i, x_j)]$, $i, j = \overline{1, n}$ can be obtained after application of complement operation

$$\mu_T(x_i, x_j) = 1 - \mu_I(x_i, x_j), \quad \forall i, j = \overline{1, n} \quad (28)$$

to the matrix of fuzzy intolerance $I = [\mu_I(x_i, x_j)]$, $i, j = \overline{1, n}$ obtained from previous operations.

3. Deriving fuzzy rules from fuzzy clusters

A technique of fuzzy rules antecedents learning is presented in the first subsection. A method of consequents learning is given in the second subsection of the section. The third subsection includes a technique of fuzzy rules tuning.

3.1. Antecedents learning

In the following, we will consider that the fuzzy inference system is a multiple inputs, multiple outputs system. The antecedent of a fuzzy rule in a fuzzy inference system defines a decision region in the m -dimensional feature space. Let us consider a fuzzy rule (1) where B_l^t , $t = \overline{1, m}$, $l \in \{1, \dots, c\}$ is a fuzzy set associate with the feature variable \hat{x}^t . Let B_l^t be characterized by the trapezoidal membership function $\gamma_{B_l^t}(\hat{x}^t)$, which is presented in Figure 1.

So, the fuzzy set B_l^t 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 [7]. We apply the D-AFC(c)-algorithm to the

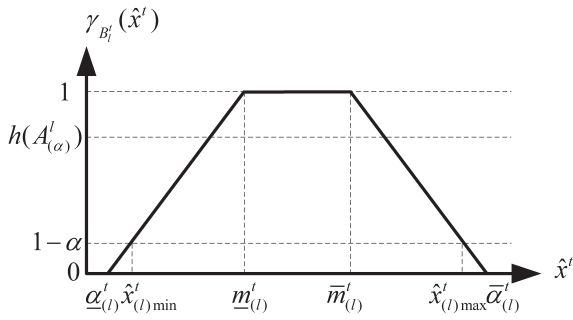


Fig. 1. A trapezoidal membership function for an antecedent fuzzy set.

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 $[\hat{x}_{(l)\min}^l, \hat{x}_{(l)\max}^l]$ of values of every attribute \hat{x}^l , $l \in \{1, \dots, m\}$, for the support A_{α}^l . The value $\hat{x}_{(l)\min}^l$ can be obtained as follows

$$\hat{x}_{(l)\min}^l = \min_{x_i \in A_{\alpha}^l} \hat{x}^l, \quad \forall l \in \{1, \dots, m\}, \quad \forall l \in \{1, \dots, c\}, \quad (29)$$

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

$$\hat{x}_{(l)\max}^l = \max_{x_i \in A_{\alpha}^l} \hat{x}^l, \quad \forall l \in \{1, \dots, m\}, \quad \forall l \in \{1, \dots, c\}. \quad (30)$$

The parameter $\underline{\alpha}^l$ can be obtained as following

$$\gamma_{B_i^l}(\hat{x}_{(l)\min}^l) = (1 - \alpha), \quad \gamma_{B_i^l}(\underline{\alpha}^l) = 0, \quad (31)$$

and the parameter $\bar{\alpha}^l$ can be obtained from the conditions

$$\gamma_{B_i^l}(\hat{x}_{(l)\max}^l) = (1 - \alpha), \quad \gamma_{B_i^l}(\bar{\alpha}^l) = 0. \quad (32)$$

We calculate the value $\hat{x}_{(l)}^l$ 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:

$$\hat{x}_{(l)}^l = \min_{\tau_e^l \in K(A_{(\alpha)}^l)} \hat{x}^l, \quad \forall e \in \{1, \dots, |I|\}, \quad (33)$$

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

$$\bar{\hat{x}}_{(l)}^l = \max_{\tau_e^l \in K(A_{(\alpha)}^l)} \hat{x}^l, \quad \forall e \in \{1, \dots, |I|\}. \quad (34)$$

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

$$\gamma_{B_i^l}(\hat{x}_{(l)}^l) = \gamma_{B_i^l}(\underline{m}_{(l)}^l) = 1, \quad (35)$$

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

$$\gamma_{B_i^l}(\bar{\hat{x}}_{(l)}^l) = \gamma_{B_i^l}(\bar{m}_{(l)}^l) = 1. \quad (36)$$

The height $h(A_{(\alpha)}^l) = \sup_{x_i \in A_{\alpha}^l} \mu_{A_{(\alpha)}^l}(x_i)$ of the fuzzy cluster

$A_{(\alpha)}^l$, $l \in \{1, \dots, c\}$, must be taken into account because the fuzzy cluster $A_{(\alpha)}^l \in R^*(X)$ can be a subnormal fuzzy set [12], [13].

So, condition $\hat{x}_{(l)}^l = \underline{m}_{(l)}^l$ and condition $\bar{\hat{x}}_{(l)}^l = \bar{m}_{(l)}^l$ are

met for all input variables \hat{x}^l , $l = 1, \dots, m$. Obviously, if condition $h(A_{(\alpha)}^l) = 1$ is met for the fuzzy cluster $A_{(\alpha)}^l$ and only one typical point is presented in the fuzzy cluster, then the condition $\hat{x}_{(l)}^l = \bar{\hat{x}}_{(l)}^l = \underline{m}_{(l)}^l = \bar{m}_{(l)}^l$ is met.

3.2. Consequents learning

The variables y_l , $l = 1, \dots, c$ are the consequents of fuzzy rules (1), represented by the fuzzy sets C_l^l , $l = 1, \dots, c$ with the membership functions $\gamma_{C_l^l}(y_l)$. Fuzzy sets C_l^l , $l = 1, \dots, c$ can be defined on the interval of memberships $[0, 1]$ and these fuzzy sets can be presented as follows: $C_l^l = (\alpha, \underline{\mu}_l, \bar{\mu}_l, 1)$, where α is the tolerance threshold, $\underline{\mu}_l = \min_{x_i \in A_{\alpha}^l} \mu_{li}$ and $\bar{\mu}_l = \max_{x_i \in A_{\alpha}^l} \mu_{li}$.

So, membership functions $\gamma_{C_l^l}(y_l)$ of fuzzy sets C_l^l , $l = 1, \dots, c$ will be trapezoidal membership functions. The situation is shown in Figure 2.

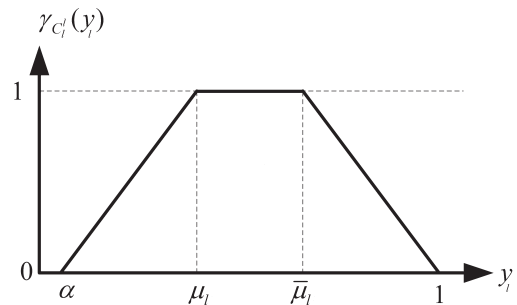


Fig. 2. A trapezoidal membership function for a consequent fuzzy set in a general case.

Fuzzy clusters can be subnormal fuzzy sets [12]. So, the case, which is presented in Figure 2, is the general case.

If the allotment $R^*(X)$ among fully separate fuzzy clusters is obtained and all fuzzy clusters $A_{(\alpha)}^l \in R^*(X)$, $l \in \{1, \dots, c\}$ are normal fuzzy sets then $h(A_{(\alpha)}^l) = 1$, $\forall A_{(\alpha)}^l \in R^*(X)$ and $\bar{\mu}_l = 1$ for each fuzzy cluster $A_{(\alpha)}^l$, $l \in \{1, \dots, c\}$. So, a trapezoidal membership function $\gamma_{C_l^l}(y_l)$ of a fuzzy set C_l^l , $l = 1, \dots, c$ in the case of normal and fully separate fuzzy clusters is presented in Figure 3.

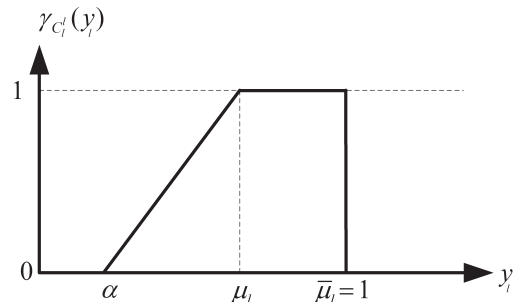


Fig. 3. A trapezoidal membership function for a consequent fuzzy set in a case of normal and fully separated fuzzy clusters.

Thus, trapezoidal membership functions $\gamma_{C_l^l}(y_l)$ for the fuzzy sets C_l^l , $l = 1, \dots, c$ can be constructed on a basis of the clustering results. The empty set $A_{\alpha}^l = \emptyset$, $l \in \{1, \dots, c\}$ can be correspond to some output variable y_l , $l \in \{1, \dots, c\}$. So, the empty fuzzy set C_l^l will be correspond to the output variable y_l , $l \in \{1, \dots, c\}$ and $\gamma_{C_l^l}(y_l) = 0$ is the membership function of the corresponding fuzzy set C_l^l .

If the allotment $R^*(X)$ among particularly separate fuzzy clusters is obtained then some non-empty fuzzy sets C_l^i will be correspond to some output variables y_l , $l = 1, \dots, c$.

3.3. Fuzzy rules tuning

The computational accuracy must be taken into account in the data processing by the D-AFC(c)-algorithm. The computational accuracy can be determined by a value of accuracy threshold $\varepsilon \in (0,1]$. If we decrease the value ε , the computational accuracy increases. Membership values μ_{ij} , $l=1, \dots, c$, $i=1, \dots, n$ the value of tolerance threshold α , and the number of typical points $|I|$ in each fuzzy cluster $A_{(\alpha)}^l \in R^*(X)$ are depending on the value of accuracy threshold ε . An allotment $R^*(X)$ can be characterized by the value of tolerance threshold $\alpha \in (0,1]$ that is increasing with decreasing accuracy, i.e., for $\varepsilon \rightarrow 1$ we have $\alpha \rightarrow 1$. From other hand, a fuzzy cluster $A_{(\alpha)}^l \in R^*(X)$ can be characterized by a kernel $K(A_{(\alpha)}^l)$ and the number of typical points of the fuzzy cluster is decreasing with increasing accuracy, i.e., for $\varepsilon \rightarrow 0$ we have $\text{card}(K(A_{(\alpha)}^l)) \rightarrow 1$. So, the accuracy threshold ε can be used as a parameter for the D-AFC(c)-algorithm. The fact was demonstrated in [14]. Moreover, the accuracy threshold ε can be considered as the analog of the fuzziness index γ in the formula (3).

Thus, the accuracy threshold ε can be useful for tuning of the rules. In particular, for $\varepsilon \rightarrow 0$ we have $(1-\alpha) \rightarrow 0$ and a crisp interval $[\underline{a}_{(l)}^i, \bar{a}_{(l)}^i]$ is increasing. Otherwise, if we decrease the accuracy threshold, $\varepsilon \rightarrow 1$, the number of typical points $|I|$ of the fuzzy cluster $A_{(\alpha)}^l \in R^*(X)$ increases, and a crisp interval $[\underline{m}_{(l)}^i, \bar{m}_{(l)}^i]$ increases.

Membership functions $\gamma_{C_l^i}(y_l)$ for consequents fuzzy sets C_l^i , $l=1, \dots, c$ depend on the value of accuracy threshold ε . For example, if we increase the value of accuracy threshold ε , the crisp interval $[\alpha, 1]$ decreases. Moreover, for $\varepsilon \rightarrow 1$ we have $\mu_{ij} \rightarrow 1$, $l=1, \dots, c$, $i=1, \dots, n$. That is why parameters $\underline{\mu}_i$ and $\bar{\mu}_i$ increases for all fuzzy sets C_l^i , i.e., for $\varepsilon \rightarrow 1$ we have $\underline{\mu}_i \rightarrow 1$ and $\bar{\mu}_i \rightarrow 1$, $\forall l=1, \dots, c$.

The proposed technique for fuzzy rules tuning will be explained by an illustrative example in the next section.

4. An illustrative example

The first subsection of the section includes the results of the Anderson's Iris data clustering by the D-AFC(c)-algorithm. The designed fuzzy inference system is presented in the second subsection and the results are compared with the results of other classifier systems.

4.1. Results of the Anderson's Iris data clustering

The Anderson's Iris data set consists of the sepal length, sepal width, petal length, and petal width measured for 150 irises [15]. The problem is to classify the plants into three subspecies on the basis of this information. The Anderson's Iris data forms the matrix of attributes $X_{4 \times 150} = [\hat{x}_i^t]$, $i=1, \dots, 150$, $t=1, \dots, 4$ where the sepal length is denoted - by \hat{x}^1 , sepal width - by \hat{x}^2 , petal length - by \hat{x}^3 and petal width - by \hat{x}^4 . The Iris database is the most known database to be found in the pattern recognition literature. The method of the data preprocessing which was described in the third section can be used for constructing the matrix of fuzzy tolerance and

the matrix of fuzzy tolerance can be processed by the D-AFC(c)-algorithm. The formula (23) and the squared normalized Euclidean distance (27) were used for the data preprocessing.

Four experiments were made for different values of the accuracy threshold ε . The allotment $R^*(X)$ among three fully separated fuzzy clusters was obtained in each experiment. The results of the Anderson's Iris data set processing by the D-AFC(c)-algorithm for different values of the accuracy threshold ε are presented in the Table 1.

Table 1. Results of the Anderson's Iris data set classification obtained from the D-AFC(c)-algorithm for different values of the accuracy threshold.

The value of accuracy threshold, ε	The value of tolerance threshold, α	The number of typical points		
		Class 1	Class 2	Class 3
$\varepsilon=0.01$	$\alpha=0.97$	42	15	12
$\varepsilon=0.001$	$\alpha=0.965$	7	2	1
$\varepsilon=0.0001$	$\alpha=0.9642$	1	1	1
$\varepsilon=0.00001$	$\alpha=0.96418$	1	1	1

By executing the D-AFC(c)-algorithm for three classes (1, 2, 3) in each experiment we obtain the following: the first class is formed by 50 elements all being Iris Setosa; the second class by 52 elements, 48 of them being Iris Versicolor and 4 Iris Virginica; the third class by 48 elements, 46 of them being Iris Virginica and 2 Iris Versicolor. In other words, the first class corresponds to the Setosa subspecies, the second class corresponds to the Versicolor subspecies and the third class corresponds to the Virginica subspecies. So, there are six mistakes of classification in each experiment.

4.2. A fuzzy inference system

Let us consider results of the experiment for the value of accuracy threshold $\varepsilon = 0.0001$. 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 $A_{(\alpha)}^l \in R^*(X)$ is equal one. So, membership functions $\gamma_{B_i^t}(\hat{x}^t)$ and $\gamma_{C_l^i}(y_l)$ for corresponding fuzzy sets B_i^t and

C_l^i , $t=1, \dots, 4$, $l=1, \dots, 3$, can be constructed immediately. The rule base induced by the D-AFC(c)-algorithm clustering result can be seen in Figure 4 where labels SLl , SWl , PLl , and SLl denote, respectively, sepal length, sepal width, petal length, and petal width, and $l=1, \dots, 3$ is the number of rule.

Note that only one typical point is presented in each fuzzy cluster. That is why membership functions $\gamma_{B_i^t}(\hat{x}^t)$, $t=1, \dots, 4$, $l=1, \dots, 3$ are triangular membership functions. Obviously that a meaningful linguistic label can be assigned to each fuzzy set B_i^t , $t=1, \dots, 4$, $l=1, \dots, 3$.

From other hand, linguistic labels Setosa, Versicolor and Virginica are associated with corresponding output variables y_l , $l=1, \dots, 3$. Note that fuzzy sets C_2^1 , C_3^1 , C_2^2 , C_3^2 , C_1^3 and C_2^3 are empty fuzzy sets.

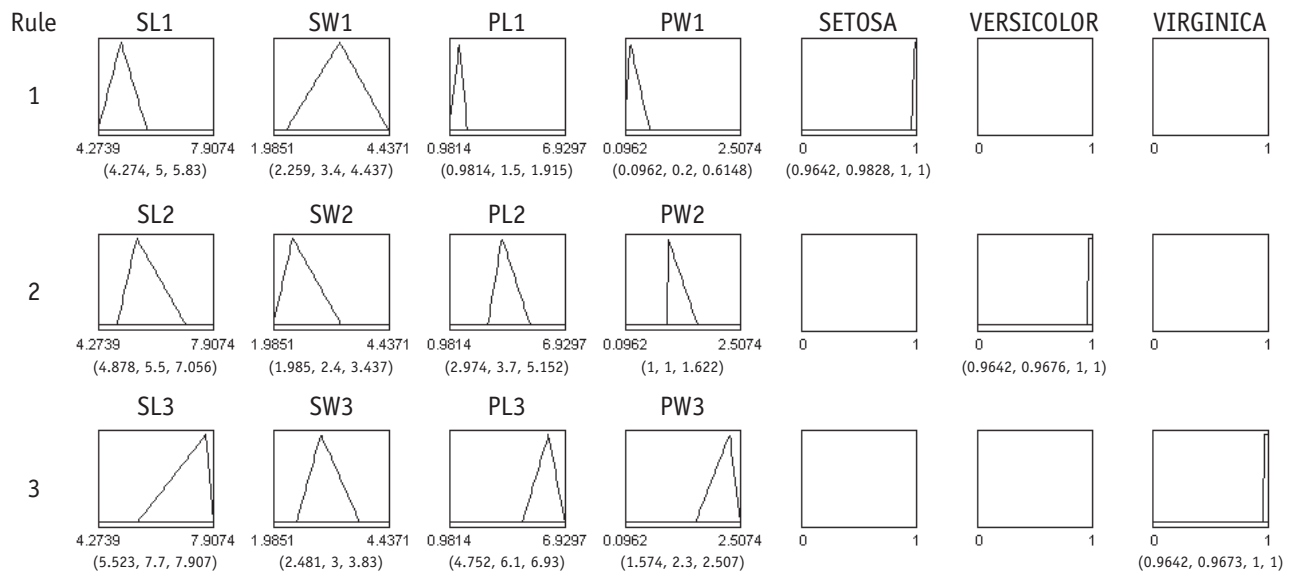


Fig. 4. The rule base induced by the clustering result.

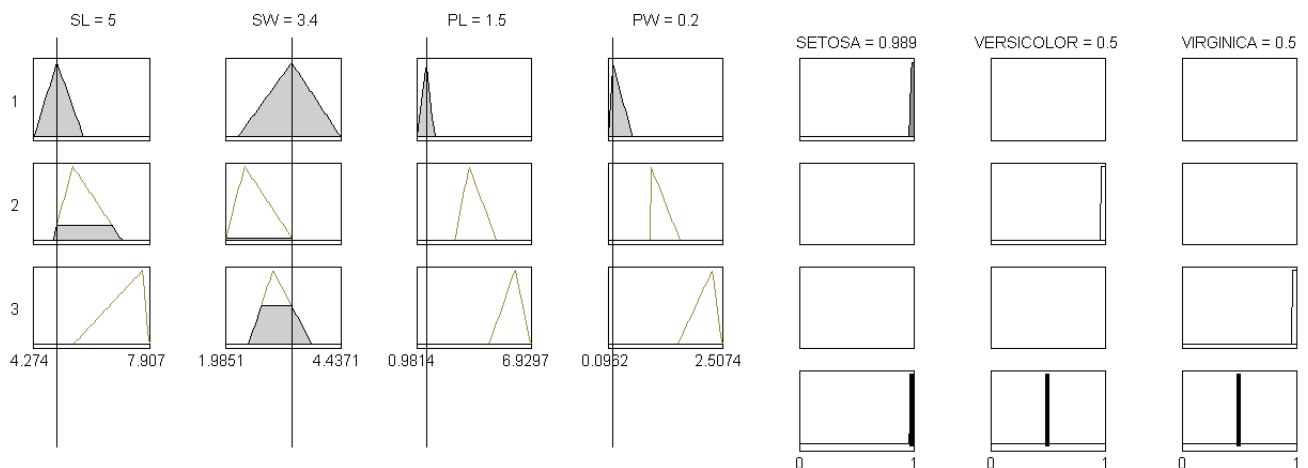


Fig. 5. The performance of the fuzzy inference system.

We show in Figure 5 a graph of the performance of the designed fuzzy inference system. The example of classification of the ninety-fifth object, which is the typical point of the first fuzzy cluster, is presented in Figure 5.

Total area is zero in the defuzzification procedure for output variables Versicolor and Virginica. That is why an average of the range of output variables Versicolor and Virginica are used as output values and these values are equal 0.5. The values can be interpreted as uncertain membership degrees.

The result, which is obtained from fuzzy inference system, is easily interpreted. Thus, the obtained model is suitable for interpretation since the rules consequents are the same or close to the actual class labels, such that each rule can be taken to describe all classes.

The Anderson's Iris data were classified using the constructed fuzzy inference system. The rules classify four objects incorrectly and two objects are rejected. Thus, the total number of misclassifications is 6. Evidently, that the results are correlated with the results, obtained from the D-AFC(c)-algorithm. So, the fuzzy inference system is accurate.

The application of the constructed fuzzy inference

system to the Anderson's Iris data was made in comparison with other approaches. Table 2 shows the results of some well-known classifier systems.

Table 2. Comparison of results of different classifier systems on the Anderson's Iris data set.

Authors	The number of rules	The number of misclassifications
Höppner et al. [2]	8	6
Roubos et al. [16]	3	4
Ishibuchi et al. [17]	5	3
Abonyi et al. [18]	3	6
Abe et al. [19]	3	2

For example, Höppner, Klawonn, Kruse and Runkler [2] applied the simplified version of the GG-algorithm of fuzzy clustering to learn a Mamdani-type fuzzy inference system for classifying the Anderson's Iris data by training on all 150 objects. An eight-rule fuzzy system was obtained. The rules classify 3 objects incorrectly and 3 more were not classified at all. So, the total number of misclassifications is 6.

From other hand, the FCM-algorithm of fuzzy cluste-

ring was applied by Roubos and Setnes [16] to obtain an initial Takagi-Sugeno model with singleton consequents. All 150 samples were used in the training process. So, the initial model with three rules was constructed from clustering results where each rule described a class. The classification accuracy of the initial model was rather discouraging, giving 33 misclassifications on the training data. A multi-objective genetic algorithm-based optimization approach was applied to the initial model. So, the number of misclassifications was reduced to 4 samples.

Ishibuchi, Nakashima and Murata [17] applied all 150 samples in the training process, and derived a fuzzy classifier with five rules. The resolution was 3 misclassifications.

Abyoni, Roubos and Szeifert [18] proposed a data-driven method to design compact fuzzy classifiers *via* combining a genetic algorithm, a decision-tree initialization, and a similarity-driven rule reduction technique. The final fuzzy inference system had three fuzzy rules and the number of misclassifications was 6.

A fuzzy classifier with ellipsoidal regions was proposed by Abe and Thawonmas [19]. They applied clustering methods to extract fuzzy classification rules, with one rule around cluster center, and then they tuned the slopes of their membership functions to obtain a high recognition rate. Finally, they obtained a three-rule fuzzy system with 2 misclassifications.

The results obtained from the constructed fuzzy inference system seem appropriate in comparison with the some well-known fuzzy systems. So, the proposed method of derivation of fuzzy classification rules from data can be considered as an effective technique of the rapid prototyping fuzzy inference systems.

5. Concluding remarks

Some conclusions are formulated in the first subsection. The second subsection deals with the perspectives on future investigations.

5.1. Discussion

Many techniques to design fuzzy inference systems from data are available; they all take advantage of the property of fuzzy inference systems to be universal approximators. This paper presents an automatic method to design fuzzy inference system for classification *via* heuristic possibilistic clustering and the method can be considered as an approach to rapid prototyping of fuzzy inference systems. The proposed method is simple in comparison with other well-known approaches. The results obtained with the proposed modeling approach for the Anderson's Iris data set case illustrate the effectiveness of the proposed method of designing fuzzy inference systems.

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.

5.2. Perspectives

Constructing a rule base from fuzzy clusters gives a first approximation for the data, which can be used as a basis for further improvements. A technique of fuzzy

rules tuning based on varying of the accuracy threshold is proposed in the paper. However, some other approaches, such as the genetic algorithm-based approach or neuro-fuzzy techniques can be used for fuzzy rules tuning.

Note that the computational complexity of the D-AFC(c)-algorithm is higher in comparison with objective function-based fuzzy clustering algorithms. For example, approximately 700 observations is the large data set for the D-AFC(c)-algorithm. Of course, the computational complexity of the D-AFC(c)-algorithm is the subject of special considerations. However, the clustering problem in cases of large data sets can be solved in the preliminary way as follows: the initial data set $X = \{x_1, \dots, x_n\}$ can be represented as a set $\tilde{X} = \{\tilde{x}_1, \dots, \tilde{x}_{\tilde{n}}\}$, where $\tilde{n} \ll n$ and each element $\tilde{x}_{\tilde{i}} \in \tilde{X}$, $\tilde{i} = 1, \dots, \tilde{n}$ is the subset of the data set X . So, the matrix of the reduced data set can be presented as the matrix $\tilde{X} = [\tilde{x}_{\tilde{i}}^{(t_{\min}, t_{\max})}]$ of the interval-valued data, and the data can be processed by the D-AFC(c)-algorithm [20]. Fuzzy rules can be extracted from the interval-valued data clustering results immediately.

From other hand, the D-AFC(c)-algorithm can be applied for classification the three-way data [13] and the fuzzy data [21]. So, the proposed method of designing fuzzy inference systems can be generalized for corresponding cases of the training data set.

These perspectives for investigations are of great interest both from the theoretical point of view and from the practical one as well.

ACKNOWLEDGMENTS

I am grateful to Prof. Janusz Kacprzyk, Prof. Jan W. Owsinski, Prof. Frank Klawonn and Prof. Valery Starovoitov for their interest in the investigations and support. I would like to thank Mr. Aliaksandr Damaratski for elaborating experimental software. I also thank the anonymous referees for their valuable comments.

AUTHOR

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

References

- [1] Mamdani E.H., Assilian S., "An experiment in linguistic synthesis with a fuzzy logic controller", *International Journal of Man-Machine Studies*, vol. 7, 1975, pp. 1-13.
- [2] Höppner F., Klawonn F., Kruse R., Runkler T., *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*, Chichester: Wiley Intersciences, 1999.
- [3] Krishnapuram R., Keller J.M., "A possibilistic approach to clustering", *IEEE Transactions on Fuzzy Systems*, vol. 1, 1993, pp. 98-110.
- [4] Sugeno M., Yasukawa T., "A fuzzy-logic-based approach to qualitative modeling", *IEEE Transactions on Fuzzy Systems*, vol. 1, 1993, pp. 7-31.
- [5] Viattchenin D.A., "A new heuristic algorithm of fuzzy

- clustering", *Control & Cybernetics*, vol. 33, 2004, pp. 323-340.
- [6] Viattchenin D.A., "A direct algorithm of possibilistic clustering with partial supervision", *Journal of Automation, Mobile Robotics and Intelligent Systems*, vol. 1, no.3, 2007, pp. 29-38.
- [7] Viattchenin D.A., "Outlines for a new approach to generating fuzzy classification rules through clustering techniques", *Proc. of the 10th International Conference on Pattern Recognition and Information Processing PRIP'2009*, Minsk, Belarus, 2009, pp. 82-87.
- [8] Viattchenin D.A., "On possibilistic interpretation of membership values in fuzzy clustering method based on the allotment concept", *Proceedings of the Institute of Modern Knowledge*, no. 3, 2008, pp. 85-90. (in Russian)
- [9] Viattchenin D.A., "Direct algorithms of fuzzy clustering based on the transitive closure operation and their application to outliers detection", *Artificial Intelligence*, no. 3, 2007, pp. 205-216. (in Russian)
- [10] Viattchenin D.A., "An algorithm for detecting the principal allotment among fuzzy clusters and its application as a technique of reduction of analyzed features space dimensionality", *Journal of Information and Organizational Sciences*, vol. 33, 2009, pp. 205-217.
- [11] Kaufmann A., *Introduction to the Theory of Fuzzy Subsets*, New York: Academic Press, 1975.
- [12] Viattchenin D.A., "Kinds of fuzzy α -clusters", *Proceedings of the Institute of Modern Knowledge*, no. 4, 2008, pp. 95-101. (in Russian)
- [13] Viattchenin D.A., "An outline for a heuristic approach to possibilistic clustering of the three-way data", *Journal of Uncertain Systems*, vol. 3, 2009, pp. 64-80.
- [14] Damaratski A., Novikau D., "On the computational accuracy of the heuristic method of possibilistic clustering", *Proc. of the 10th International Conference on Pattern Recognition and Information Processing PRIP'2009*, Minsk, Belarus, 2009, pp. 78-81.
- [15] Anderson E., "The irises of the Gaspé Peninsula", *Bulletin of the American Iris Society*, vol. 59, 1935, pp. 2-5.
- [16] Roubos H., Setnes M., "Compact and transparent fuzzy models and classifiers through iterative complexity reduction", *IEEE Transactions on Fuzzy Systems*, vol. 9, 2001, pp. 516-524.
- [17] Ishibuchi H., Nakashima T., Murata T., "Three-objective genetic-based machine learning for linguistic rule extraction", *Information Sciences*, vol. 136, 2001, pp. 109-133.
- [18] Abonyi J., Roubos J.A., Szeifert F., "Data-driven generation of compact, accurate and linguistically sound fuzzy classifiers based on a decision-tree initialization", *International Journal of Approximate Reasoning*, vol. 32, 2003, pp. 1-21.
- [19] Abe S., Thawonmas R., "A fuzzy classifier with ellipsoidal regions", *IEEE Transactions on Fuzzy Systems*, vol. 5, 1997, pp. 516-524.
- [20] Viattchenin D.A., Damaratski A. "Constructing of allotment among fuzzy clusters in case of quasi-robust cluster structure of set of objects", *Doklady BGUIR*, no. 1, 2010, pp. 46-52. (in Russian)
- [21] Viattchenin D.A., "A heuristic approach to possibilistic clustering for fuzzy data", *Journal of Information and Organizational Sciences*, vol. 32, 2008, pp. 149-163.