Knowledge Acquisition Tool for Learning Membership Function and Fuzzy Classification Rules from Numerical Data

FadlMutaher Ba-Alwi
Faculty of Computer and IT
Sana'a University – Yemen

ABSTRACT
Generating suitable membership function (MF) is the core step of fuzzy classification system. This paper presents a novel learning algorithm that generates automatically reasonable MFs for quantitative attributes. In addition, a set of an appropriate fuzzy classification rules (FCRs) are discovered from a given numerical data. Each fuzzy rule (FR) is of the form IF-THEN rule. The antecedent IF-part and consequent THEN-part contain fuzzy sets. Since MFs are generated automatically, the proposed fuzzy learning algorithm can be viewed as a knowledge acquisition tool for classification problems. Experimental results on Iris dataset are presented to demonstrate the contribution of the proposed approach for generating MFs.

General Terms
Fuzzy Classification Systems, Knowledge Discovery in Databases (KDD), Knowledge Acquisition tool,Data Mining, Pattern Recognition.

Keywords
Fuzzy Classification Rule (FCR), knowledge acquisition tool, Learning algorithm, Membership function (MF).

1. INTRODUCTION
A classification problem involves learning a set of classification rules (CRs) from a training dataset. These CRs are extracted in classical form IF-THEN production rules (PRs); i.e. IF *antecedent-part* THEN *consequent-part*. The antecedent part contains conjunction of n conditions on values of predictor attributes, whereas the rule consequent contains a prediction about the value of a target attribute. The limitation of this classical representation is that they are unable to capture the cognitive human process because the only absolutely true statements are considered. More precisely, PRs are not capable of dealing with cognitive uncertainties to real world decision making [5]. Fuzzy sets is an adequate theory to develop tools for modeling cognitive human processes related to the aspects of recognition [1] because of using the notions of truth and falsehood in a graded fashion [2]. Fuzzy logic (FL),which is based on fuzzy sets theory, provides a natural way for constructing fuzzy IF-THEN rules that are closer to the human decision making process by using linguistic interpretations in a mathematical framework [3]. Integrating FL into the design of rule based systems will result a Fuzzy Classification Rule System (FCRS). From machine learning literature, Fuzzy Classification Rules (FCRs) have three different cases [1]. Case1: Rules with a single class in the consequent, case2: Rules with a single class and a rule weight associated to this class in the consequent, and case3: Rules with rule weights associated to each one of the class of the consequent. FCR with first case is simple and intuitive for human users [4]. And hence, it is adopted for discovering FCRs in this paper. Given dataset with quantitative attributes (n predicting attributes and one target attribute, which can take m classes), and t training patterns: $v_p = (v_{p1}, v_{p2},\ldots,v_{pn})$, p = 1, 2, . . .t, the fuzzy rule (FR) for p-th training pattern takes the following form:

$$FR_p\text{:IF } v_{p1} \text{ is } R_{k1} \text{ AND } \ldots\text{AND } v_{pn} \text{ is } R_{kn} \text{ THEN class}_p \text{ is } C_k$$

where FR_p is the label of the p-th fuzzy rule, v_{pi} is the value of the i-th attribute (i = 1…n) of the p-th training pattern,R_{ki} is the k-th fuzzy set (i.e. fuzzy region) of the i-th attribute (k=1...number of fuzzy sets of i-th attribute), and C_k is a consequent fuzzy set of the class attribute of the p-th pattern (k=1...number of fuzzy sets of class attribute). Several approaches are proposed in the literature for mining of FCRs [5-11]. The core step of fuzzy classification system is building suitable membership functions (MFs) of each attribute to generate a set of FRs suitable to deal with a specific classification problem. Lots of procedures for building MFs have been devised and can be easily found in literature, ranging from very simple heuristics that start slicing the domain of values of a linear variable to the extensive use of evolutionary algorithms and well-known numerical procedures such as interpolation [12]. This paper presents a novel learning algorithm that generates automatically reasonable MFs for quantitative attribute. In addition, a set of an appropriate FCRs are discovered from a given numerical data. Each generated FR is of the form IF-THEN rule. The antecedent IF-part and consequent THEN-part contain fuzzy sets. Since MFs are generated automatically, the proposed fuzzy learning algorithm can be viewed as a knowledge acquisition tool for classification problems. The rest of this paper is organized as follows: Section 2 briefly describes the related works about generating MFs and discovery of FCRs. Section 3 presents new proposed learning algorithms for constructing fuzzy MFs and discovery of FCRs. Section 4 describes the computational results of the proposed approach. Finally, the conclusions are discussed in section 5.

2. RELATED WORK
In recent years, several approaches have been proposed in the literature for generating Membership Functions (MFs). The authors in [14] divided each domain interval of input and output spaces into 2N+1 fuzzy regions (N can be different for different variables, and the length of these regions can be equals or unequal) denoted by SN (Small N),…,S1 (Small 1), CE (Center), B1 (Big 1),…,BN (Big N), and assign each region a fuzzy MF. The authors in [15] proposed a general learning method as a framework for automatically deriving MFs and fuzzy IF-THEN rules from a set of given training

examples. The output values of all training instances are appropriately grouped by applying a proposed clustering procedure, and appropriate MFs for output values are derived. In [16-18] it is proposed an algorithm that extracts MFs by means of a genetic learning of the MFs. Based on the correlation coefficient threshold value a methods in [19, 20] are proposed to construct and tune MFs and generate FCRs from training instances for handling the Iris data classification problem. Fuzzy C-Means algorithm proposed in [21] partitions a quantitative attributes into several fuzzy sets to generate membership values. First, the algorithm divide the domain of each input variable into several overlapping length intervals and each interval is associated with a fuzzy triangular MF. Then, a label is assigned to each MF. Several methods based on heuristics, hybrid methods, fuzzy clustering algorithms, neural networks, genetic algorithms, and the entropy for the automatic generation of MFs were discussed in [22]. In the present work, a knowledge acquisition tool for classification problem is defined byproposing a novel learning algorithm to generate automatically reasonable MFs for quantitative attribute. In addition, a set of an appropriate FCRs are discovered from a given numerical data.

3. THE PROPOSED FCRS

The process of fuzzy systems design involves the following steps [23]:

Step1: Define the input (I/P) and output (O/P) attributes.

Step2: Constructing membership functions (MF) for I/P and O/P attributes.

Step3: Generate fuzzy rule base for the system.

Step4: Fuzzy inference engine.

Step5: Select the defuzzification technique for generating a crisp O/P value.

In step1, the input and output attributes can be identified by many ways either by the experts, forward selection procedure, backward elimination procedure, or by statistical selection procedures [23]. In this paper, the output attribute, i.e. class attribute, is predefined. Step2 is an important task in designing fuzzy classification system. In this paper, a new learning algorithm is proposed to construct MFs for each quantitative attribute to generate a set of FRs suitable to deal with fuzzy classification system. In step3, suitable FRs can be generated either by domain experts, or by rule base generation techniques: genetic algorithm based methods [24], and neural networks based methods [25]. This paper presents learning algorithm for generating FRs and discovery of FCRs. Some literature in fuzzy systems considers the above first three steps as one step called fuzzification step. In step4, lots of fuzzy inference engine for evaluating the rules in the rule base are found in the literature [1, 13]. Finally, amongst large number of defuzzification techniques which are available in the literature, the Center of Average Defuzzifier (COA) technique is the most commonly used in fuzzy systems and fuzzy control because it is computationally simple and intuitively plausible.As mention above, MF is the core step of designing fuzzy classification systems. Generating MFs for a quantitative attribute A_i is based on three factors: shape of fuzzy MF, number of fuzzy regions assigned to attribute A_i (i.e. fuzzy partition), and fuzzy regions overlapping. For the first factor, this paper used the most popular shape the so-called triangular MF. Second factor is based on the domain in which the MF is defined. A learning algorithm is proposed to generate proper fuzzy regions for values of A_i by mining the proper interval values, boundaries, of each region in axis of A_i. Finally, a measure is proposed for computing the degree of overlap between the adjacent fuzzy regions. The two last factors are based on the following definitions:

Definition 1. Let Incr(A_i) denotes for the increment unit in the axis of A_i. This measure can be defined as follow:

$$\text{Incr}(A_i) = \left\lfloor \frac{A_i^{max} + A_i^{min}}{|A_i|} * w \right\rfloor \qquad (1)$$

where A_i is the i-th attribute, $A_i^{min}$is the minimum value of A_i, $A_i^{max}$is the maximum value of A_i, $|A_i|$ is the number of distinct values of A_i, and w is a positive integer user-defined weight. The user-defined weight is used to control the number of regions needs to be created. Larger w will have a small number of fuzzy regions. This measure helps on defining a proper vector of interval values (shortly named "vector unit") of fuzzy regions in the universe of discourse X=[X_{min}, X_{max}] as follow:

vector unit(A_i) ={X_{min}, Incr(A_i), Incr(A_i)*2,…., Incr(A_i)*k}

whereX_{min} and X_{max} are the minimum and maximum values of X respectively, and k is a positive integer number such that Incr(A_i)*(k-1) <A_i^{max} ≤ Incr(A_i)*k. For simplicity, it can be written as: vector unit(A_i)={u_1, u_2,….,u_{max}},where u_1=X_{min}, u_2=Incr(A_i),…,and u_{max}=Incr(A_i)*k, and u_{max-1}<A_i^{max} ≤ u_{max}.

Example 1: Suppose X=[0...100] is the universe of discourse of attribute Age = {5, 10, 11, 30, 65, 70}, then Age^{min}=5, Age^{max}=70, and |Age| =6. The increment unit with w=1 is Incr(Age) =$\left\lfloor \frac{5+70}{6} \right\rfloor$ = 12. So, the vector unit(Age)={0, 12, 24, 36, 48, 60, 72} as shown in Figure 1. Note that the last unit 72 with k=6, satisfy the condition: 60 <A_i^{max}=70 ≤ 72. But for k=7 the condition is not satisfied because 72 ≮ 70 ≤ 84.

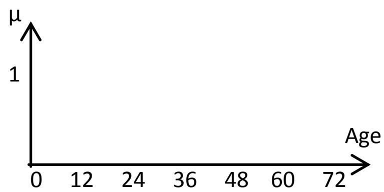

Fig 1: Increment unit for attribute Age

Definition 2. Let $C^i(R_l, j, k)$, j ≠ k, denotes for the set of all classes for whichV_{ij} and V_{ik} are belongs to, where both V_{ij} and $V_{ik}$are in the same region R_l. This measure can be defined as follow:

Suppose V_{ij} belongs to the set of classes C_j^i={C_{j1}^i, C_{j2}^i,… }, and V_{ik} belongs to the set of classes C_k^i={C_{k1}^i, C_{k2}^i,… }, then $C^i(R_l, j, k)$ = {C_j^i} ∪ {C_k^i}. Note that this definition can also be applied for one value. In other word, one of these two values can be absent.

Example 2: Given dataset with 10 classes (C_1, C_2 ,..., C_{10}) and attribute Age = {5, 10, 11, 30, 65, 70}.Suppose values 5 and 10 are in the same region R. Also, assuming that value 5 belongs to classes {C_1, C_2, C_6} and value 10 belongs to classes {C_2, C_5}. Therefore,C^{age}(R, 5, 10) = {C_1, C_2, C_5, C_6}.

Definition 3. Let $C_{all}^i(R_l, R_s)$ denotes for the set of all classes of regions R_l and R_s. This measure can be defined as follow:

Suppose region R_l contains the values {V_{ij},V_{ik}}, and region R_s contains the values {V_{id},V_{ir}}, according to definition 2, we can get $C^i(R_l, j, k)$ and $C^i(R_s, d, r)$. Therefore, $C^i_{all}(R_l, R_s)$ = $\{C^i(R_l, j, k)\} \cup \{C^i(R_s, d, r)\}$

Example 3: Referring to the attribute Age in example 2.Suppose values (5, 10) are in the same region R_l and value 30 is in other region R_2. Assume that $C^{age}(R_1, 5, 10) = \{C_1, C_2, C_5, C_6\}$ and $C^{age}(R_2, 30) = \{C_3, C_4, C_6\}$. Therefore, $C^{Age}_{all}(R_1, R_2) = \{C_1, C_2, C_3, C_4, C_5, C_6\}$.

Definition 4. Let $C^i_{com}(R_l, R_s)$denotes for the set of common classes of the regions R_l and R_s. This measure can be defined as follow:
Suppose region R_l contains the values $\{V_{ij}, V_{ik}\}$, and region R_s contains the values $\{V_{id}, V_{ir}\}$, according to definition 2, we can get $C^i(R_l, j, k)$ and $C^i(R_s, d, r)$. Therefore, $C^i_{com}(R_l, R_s)$ = $\{C^i(R_l, j, k)\} \cap \{C^i(R_s, d, r)\}$.

Example 4: Referring to all information in example 3, $C^{Age}_{com}(R_1, R_2) = \{C_6\}$.

From the above definitions 2, 3 and 4, the proposed measure for computing the degree of overlap between the adjacent fuzzy regions R_l and R_s is computed as:

$$Overlap(R_l, R_s) = \frac{|\mathbf{C}^i_{com}(\mathbf{R}_l, \mathbf{R}_s)|}{|\mathbf{C}^i_{all}(\mathbf{R}_l, \mathbf{R}_s)|} \quad (2)$$

Example 5: The degree of overlap between the adjacent fuzzy regions R_1 and R_2 (given in examples 2 and 3) is 1/6 = 0.2.

3.1 The Proposed Learning Algorithm for Constructing MFs

A triangular MF is specified by three parameters (boundaries): lower limit value (left vertex), upper limit value (right vertex), and modal value (center vertex). In the present work, fuzzy regions are represented by triangular MFs with highest 1 at the center of the fuzzy region. The distinct values of an attribute A_i are represented by a set of regions $\{R_1, R_2, \ldots, R_k, R_{k+1}, \ldots\}$. Constructing fuzzy region R_k needs to identify its three parameters: lower limit value (R^l_k), upper limit value (R^u_k), and modal value (M_k). The following proposed learning algorithm discovers the initial parameters of all fuzzy regions for the distinct values of attribute A_i.

Input: $A\grave{_i}$ values of $A\grave{_i}$are sorted in ascending order, where $A\grave{_i}$= A_i.
Vector unit($A\grave{_i}$)=$\{u_1, u_2, \ldots, u_{max}\}$ represents the intervals values of A_i axis.
Output: R_{Ai}: set of fuzzy regions with initial lower and upper parameters for values of A_i. Initially R_{Ai}= Ø.
Construct MF algorithm
Begin
Step1: /*Assign values to lower and upper limit boundaries of fuzzy region R_k, initially k=1*/
Let R^l_k = u_1 (first unit) and R^u_k= u_2 (next unit).
Let G_k is a group of closest units for region R_k, initially G_k = Ø.
Step2: Repeat
Let G_k contains all values of $A\grave{_i}$ which ∈ [R^l_k,R^u_k].
If G_k = Ø and $R^u_k \neq u_{max}$ Then
Expand the upper limit of R_k as:R^u_k = next unit.
If G_k = Ø and R^u_k = u_{max} Then G_k = all values in $A\grave{_i}$.
Else Begin
Create region R_k(R^l_k,R^u_k) for the values in G_k.
$R_{Ai} = R_{Ai} \cup \{R_k\}$
End
Until region R_k is created.
Step3: Remove from $A\grave{_i}$ all values ∈ [R^l_k,R^u_k].
Remove from the Vector unit($A\grave{_i}$) all units which are less thanR^u_k.
Step4: If all values of $A\grave{_i}$are covered i.e. $A\grave{_i}$= Ø Then returnR$_{Ai}$ and Stop
Else back to step1 with k=k+1to generatenext region.
End /*algorithm*/

The above learning algorithm discovers the initial lower and upper parameters of all fuzzy regions of A_i. The modal value M_k for any region R_k(R^l_k,R^u_k) is computed as: $M_k = \frac{R^l_k + R^u_k}{\mathbf{2}}$**.** In this paper, the degree of overlap between adjacent functions of output attribute is assumed to be 50%. Whereas, MFs of an input attribute may not overlap each other. The shape of MFs for an input attribute is affected by the result of overlap measure (formula 2). Given two adjacent fuzzy regions R_1 and R_2, if the Overlap(R_1, R_2)=0 then the lower and upper limit values of both regions are the same as the initial values discovered in the above learning algorithm. If the Overlap(R_1, R_2)=1 then merge R_1 and R_2into one region such that lower and upper limit values of the new region are R^l_1 and R^u_2, respectively. But in case degree of Overlap(R_1, R_2)=£, where 0<£<1, then the upper limit value of fuzzy region R_1 will be change as R^u_1=(R^u_2 - R^l_2)*£ +R^l_2, i.e. the percentage amount of shifting R^u_1 to inside of region R_2. Further, the lower limit value of fuzzy region R_2 will be also change as R^l_2= R^u_1-(R^u_1 - R^l_1)*£, i.e. the percentage amount of shifting R^l_2 to inside of region R_1. In all cases no change on the modal value.Finally, the membership degree of an input value, say x, can be decided either subjectively or by a MF defined on the range of numeric values of the attribute. In this paper, the following triangular MF is employed for computing the degree of value x which belongs to region R_k [2]:

$$\mu(x, R^l_k, M_k, R^u_k) = \max(0, \min(\frac{x - R^l_k}{M_k - R^l_k}, \frac{R^u_k - x}{R^u_k - M_k})).$$

Example 6: Consider the sorted values of attribute Age`= {5, 10, 11, 30, 65, 70}, andvector unit =$\{u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$={0, 12, 24, 36, 48, 60, 72}.In step1: R^l_1 = 0, R^u_1= 12, and G_1 = Ø.In step2: Values 5, 10, and 11 are ∈ [0, 12] and hence, G_1 = {5, 10, 11}.$G_k \neq$ Ø, therefore, region R_1 is created with parameters:R^l_1=0,R^u_1=12, and M_1= 6 (i.e. (0+12)/2) as shown in figure 2.

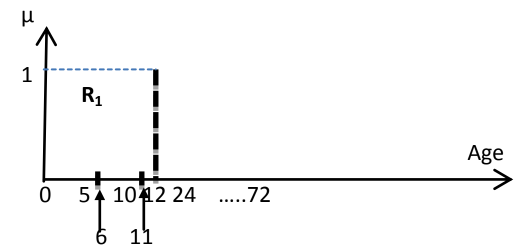

Fig 2: Generating region R_1.

In step3: Removing values of G_1 from Age`={30, 65, 70}. The first unit=0 <R^u_1=12 is removed from vector unit to be = {12, 24, 36, 48, 60, 72}.
In step4: The steps are repeated to generate the next fuzzy regions. Region R_2(R^l_2=12,R^u_2=36, M_2=24) is created for value

Age=30. The last region R_3(R_3^l=36, R_3^u=72, M_3=54) is created for values Age= {65, 70}. Figure 3 shows the three created fuzzy regions with their initial parameters for attribute Age.

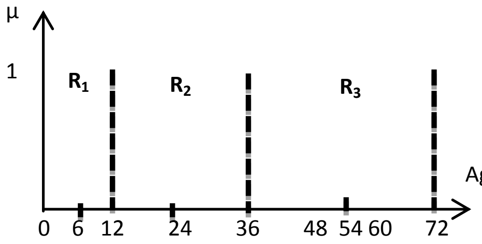

Fig 3: Three fuzzy regions for attribute Age.

Figure 4 shows the three triangular MFs for the Age attribute without overlapping between the regions. While, figure 5 shows the MFs with Overlap(R_1, R_2)= 0.5 and Overlap(R_2, R_3)= 0.3. The Overlap(R_1, R_2)= 0.5 will cause changing the values of $R_1^u$and R_2^las follows: R_1^u=(36-12)*0.5+12= 24 and R_2^l=12-(12-0)*0.5= 6. Further, the Overlap(R_2, R_3)= 0.3 will cause changing the values $R_2^u$and R_3^las follows: R_2^u=(72-36)*0.3+36= 46.8 and R_3^l=36-(36-12)*0.3=28.8.

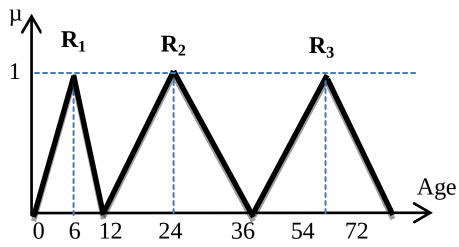

Fig 4: MFs without overlap.

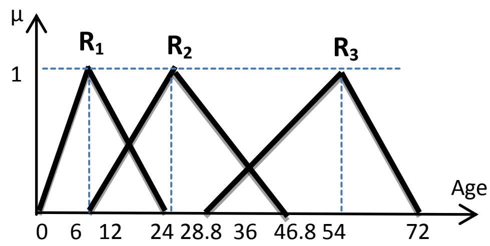

Fig 5: MFs with overlapping.

Referring to figure 5, the value Age=30 belongs to fuzzy regions R_2 and R_3 with different degrees: μ(30, R_2) = μ(30, 6, 24, 46.8) = 0.74 and μ(30, R_3) = μ(30, 28.8, 54, 72) = 0.05.

3.2 The Proposed Discovery of FCRs

As mention earlier a multi-input single-output fuzzy rule based system is considered in the present work. Given dataset D with quantitative attributes (n predicting attributes and one target attribute, which can take m classes), and t training patterns: v_p = (v_{p1}, v_{p2},…,v_{pn}), p = 1, 2, . . .t, the FR for p-th training pattern takes the following form:

FR_p: IF v_{p1} is R_{k1} AND ….AND v_{pn} is R_{kn} THEN $class_p$ is C_k

where FR_p is the label of the p-th fuzzy rule, v_{pi} is the value of the i-th attribute (i = 1…n) of the p-th training pattern,R_{ki} is the k-th fuzzy set (i.e. fuzzy region) of the i-th attribute (k=1...number of fuzzy sets of i-th attribute), and C_k is a consequent fuzzy set of the class attribute of the p-th pattern (k=1...number of fuzzy sets of class attribute).Note that each fuzzy set has three parameters e.g. R_{ki} (R_k^l, M_k, R_k^u). Generating FR for each pattern v_p = (v_{p1}, v_{p2},…,v_{pn}; $class_p$) needs to determine the membership values of v_{pi} in fuzzy sets R_{ki} (k=1,...,number of MFs of i-th attribute) and the membership values of $class_p$ in fuzzy sets C_k (k=1,..,number of MFs of class attribute). Further, for each attribute A_{pi} (i = 1,..., n), determine the fuzzy set in which v_{pi} has the largest membership value, that is, determine R_{ki}^* such that $\mu_{R_{ki}^*}(v_{pi}) \geq \mu_{R_{ki}}(v_{pi})$ for (k=1,..., number of MFs of A_{pi}). Furthermore, determine the fuzzy set in which $Class_p$ has the largest membership value, that is, determine C_k^* such that $\mu_{C_k^*}(class_p) \geq \mu_{C_k}(class_p)$ for (k=1,..., number of MFs of class attribute). Finally, generate the following FR [18]:

FR_p^* : IF v_{p1} is R_{k1}^* AND v_{p2} is R_{k2}^* AND ….AND v_{pn} is R_{kn}^* THEN $class_p$ is C_k^* (3)

Since there are t patterns in the dataset D, the maximum number of FRs of form (3) that can be generated is t which is large number. Removing conflict rules, that is, rules with the same IF-parts but different THEN-parts will reduce the number of generated FRs.And that can be done by assigning degree to each generated FR_p^* and keep only one rule from a conflicting group that has the maximum degree. The degree of each FR_p^* is computed as follow [13]:

$$\text{Degree}(FR_p^*) = \prod_{i=1}^{n} \mu_{R_{ki}^*}(v_{pi})\,\mu_{C_k^*}(class_p)$$

where$\mu_{Rki} \in$ [0,1] is the membership degree of the v_{pi}. Although, the generated fuzzy rules are reduced by removing conflict rules, also they can be more reducing by classifying them. The generated fuzzy rules are grouped according to the class valuein the consequent part, i.e. FRs with the same class value are grouped. From each group, keep only one rule that has the maximum classification measure (CM) which proposed as follow:

$$\text{CM}(FR_p^*) = \frac{w * Degree(FR_p^*) * Cov(FR_p^*)}{NFS(FR_p^*)}$$

where w is a user-defined weight ∈ [0, 1], it can be different for different groups,Cov(FR_p^*) is the number of patterns covered by rule FR_p^*, and NFS(FR_p^*) is the number of antecedent fuzzy sets in the fuzzy rule FR_p^*, i.e. the simplicity of the rule.Rule with the lowest number of fuzzy sets is the better.

4. EXPRIMENTS

The performance of the proposed algorithm is investigated on the public domain dataset, the so-called Iris. TheIris dataset contains 150 instances with four input quantitative attributesi.e., SepalLength (SL), Sepal Width (SW), Petal Length (PL) and Petal Width (PW). There are three species of flowers in the Iris data, i.e., Iris-Setosa (class c1), Iris-Versicolor (class c2)and Iris-Virginica (class c3). Table 1 shows the maximum attribute value and the minimum attribute value of the input attributes of the Iris data, respectively.

Table 1.Maximum and minimum attribute values of the input attributesof the Iris dataset.

Attribute (A_i)	A_i^{max}	A_i^{min}
Sepal Length (SL)	7.9 cm	4.3 cm
Sepal Width (SW)	4.4 cm	2.0 cm
Petal Length (PL)	6.9 cm	1.0 cm
Petal Width (PW)	2.5 cm	0.1 cm

The initial MFs with their parameters for the four attributes are given in figure 6. For each attribute, the computations of overlap degree with w=10 between the adjacent regions are given in Table 2. In this table,column "Region" shows the set of fuzzy regions of each attribute. The values in these regions are belonging to set of classes given in column "Classes". Based on the value of overlap measure between any two adjacent regions R_k and R_{k+1}, the lower $R_{k+1}^l$and upper R_k^u values of these regions may change or both regions may merge. The final MFs are given in figure 7. Note that, for the attribute PL, regions R_2 and R_3 are merged as well as regions R_4 and R_5 (see sub-figure c-1). The overlapping between the fuzzy regions of attribute PL is given in the sub-figure (c-2).

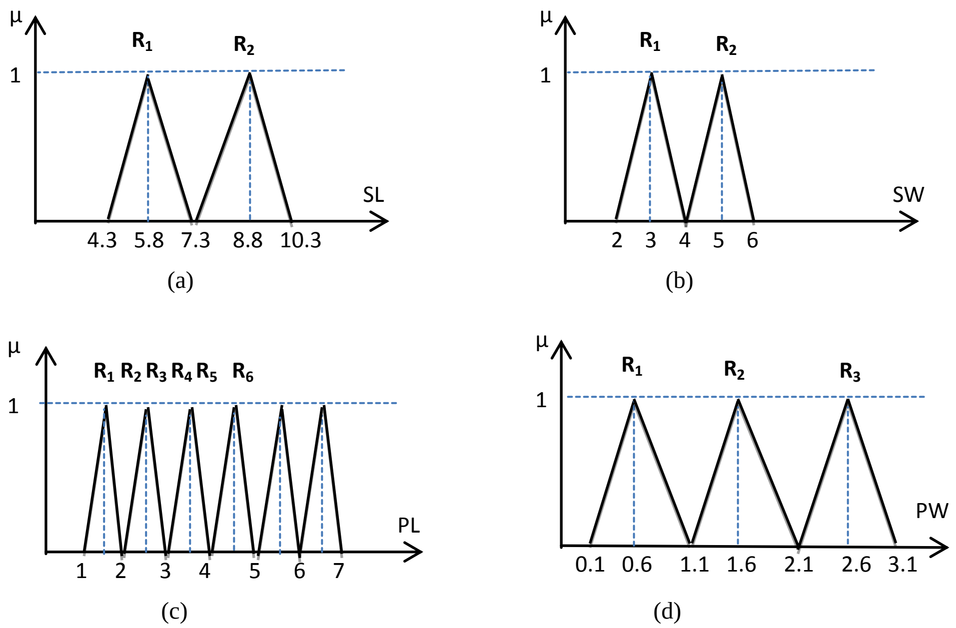

Fig 6: The initial MFs: (a) SL attribute (b) SW attribute (c) PL attribute (d) PW attribute.

Table 2.The overlap degree between adjacent regions.

A_i	Region R_k	Classes	Overlap (R_k, R_{k+1})	R_k^u	R_{k+1}^l	Merge (R_k, R_{k+1})
SL	R_1	{c1, c2, c3}	(R_1, R_2) = 0.3	R_1^u=8.2		
	R_2	{c3}			R_2^l= 6.4	
SW	R_1	{c1, c2, c3}	(R_1, R_2) = 0.3	R_1^u= 4.6		
	R_2	{c1}			R_2^l= 3.4	
PL	R_1	{c1 }	(R_1, R_2) = 0.0			
	R_2	{c2}				
	R_3	{c2}	(R_2, R_3) = 1.0			Merge(R_2, R_3)
	R_4	{c2, c3}	(R_3, R_4) = 0.5			
	R_5	{c2, c3}	(R_4, R_5) = 1.0			Merge(R_4, R_5)
	R_6	{c3}	(R_5, R_6) = 0.5			
PW	R_1	{c1, c2}	(R_1, R_2) = 0.3	R_1^u= 1.4		
	R_2	{c1, c2}	(R_2, R_3) = 0.5	R_2^u=2.6	R_2^l= 0.8	
	R_3	{c3}			R_3^l= 1.6	

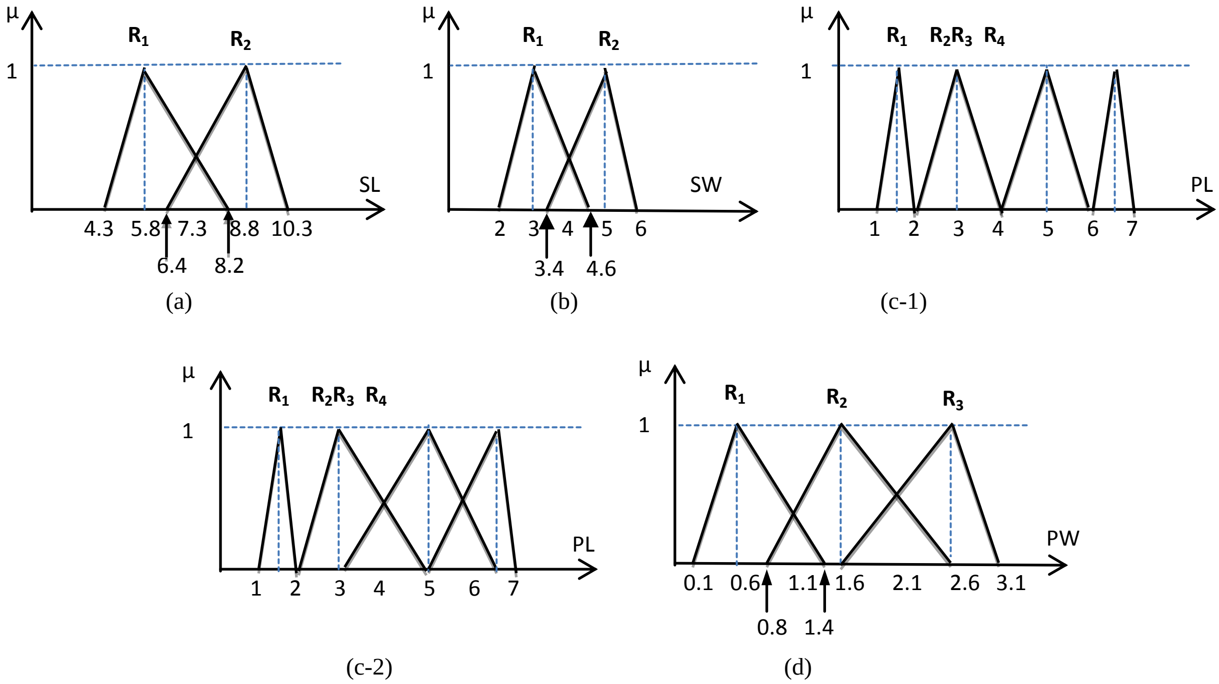

Fig 7: The final MFs for (a) SL attribute (b) SW attribute (c) PL attribute (d) PW attribute.

The final MFs are given in figure 7. Note that, regions R_2 and R_3 of attribute PL are merged as well as regions R_4 and R_5 (see sub-figure c-1). The overlapping between the fuzzy regions of attribute PL is given in the sub-figure (c-2).

4.1 Comparative Study

The constructed MFs, for Iris dataset, by the proposed learning algorithm are compared with the results obtained by Learning Algorithm (LA) used in [15, 20]. The comparison given in Table 3 is based on the number of fuzzy sets generated for PL and PW attributes by the three LAs. In Table 3, LA_1, LA_2, and LA_p means the learning algorithms developed in [15], in [20], and proposed learning algorithm, respectively. The comparisons show that the proposed learning algorithm gives rational result.

Table 3.Comparisons between proposed and others LA.

LAs	A_i	R_1 (R_1^l, R_1^u)	R_2 (R_2^l, R_2^u)	R_3 (R_3^l, R_3^u)	R_4 (R_4^l, R_4^u)
LA_1	PL	(0, 4.9)	(4, 4.8)	(4.9, ∞)	
LA_2		(0.8,2.4)	(2.9,5.3)	(4.46,7.2)	
LA_p		(1, 2)	(2, 5)	(3, 6.5)	(5, 7)
LA_1	PW	(0, 1.1)	(0.3,1.7)	(1.6,1.8)	(1.7,∞)
LA_2		(-1.2,0.8)	(0.66,2)	(1.06,3.2)	
LA_p		(0.1,1.4)	(0.8,2.6)	(1.6,3.1)	

5. CONCLUSION

In this paper, a novel learning algorithm is developed for automatically driving reasonable membership functions (MFs) and reduces the number of generated fuzzy classification rules from numerical data. The proposed fuzzy learning algorithm can be viewed as a knowledge acquisition tool for classification problems. Experimental results on Iris dataset have been demonstrated the contribution of the proposed approach and the results are satisfactory. Fuzzy genetic algorithm may be introduced for generating membership functions suitable to deal with a specific classification problem as an ulterior development direction.

6. REFERENCES

[1] Fernandez A. and Herrera F., 2012 Linguistic Fuzzy Rules in Data Mining: Follow-Up Mamdani Fuzzy Modeling Principle. Springer-Verlag Berlin Heidelberg.

[2] Spada L., 2009 Introduction to Fuzzy Sets and Fuzzy Logic, Reasonpark.Foligno, 17-19 Sep.

[3] Zadeh L. A., 1996 Fuzzy Logic = Computing with Words, IEEE Transactions on Fuzzy Systems, vol. 4, No. 2, 103–111.

[4] Ishibuchi H., 2000 Effect of Rule Weights in Fuzzy Rule-Based Classification Systems, IEEE Transactions on Fuzzy Systems, vol.9, No.4, 506-515.

[5] Bala, B. and Saroj, 2012 Discovering Fuzzy Censored Classification Rules: a Genetic Algorithm Approach, International Journal of Artificial Intelligence & Applications (IJAIA), vol.3, No.4, 175-188.

[6] Mansoori E.G., Zolghadri M.J., and Katebi S.D., 2008 SGERD: A Steady-State Genetic Algorithm for Extracting Fuzzy Classification Rules from Data, IEEE Transactions on Fuzzy Systems, vol. 16, No. 4.

[7] Hu Y., Chen R.S., Tzeng G.H., 2003 Finding Fuzzy Classification Rules using Data Mining Techniques, Pattern Recognition Letters 24, 509–519.

[8] Hu Y., Chena R., Tzengb G., 2002 Mining Fuzzy Association Rules for Classification Problems, Computers & Industrial Engineering, vol.43, 735–750.

[9] S´aez J.A., Luengo J., Herrera F., 2011 Fuzzy Rule Based Classification Systems versus Crisp Robust Learners Trained in Presence of Class Noise's Effects: a Case of Study, IEEE, 11-th International Conference on Intelligent Systems Design and Applications, 1229-1234.

[10] Roubos J.A., Setnes M.,, Abonyi J., 2003 Learning Fuzzy Classification Rules from Labeled Data, Information Sciences, 150, 77–93.

[11] Dehzangi O., Zolghadri M. J., Taheri S., and Fakhrahmad S.M., 2007 Efficient Fuzzy Rule Generation: A New Approach using Data Mining Principles and Rule Weighting. 4-th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)-IEEE.

[12] Rodr´ıguez Y., Falc´on R., Varela A., and Garc´ıa M.M., 2008 Learning Membership Functions for an Associative Fuzzy Neural Network, Springer-Verlag Berlin Heidelberg, 151-161.

[13] Wang L., 1997 A Course in Fuzzy System and Control, International edition, Prentice-Hall International, Inc.

[14] Wang L., and Mendel M., 1992 Generating Fuzzy Rules by Learning from Examples, Transactions on Systems and Cybernetics, vol.22, No.6, 1414-1427.

[15] Hong T., and Leeb C., 1996 Induction of Fuzzy Rules and Membership Functions from Training Examples, Fuzzy Sets and Systems, vol.84, 33-47.

[16] Chen S.M., and Chen Y.C., 2002 Automatically Constructing Membership Functions and Generating Fuzzy Rules using Genetic Algorithms. Cybernetics and Systems: An International Journal, vol.33, No. 8, 841-863.

[17] Alcal´a-Fdez J., Alcal R., Gacto M., and Herrera F., 2009 Learning the Membership Function Contexts for Mining Fuzzy Association Rules by Using Genetic Algorithms,Fuzzy Set and Systems vol.160, No.7, 905-921.

[18] Huynh T., Nguyen H., Le B., 2012 A Unified Design for the Membership Functions in Genetic Fuzzy Systems, International Journal of Computer Science Issues, vol.9(3), No.1, 1694-0814.

[19] Chen S., and Fang Y., 2005 A New Method to deal with Fuzzy Classification Problems by Tuning Membership Functions for Fuzzy Classification Systems, Journal of the Chinese Institute of Engineers, vol.28, No. 1.

[20] Chen S., and Tsai F., 2005 A New Method to Construct Membership Functions and Generate Fuzzy Rules from Training Instances, Information and Management Sciences, vol.16, No. 2, 47-72.

[21] Radha R., and Rajagopalan S.P., 2010 Generating Membership Values and Fuzzy Association Rules from Numerical Data, vol.02, No.08, 2705-2715.

[22] Cintra M.E., Camargo H.A., and Monard M.C., 2008 A Study on Techniques for the Automatic Generation of Membership Functions for Pattern Recognition, Ibero-American Conference Publications, vol.1, 1-10.

[23] Kumar S., Bhalla P., and Singh A., 2009 Fuzzy Rule base Generation from Numerical Data using Biogeography-based Optimization, IE(I) Journal–ET, vol.90, 1-6.

[24] Juang C., 2005 Combination of Online Clustering and Q-value based GA for Reinforcement Fuzzy System Design. IEEE Transactions on Fuzzy Systems, vol.13, No 3. 289-302.

[25] Ishibuchi H., et al.1993 Neural Networks that Learn from Fuzzy If Then Rules, IEEE Transaction on Fuzzy Systems, vol.1, 85-97.