NEW APPROACH TO THE FUZZY CLASSIFICATION VIA AFS THEORY

XIANCHANG WANG, RUI DING, AND XIAODONG LIU

Abstract. In this paper, we propose a new classifier design based on fuzzy clustering approaches via AFS theory proposed by Ren Yan et. al. (2006). Firstly, the new fuzzy classifier, which is based on the fuzzy descriptions of the classes, is proposed. The classifier has two main advantages: One is that it can mimic the human reasoning processes and offer an interpretable classifier which is represented by some fuzzy sets with definitely semantic interpretations. Another is the data types of the attributes can be various types or sub-preference relations, even descriptions of human intuitions. Finally, Iris data-set is used to illustrate accuracy and stability of the new classifier. A number of illustrative examples show that this approach offers a far more flexible and effective means for the intelligent systems in real-world applications.

Key Words. AFS algebras, Molecular lattices, AFS structures, Fuzzy classifier.

1. Introduction

Classification of objects are important areas in a variety of fields, such as pattern recognition, artificial intelligence and vision analysis. Roughly speaking, we can classify the current approaches of classifications as five types: 1)Template matching [13, 14]. 2)Statistical classification [2]. 3)Syntactic or structured matching [12]. 4)Neural Networks [1, 6]. 5)Fuzzy classification. In this paper, we mainly focus on the fuzzy classification. Conventional non-fuzzy or crisp classification techniques assume that a pattern \( \alpha \) belongs to only one class. In real world applications, a pattern \( \alpha \) always belongs to more than one class at different degrees. It is natural to apply fuzzy set theory in cluster analysis and classification. In [29], the authors have proposed a new fuzzy clustering algorithm in the framework of AFS theory. Compared with the current fuzzy clustering algorithm, the new algorithm has the following advantages: 1. The attributes of objects can be various data types or sub-preference relations, even human intuition descriptions. 2. The distance function and objective function are not required, and the cluster number or the class label need not be given beforehand. 3. Each cluster is described by a fuzzy set in \( EM \), which is the AFS fuzzy logical compound of the simple concepts on some features with definitely semantic interpretation and determines the degree of each pattern belonging to this cluster.

In this paper, we propose a new classifier based on the improved fuzzy clustering analysis algorithm [30] and study the accuracy and stability of the new classifier by
the Iris dataset, which is the machine-learning database at University of California, Irvine via an anonymous ftp server (ftp://ftp.ics.uci.edu/pub/machine-learning-databases/). This new classifier is simply the fuzzy sets $\varsigma_{C_1}, \varsigma_{C_2}, \ldots, \varsigma_{C_l} \in EM$, where each $\varsigma_{C_i}$, $i = 1, 2, \ldots, l$ is the fuzzy description of class $i$. For each pattern $\alpha$, the degree of $\alpha$ belonging to fuzzy set $\varsigma_{C_i}$ is the degree of $\alpha$ belonging to the class $i$. $\varsigma_{C_i}$, the fuzzy description of class $i$, is learned from the training data by the improved fuzzy clustering algorithm\cite{30}. The accuracy and stability of the fuzzy description of class $i$, $\varsigma_{C_i}$, is studied by different partitions of Iris dataset. The examples show that the fuzzy descriptions $\varsigma_{C_i}$ are the representations of the constant models contained in the data. Compared to the current fuzzy classifiers mentioned above, it has the following advantages:

1. Instead for control problems and function approximation problems, the proposed fuzzy classifiers are designed for classification problems to determine the class of new pattern. The classifiers can be trained by using part of all samples as training data and the correct rate evaluated by testing samples.

2. The design of the proposed classifier, in which each class is represented by a fuzzy set in $EM$ which determines the degree of the new pattern belonging to the class, is simple, comprehensible and similar to human recognition habit without model but the fuzzy sets in $EM$, instead of fuzzy clustering model, genetic algorithm and T-S fuzzy model. Because each fuzzy set in $EM$ which represents a class has a definitive semantic signification in AFS fuzzy logic, the linguistic interpretation of the proposed classifiers is very comprehended. The fuzzy sets describing the classes are on some special features, in stead of all features, hence the proposed classifiers still can determine the class of a new pattern when the new pattern loses data of partial features.

3. The design of the proposed classifier is valid without assumptions of data set $X \subset R^{p \times n}$, since the AFS structure $(M, \tau, X)$ by which the membership functions of any fuzzy sets in $EM$ can be obtained, can be established by database with any types of attribute, even linguistic description based on human intuition.

4. Since the new classifier is made up of and amended by separably adding the information of each training sample, hence the classifier can be obtained by parallel processing and amended online.

The interested things are that the design of the new classifiers imitate the recognition process in which human classifies a set of objects by some given fuzzy concepts, attributes and features, therefore this research is an approach to knowledge representations and inference that is essential to any intelligence systems. It offers a far more flexible and powerful framework for representing human knowledge and studying large-scale intelligence systems in real world applications.

This paper is organized as follows: Section 2 introduces some basic concepts and presents several pertinent results on AFS theory. Section 3 presents framework of fuzzy classifiers based on Fuzzy Clustering Approaches via AFS Theory proposed by Ren Yan. Section 4 studies the accuracy and stability of the new classifier. Section 5 concludes this paper.

2. Preliminaries of the AFS theory

In this section, we recall the notations and definitions, present several pertinent results of AFS theory concerning the fuzzy clustering analysis and classification. We employ the notations, definitions and the symbols in \cite{29} in what follows. In essence, the AFS framework supports the studies on how to convert the information in the training examples or databases into the membership functions and their
fuzzy logic operations. AFS theory is made of AFS structures which is a special kind of combinatorics objects and AFS algebra which is a family of completely distributivity lattices. About the detail mathematical properties of AFS structure and AFS algebra, please see [18, 19, 26, 29].

In general, we explain the fuzzy sets and crisp subsets on $X$ as the followings:

For a fuzzy set $\zeta$ on universe of discourse $X$, any $x \in X$, either $x$ belongs to $\zeta$ at some degree or does not belong to $\zeta$ at all. While for a crisp subset $A$ of $X$, any $x \in X$, either $x$ belongs to $A$ or does not belong to $A$ at all.

Based on this statement, both a fuzzy set and a crisp subset on $X$ can be represented by a binary relation $R$ on $X$ through comparing the degrees of each pair of $x, y$ in $X$ belonging to the concept as described below.

**Definition 1.** ([29]) Let $\zeta$ be any concept on the universe of discourse $X$. $R_\zeta$ is called a binary relation (i.e., $R_\zeta \subset X \times X$) of $\zeta$ if $R_\zeta$ satisfies: $x, y \in X$, $(x, y) \in R_\zeta \Leftrightarrow x$ belongs to concept $\zeta$ at some degree and the degree of $x$ belonging to $\zeta$ is larger than or equal to that of $y$, or $x$ belongs to concept $\zeta$ at some degree and $y$ does not at all.

Although fuzzy concepts are ambiguous, the binary relations corresponding to them are crisp subset of $X \times X$. In real world applications, $R_\zeta$ can also be obtained by comparing the degrees of each pair objects belonging to concept $\zeta$ through human intuitions without necessary to represent them in $[0, 1]$ or a lattice in advance. We should notice that $(x, x) \in R_\zeta$ means that $x$ belongs to concept $\zeta$ at some degree and $(x, x) \notin R_\zeta$ means that $x$ does not belong to concept $\zeta$ at all. For example, let $X$ be a set of persons and $m$ be the simple concept “tall”, even if we do not know the exact height of $x$ and $y$, the degrees of $x$ and $y$ belonging to a simple concept $m$ can be compared. Now that each human concept corresponds a unique binary relation, how can we study human concepts by binary relations? First we should dealt with a class of simple concepts whose membership functions and logic operations are simple enough to be obtained, then the complex concepts are represented by these simple concepts.

**Definition 2.** ([29]) Let $X$ be a set and $R$ be a binary relation on $X$. $R$ is called a sub-preference relation on $X$ if for $x, y, z \in X, x \neq y$, $R$ satisfies the following conditions:

- $D2$-1. If $(x, y) \in R$, then $(x, x) \in R$;
- $D2$-2. If $(x, x) \in R$ and $(y, y) \notin R$, then $(x, y) \in R$;
- $D2$-3. If $(x, y), (y, z) \in R$, then $(x, z) \in R$;
- $D2$-4. If $(x, x) \in R$ and $(y, y) \in R$, then either $(x, y) \in R$ or $(y, x) \in R$.

A concept $\zeta$ is called a simple concept or simple attribute on $X$ if $R_\zeta$ is a sub-preference relation. Otherwise $\zeta$ is called a complex concept or a complex attribute on $X$.

**Definition 3.** ([19]) Let $X_1, ..., X_n, M$ be $n + 1$ non-empty sets. Then the set $EX_1,...,X_n,M^*$ is defined by

$$EX_1,...,X_n,M^* = \{ \sum_{i \in I} (u_{1i}...u_{ni}A_i) | A_i \in 2^M, u_{ri} \in 2^{X_r}, r = 1, ..., n, i \in I, I \text{ is a non-empty indexing set} \}.$$ 

In the case $n = 0$,

$$EM^* = \{ \sum_{i \in I} A_i | A_i \in 2^M, i \in I, I \text{ is a non-empty indexing set} \},$$

where the element $\sum_{i \in I} (u_{1i}...u_{ni}A_i)$ is composed of terms $(u_{1i}...u_{ni}A_i)$'s, $i \in I$, separated by “$+$”. $\sum_{i \in I} (u_{1p(i)}...u_{np(i)}A_{p(i)})$ and $\sum_{i \in I} (u_{1p(i)}...u_{np(i)}A_{p(i)})$ are the same.
element of $EX_1...X_n M^*$ if $p$ is a bijection from $I$ to $I$. When $I$ is finite, 
$\sum_{i=1}^{q}(u_{i1}...u_{in}A_i)$ is also denoted as $(u_{11}...u_{n1}A_1) + ... + (u_{1q}...u_{nq}A_q)$.

For a set, we know that the subsets of a set often contain or represent some useful information and knowledge. In real world applications, instead of one set, often many sets are involved and the information and knowledge represented by the subsets of different sets may have some kinds of relations. In order to study the complicated relations among the information and knowledge associated to different sets, we introduce the notation of $EX_1...X_n M^*$. Every element of $EX_1...X_n M^*$ is a “formal sum” of the terms constituted by the subsets of $X_1, X_2, ..., X_n, M$.

For $\gamma = \sum_{i\in I}(u_{i1}...u_{in}A_i) \in EX_1...X_n M^*$, $\gamma$ can be regarded as the “synthesis” of the information represented by all terms $u_{i1}...u_{in}A_i$’s. In practice, $M$ is a set of simple concepts, and $X_1, X_2, ..., X_n$ are the sets associated the concepts in $M$. For example, let $X$ be a set of persons and $M$ be a set of concepts such as “male”, “female”, “age”, “height”, “salary”, “hair black”, “hair white”, etc. For $\sum_{i\in I}(u_{i1}...u_{in}A_i) \in EX M^*$, every term $u_{Ai}, i \in I$, may mean that the persons in set $u_i \subset X$ satisfy some “condition” described by the attributes in $A_i \subset M$. AFS theory supports studies on how to convert the information represented by the elements of $EX_1...X_n M^*$ for the training examples and databases into the membership functions and their fuzzy logic operations.

**Definition 4.** ([19]) Let $X_1,...,X_n, M$ be $n+1$ non-empty sets. A binary relation $R$ on $EX_1...X_n M^*$ is defined as follows. For any $\sum_{i\in I}(u_{i1}...u_{in}A_i), \sum_{j\in J}(v_{j1}...v_{nj}B_j) \in EX_1...X_n M^*$,

$[\sum_{i\in I}(u_{i1}...u_{in}A_i)] R [\sum_{j\in J}(v_{j1}...v_{nj}B_j)] \iff$

(i) $\forall(u_{i1}...u_{in}A_i) (i \in I), \exists(v_{j1}...v_{nj}B_j) (j \in J)$ such that $A_i \supset B_h, u_{ri} \subseteq v_{rh}$, $1 \leq r \leq n$;

(ii) $\forall(v_{j1}...v_{nj}B_j) (j \in J), \exists(u_{1k}...u_{nk}A_k) (k \in I)$, such that $B_j \supset A_k, v_{rj} \subseteq u_{rk}$, $1 \leq r \leq n$.

**Proposition 1.** ([19]) Let $X_1,...,X_n, M$ be $n+1$ non-empty sets. If $A_i \subseteq A_t$, $u_{ri} \supseteq u_{rs}, r = 1, 2, ..., n$, $t, s \in I$, $t \neq s$, $\sum_{i\in I}(u_{i1}...u_{in}A_i) \in EX_1...X_n M^*$, then

$$\sum_{i\in I}(u_{i1}...u_{in}A_i) = \sum_{i \in I - \{s\}}(u_{i1}...u_{in}A_i).$$

**Theorem 1.** ([19]) Let $X_1,...,X_n, M$ be $n+1$ non-empty sets. Then $(EX_1...X_n M, \lor, \land)$ forms a completely distributive lattice under the binary compositions $\lor$ and $\land$ defined as follows. For any $\sum_{i\in I}(u_{i1}...u_{in}A_i), \sum_{j\in J}(v_{j1}...v_{nj}B_j) \in EX_1...X_n M^*$,

$$\sum_{i\in I}(u_{i1}...u_{in}A_i) \lor \sum_{j\in J}(v_{j1}...v_{nj}B_j) = \sum_{k\in I \cup J}(w_{1k}...w_{nk}C_k),$$

$$\sum_{i\in I}(u_{i1}...u_{in}A_i) \land \sum_{j\in J}(v_{j1}...v_{nj}B_j) = \sum_{i \in I, j \in J}[(u_{i1} \land v_{j1} ... u_{in} \land v_{nj})(A_i \cup B_j)],$$

where for any $k \in I \cup J$ (the disjoint union of $I$ and $J$), $C_k = A_k, w_{rk} = u_{rk}$ if $k \in I$, and $C_k = B_k, w_{rk} = v_{rk}$ if $k \in J, r = 1, 2, ..., n$.

$(EX_1...X_n M, \lor, \land)$ is called the $EI^{n+1}$ (expanding $n + 1$ sets $X_1,...,X_n, M$) algebra over $X_1,...,X_n$ and $M$. For $\alpha = \sum_{i\in I}(u_{i1}...u_{in}A_i), \beta = \sum_{j\in J}(v_{j1}...v_{nj}B_j) \in EX_1...X_n M, \alpha \leq \beta \iff \alpha \lor \beta = \beta \iff \forall(u_{i1}...u_{in}A_i) (i \in I), \exists(v_{j1}...v_{nj}B_j) (h \in J)$ such that $A_i \supset B_h, u_{ri} \subseteq v_{rh}, 1 \leq r \leq n$. 
Let \( s \) satisfies the following conditions:

\[
\sum_{i \in I} A_i \in EM.
\]

The elements of \( M \) are viewed as “elementary” (or “simple”) concepts. For any \( i \in I \), \( A_i \subset M \) represents conjunction of the concepts in \( A_i \), and \( \sum_{i \in I} A_i \) is the disjunction of the conjunctions represented by \( A_i \)'s (i.e., every element of \( EM \) corresponds to the disjunctive normal form of a formula representing a concept). For example, let \( M = \{ m_1, m_2, \ldots, m_8 \} \), where \( m_1 = \text{Color red} \), \( m_2 = \text{Color green} \), \( m_3 = \text{Color blue} \), \( m_4 = \text{Weight small} \), \( m_5 = \text{Weight medium} \), \( m_6 = \text{Weight large} \), \( m_7 = \text{Textured yes} \) and \( m_8 = \text{Textured no} \). Suppose the following fuzzy rules 1-4 describe a class of objects denoted by \( C \):

**Rule 1:** If \( x \) is [\text{Color is blue}], [\text{Textured is no}], then \( x \) belongs to \( C \).

**Rule 2:** If \( x \) is [\text{Color is red}], [\text{Weight is small}], then \( x \) belongs to \( C \).

**Rule 3:** If \( x \) is [\text{Color is red}], [\text{Weight is small}],[\text{Textured is no}], then \( x \) belongs to \( C \).

**Rule 4:** If \( x \) is [\text{Color is red}], [\text{Weight is large}], [\text{Textured is yes}], then \( x \) belongs to \( C \).

Rules 1-4 can be represented through a single rule by using the algebra operations \( \lor, \land \) in \( EM \), “if \( x \) is \( \gamma \), then \( x \) is \( C' \)”. where \( \gamma = \{ m_3, m_8 \} + \{ m_1, m_4 \} + \{ m_1, m_4, m_8 \} + \{ m_1, m_6, m_7 \} \in EM \) states that “\( m_3 \) and \( m_8 \)” or “\( m_1 \) and \( m_4 \)” or “\( m_1, m_4 \) and \( m_8 \)” or “\( m_1, m_4 \) and \( m_7 \)”.

Although \( M \) may be a set of fuzzy or crisp attributes, every element of \( EM \) has a well-defined meaning like the one we have discussed above. In virtue of (1) we know that

\[
\gamma = \{ m_3, m_8 \} + \{ m_1, m_4 \} + \{ m_1, m_6, m_7 \}
\]

This implies the left side and right side are equivalent as the conditions of the two rules. Considering the terms of \( \gamma \), any \( x \) which satisfies \( \{ m_1, m_4, m_8 \} \) also satisfies \( \{ m_1, m_4 \} \). Therefore, term \( \{ m_1, m_4, m_8 \} \) is redundant when forming the left side \( \gamma \) of the rule. Thus rules 1-4 can also be simply represented through a single rule by

\[
\text{If } x \text{ is } \{ m_3, m_8 \} + \{ m_1, m_4 \} + \{ m_1, m_6, m_7 \}, \text{ then } x \text{ is } C'.
\]

Since by Theorem 1, we have

\[
(\{ m_3, m_8 \} + \{ m_1, m_4 \} + \{ m_1, m_6, m_7 \}) \lor (\{ m_3, m_8 \} + \{ m_1, m_4, m_8 \} + \{ m_1, m_6, m_7 \}) = \{ m_3, m_8 \} + \{ m_1, m_4 \} + \{ m_1, m_6, m_7 \}
\]

Hence

\[
\{ m_3, m_8 \} + \{ m_1, m_4 \} + \{ m_1, m_6, m_7 \} \geq \{ m_3, m_8 \} + \{ m_1, m_4, m_8 \} + \{ m_1, m_6, m_7 \}
\]

in lattice \( EM \). This implies that the right side of the inequality is stricter than left side as the conditions of some rules. By (2) and (3), we observe that the operations \( \lor, \land \) of the elements of \( EM \) correspond to the “or”, “and” of the corresponding rules respectively.

In the following, we define an AFS structure, a triple \((M, \tau, X)\), which gives rise to the lattice representations of the membership degrees and fuzzy logic operations for each concepts in \( EM \). In general, where \( X \) is universe of discourse and \( M \) is a set of some simple concepts on \( X \). The map \( \tau : X \times X \to 2^M \) not only describes all binary relations corresponding to the (fuzzy or crisp) concepts in \( M \), but also constructs a combinatorics object [9].

**Definition 5.** ([18,19]) Let \( X, M \) be sets and \( 2^M \) be the power set of \( M \), \( \tau : X \times X \to 2^M \). \((M, \tau, X)\) is called an AFS structure if for any \( x_1, x_2, x_3 \in X \), \( \tau \) satisfies the following conditions:
Let \( X \) be a measurable space, where \( \tau \) is a measurable function and \( \sum x \) is defined by
\[
\tau(x_i, x_j) = \{ m \mid m \in M, (x_i, x_j) \in R_m \}, x_i, x_j \in X.
\]

**Theorem 2.** ([18]) Let \((M, \tau, X)\) be an AFS structure. For \( x \in X \), \( A \subseteq M \), we define the symbol
\[
I\{\{x\}\} = \{ y \mid y \in X, \tau(x, y) \supseteq A \}.
\]
For any given \( x \in X \), if we define \( \phi_x : EM \rightarrow EXM \) as follows. For any \( \sum_{i \in I} A_i \in EM \),
\[
\phi_x(\sum_{i \in I} A_i) = \sum_{i \in I} A_i(\{x\}) A_i \in EXM,
\]
then \( \phi_x \) is a lattice homomorphism from \((EM, \vee, \wedge)\) to \((EXM, \vee, \wedge)\).

Theorem 2 implies that for any given concept \( \sum_{i \in I} A_i \in EM \), we get a map \( \sum_{i \in I} A_i : X \rightarrow EXM, \forall x \in X \),
\[
(\sum_{i \in I} A_i)(x) = \sum_{i \in I} A_i(\{x\}) A_i \in EXM.
\]
Since \((EXM, \vee, \wedge)\) is a lattice, hence map \( \sum_{i \in I} A_i \) is a L-fuzzy set with membership degrees valued by \( EII \) algebra \( EXM \). For \( \alpha, \beta \in EM, \alpha \vee \beta \) and \( \alpha \wedge \beta \) are logic “or” and “and” of L-fuzzy sets \( \alpha \) and \( \beta \) respectively.

**Definition 6.** ([18]) Let \( X, M \) be sets, \((M, \tau, X)\) be an AFS structure and \((X, S, m)\) be a measurable space, where \( S \) is the \( \sigma \)-algebra on \( X \) and \( m \) be a finite and positive measure on \( S \) with \( 0 < m(X) < \infty \). Then \((M, \tau, X, S, m)\) is called a semi-congnitive field. For fuzzy concept \( \gamma = \sum_{i \in I} A_i \in EM \), if \( \forall i \in I, \forall x \in X, A_i(\{x\}) \in S \) (refer to (21)), then fuzzy concept \( \sum_{i \in I} A_i \) is called measurable in the semi-congnitive field \((M, \tau, X, S, m)\) and its membership function is defined as follows:
\[
\mu_\gamma(x) = \sup_{i \in I} \left\{ \frac{m(A_i(\{x\}))}{m(X)} \right\}.
\]

For \( M \) a given set of simple concepts on universe of discourse \( X \), using the AFS structure \((M, \tau, X)\) and the AFS algebras, we get the membership functions and their fuzzy logic operations for the fuzzy concepts in \( EM \) by (6). The lattice value membership degrees of the fuzzy concepts and their logical operations are determined by the original data and facts, in stead of triangular norm and human intuition. This is totally different from the other fuzzy theories. We should notice that the membership functions of the fuzzy concepts in \( EM \) can be obtained as long as the binary relations of the simple concepts in \( M \) are given, and the real number descriptions of the objects on each attribute are not necessary. In [28], the authors defined the logic operator \( ' \) (negation) as follows: For fuzzy concept \( \sum_{i \in I} A_i \in EM \),
where $a'$ is the negation of simple concept $a \in M$ and the algorithm of obtaining $a'$ can be found in [28]. The logic system $(EM, \lor, \land, \neg')$ is called AFS fuzzy logic.

3. The Fuzzy Classifiers Based on Fuzzy Clustering Analysis

In this section, we propose a new design of fuzzy classifiers based on the fuzzy descriptions of the clusters learned by the fuzzy clustering algorithm proposed in [30]. The well-known Iris dataset is provided by Fisher in 1936. The data have been obtained from the UCI ML repository [15]. The Iris data has $150 \times 4$ matrix $W = (w_{ij})_{150 \times 4}$ evenly distributed in three classes: iris-setosa, iris-versicolor, and iris-virginica. Vector of sample $i$, $(w_{i1}, w_{i2}, w_{i3}, w_{i4})$ has four features: sepal length and width, and petal length and width (all given in centimeters). Let $X = \{x_1, x_2, ..., x_{150}\}$, where $x_i = (w_{i1}, w_{i2}, w_{i3}, w_{i4}), i = 1, 2, ..., 150$, be the set of the 150 samples and $M = \{m_1, m_2, ..., m_8\}$ be the set of concepts on the features petal length or width, where $m_1$ is the concept: the petal length is mid, i.e., close to $5$ inches, $m_2$ is the concept: the petal is not long, i.e., $m_2 = m_1'$, the negation of $m_1$, $m_3$ is the concept: the petal length is mid, i.e., close to $\sum_{1 \leq i \leq 150} w_{i3}/150 = 3.7587$, $m_4 = m_3'$, $m_5$ is the concept: the petal width is wide, $m_6 = m_5'$, $m_7$ is the concept: the petal width is mid, i.e., close to $\sum_{1 \leq i \leq 150} w_{i4}/150 = 1.1987$, $m_8 = m_7'$. For each $m \in M$, according to the values of each pair $x_i$ and $x_j$ on the feature associating to concept $m$, the degrees of $x_i$ and $x_j$ belonging to concept $m$ can be compared, i.e., either $(x_i, x_j) \in R_m$ or $(x_j, x_i) \in R_m$ (refer to Definition 1). For example, if the petal length of $x_i$ is longer than or equal to that of $x_j$, then $(x_i, x_j) \in R_m$. Note that the petal lengths of $x_i$ and $x_j$ can be compared by human intuition, even if the exact petal lengths of $x_i$ and $x_j$ are not measured. In other words, to obtain the binary relation $R_m$ for a concept $m$, we only need the order relation description of the concept $m$. By Definition 2, one can verify that each $m \in M$ is a simple concept and for any $x, y \in X$, if define $\tau(x, y) = \{m | m \in M, (x, y) \in R_m\}$, then $(M, \tau, X)$ is an AFS structure. Let $S = 2^X$ be the $\sigma$-algebra on $X$ and $m$ be a measure on $S$ defined by: For any $A \subseteq X$, $m(A) = |A| (\bar{A} = \{x \in X | x / \in A\}$ is the number of elements of set $A$). By Definition 6, one knows that $(M, \tau, X, S, m)$ is a semi-cognitive field and for any fuzzy concept $\gamma = \sum_{i \in I} A_i \in EM$, since $S = 2^X$, then $\gamma$ is measurable in the semi-cognitive field $(M, \tau, X, S, m)$ and its membership function is defined by (6).

The classifier design algorithm

Step 1: Based on the given dataset $X$, learn the fuzzy descriptions $\zeta_{C_i}, i = 1, 2, ..., l$, through the fuzzy clustering analysis[30] with the AFS structure $(M, \tau, X)$ and $A = \{\{m\}|m \in M\} \subseteq EM$. Refer to the given classes $C_1, C_2, ..., C_l$, find appropriate parameters $\alpha \in (0, 1), \varepsilon > 0, \delta \in (0, 1)$ to optimize the accurate rate (refer to (24)) of the clustering results by the fuzzy clustering algorithm in Section 3 of paper [30]. Let fuzzy descriptions $\zeta_{C_i}, i = 1, 2, ..., l$ yield the best accurate rate of the clustering.

Step 2: For each test or new sample $y \notin X$, the membership degree of $y$ belongs to fuzzy concept $\eta = \sum_{i \in I} A_i \in EM$ is defined as

\[(8) \quad \mu_\eta(y) = \sup_{i \in I} \frac{|A_i(y)|}{|X|},\]
where

\[ A_i(\{y\}) = \{x \in X | \forall m \in A_i, \text{ the degree of } y \text{ belonging to } m \text{ is larger than or equal to that of } x\}. \]

The fuzzy classifier is denoted as \( \{\zeta_i| i = 1, 2, \ldots, l\} \). For each test or new sample \( y \notin X \), \( y \) is classified to class or pattern \( C_k \), if \( \mu_{\zeta_{C_k}}(y) = \max_{1 \leq i \leq l}\{\mu_{\zeta_{C_k}}(y)\} \).

**Discussion:**

Compared to the other fuzzy classifiers, the proposed fuzzy classifier has the following advantages:

1. Instead for control problems and function approximation problems, the proposed fuzzy classifiers are designed for classification problems to determine the class of new pattern. The classifiers can be trained by using part of all samples as training data and the correct rate evaluated by testing samples.

2. The design of the proposed classifier \( \{\zeta_i| i = 1, 2, \ldots, l\} \), in which each class is represented by a fuzzy set \( \zeta_{C_i} \in EM \) which determines the degree of the new pattern belonging to the class, is simple, comprehensible and similar to human recognition habit without model but the AFS fuzzy logic of the fuzzy sets in \( EM \), instead of fuzzy clustering model, genetic algorithm and T-S fuzzy model. Because each fuzzy set \( \zeta_{C_i} \in EM \) which represents a class has a definitive semantic interpretation in AFS fuzzy logic, the linguistic interpretation of the proposed classifiers is very comprehensible. Some molecular elements of the fuzzy sets \( \zeta_{C_i} \), describing the classes are on some special features, in stead of all features, hence the proposed classifiers still can determine the class of a new pattern when the new pattern loses data of partial features. For example, a fuzzy description \( \zeta_{C1} = \{m_6\} + \{m_2, m_4\} \), where molecular element \( \{m_6\} \) is on feature petal width and \( \{m_2, m_4\} \) is on feature petal length. For a sample \( x \), the degree of \( x \) belonging to the fuzzy description \( \zeta_{C1} = \{m_6\} + \{m_2, m_4\} \) can be obtained, as long as we know one of the features petal width and petal length, e.g. by the membership degree of \( x \) belonging to \( \{m_6\} \), the degree of \( x \) belonging to \( \zeta_{C1} = \{m_6\} + \{m_2, m_4\} \) can be estimated if the feature petal length is lost.

3. The design of the proposed classifier is valid without assumptions of data set \( X \subset R^n \), since the AFS structure \( (M, \tau, X) \), by which the membership functions of any fuzzy sets in \( EM \) can be obtained, can be established by the binary relations on \( X \). The binary relations can be abstracted from the database with any types of attribute, even linguistic description based on human intuitions.

4. Since the computation of the fuzzy descriptions of the training samples are independent each others, hence the classifier can be obtained by parallel processing and can be amended online.

In the following, using Iris dataset, we take 10 experiments. In each experiment, 60% samples of each classes are randomly selected as training samples and the other 40% samples are served as test samples. Figure 1 to 3 show the membership degrees of the test samples belonging to class 1 to 3 in the 1th, 4th and 7th experiments. The following are the detail results of the experiments:

<table>
<thead>
<tr>
<th>Fuzzy description</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-th experiments</td>
<td>( {m_6} + {m_2, m_4} )</td>
<td>( {m_3} + {m_7} )</td>
<td>( {m_1} + {m_5} )</td>
</tr>
<tr>
<td>4-th experiments</td>
<td>( {m_6} + {m_2} )</td>
<td>( {m_3} + {m_7} )</td>
<td>( {m_1} + {m_5} )</td>
</tr>
<tr>
<td>7-th experiments</td>
<td>( {m_6} + {m_2} )</td>
<td>( {m_3} + {m_7} )</td>
<td>( {m_1} + {m_5} )</td>
</tr>
</tbody>
</table>

Table 2: The incorrectly classified test samples in the 10 experiments
Figure 1. The membership degrees of the test samples belong to $\zeta_{C_1}, \zeta_{C_2}, \zeta_{C_3}$ in 1th experiment.

Figure 2. The membership degrees of the test samples belong to $\zeta_{C_1}, \zeta_{C_2}, \zeta_{C_3}$ in 4th experiment.

<table>
<thead>
<tr>
<th>No. experiment</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number error clustering</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>
4. the stability of the new classifier

In order to study the stability of the classifier, we study the fuzzy descriptions of each class obtained by the two training dataset which are randomly parted 150 samples into two equal number of sample set, \(X = X_1 \cup X_2, X_1 \cap X_2 = \emptyset, |X_1| = |X_2|\). We take 4 experiments. Figure 4, 5 show the membership functions of the different fuzzy descriptions of class 1 and 3. By the figures, we can observe that although the fuzzy sets in EM to describe a class may be different, i.e., with different semantic interpretations, the membership functions of them are very similar. For a class of objects, different persons may have different descriptions, but the same classification results often are obtained by the different descriptions of the classes. This implies that the proposed classifier is very stable and is the un-variable pattern of the dataset.

<table>
<thead>
<tr>
<th>Table 3: The compared classifiers of the 4 experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>fuzzy descr.</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>1st exp. (X_1)</td>
</tr>
<tr>
<td>1st exp. (X_2)</td>
</tr>
<tr>
<td>2nd exp. (X_1)</td>
</tr>
<tr>
<td>2nd exp. (X_2)</td>
</tr>
<tr>
<td>3rd exp. (X_1)</td>
</tr>
<tr>
<td>3rd exp. (X_2)</td>
</tr>
<tr>
<td>+ {m_2, m_6}</td>
</tr>
<tr>
<td>4th exp. (X_1)</td>
</tr>
<tr>
<td>4th exp. (X_2)</td>
</tr>
</tbody>
</table>
5. Conclusion

In this paper, in the framework of AFS theory, we propose a new design of fuzzy classifiers and fuzzy descriptions based on the fuzzy clustering analysis\(^{(30)}\) in which the attributes of objects can be various data types or sub-preference relations (i.e., simple fuzzy concepts) while by other classification algorithms, it is difficult
or impossible to study clustering problems of these kinds. The stability and the robust property of the proposed classifiers are demonstrated by the similarity of the membership functions of the fuzzy sets describing a given class, which are learned from the randomly selected samples from the dataset. The proposed classifier is simply represented as \( \{ \zeta_{C_i} | i = 1, 2, \ldots, l \} \), which are fuzzy sets in \( EM \). The design algorithm is comprehensible and similar to human recognition habit without model but the AFS fuzzy logic of the fuzzy sets in \( EM \), instead of fuzzy clustering model, genetic algorithm and T-S fuzzy model. Because each fuzzy set \( \zeta_{C_i} \in EM \) which represents a class has a definitive semantic interpretation in AFS fuzzy logic, the linguistic interpretation of the proposed classifiers is very comprehensible. Indeed, this approach also can be regard as the knowledge representation of the training data. This implies that the approaches in this paper are also a new data mining method. We would like to share this new idea with more mathematicians, scientists and engineers.

References

Department of Mathematics, Dalian Maritime University, Dalain 116026, P. R. China
Research Center of Information and Control, Dalian University of Technology, Dalian 116024, P. R. China
E-mail: wxcix11@sohu.com