A Comparison between Two Fuzzy Clustering Algorithms for Mixed Features

Irene Olaya Ayaquica-Martínez and José F. Martínez-Trinidad

Instituto Nacional de Astrofísica, Óptica y Electrónica
Luis Enrique Erro No. 1
Santa María Tonantzintla, Pue. C.P. 72840, México
{ayaquica, fmartine}@inaoep.mx

Abstract In this paper, a comparative analysis of the mixed-type variable fuzzy c-means (MVFCM) and the fuzzy c-means using dissimilarity functions (FCMD) algorithms is presented. Our analysis is focused in the dissimilarity function and the way of calculating the centers (or representative objects) in both algorithms.

1 Introduction

Restricted unsupervised classification (RUC) problems have been studied intensely in Statistical Pattern Recognition (Schalkoff, 1992). The fuzzy c-means algorithm is based on a metric over a n-dimensional space. It has shown its effectiveness in the solution for many unsupervised classification problems.

The fuzzy c-means algorithm starts with an initial partition then it tries all possible moving or swapping of data from one group to others iteratively to optimize the objective measurement function. The objects must be described in terms of features such that a metric can be applied to evaluate the distance. Nevertheless, the conditions in soft sciences as Medicine, Geology, Sociology, Marketing, etc., are quite different. In these sciences, the objects are described in terms of quantitative and qualitative features (mixed data). For example, if we look at geological data, features such as age, porosity, and permeability, are quantitative, while others such as rock types, crystalline structure and facies structure, are qualitative. Likewise, missing data is common in this kind of problems. In these circumstances, it is not possible measure the distance between objects; only the degree of similarity can be determined.

Nowadays, the mixed-type variable fuzzy c-means algorithm (MVFCM) of Yang et al. (2003) and the fuzzy c-means using dissimilarity functions (FCMD) of Ayaquica (2002) (see also Ayaquica and Martínez (2001)) are the most recent works that solve the RUC problem when mixed data appear.

In this paper, the mixed-type variable fuzzy c-means and the fuzzy c-means using dissimilarity functions algorithms are analyzed. In addition, a comparison between them is made.

A. Sanfeliu and J. Ruiz-Shulcloper (Eds.): CIARP 2003, LNCS 2905, pp. 472–479, 2003.

2 Mixed-Type Variable Fuzzy C-Means Algorithm (MVFCM)

In this section, the mixed-type variable fuzzy c-means algorithm (MVFCM) of Yang et al. (2003) is presented. They proposed a dissimilarity function to handle symbolic and fuzzy features.

The dissimilarity function used to evaluate the dissimilarity between symbolic features is the function proposed by Gowda and Diday, with some modifications. According to Gowda and Diday, the symbolic features can be divided into quantitative and qualitative in which each feature can be defined by $d_p(A_k,B_k)$ due to position p, $d_s(A_k,B_k)$ due to span s and $d_c(A_k,B_k)$ due to content c (see Yang et al. (2003) for details).

a) Quantitative features $d(A_k,B_k) = d_p(A_k,B_k) + d_s(A_k,B_k) + d_c(A_k,B_k)$

b) Qualitative features $d(A_k,B_k) = d_s(A_k,B_k) + d_c(A_k,B_k)$

Let $A = m(a_1,a_2,a_3,a_4)$ and $B = m(b_1,b_2,b_3,b_4)$ be any two fuzzy numbers. The dissimilarity $d_f(A,B)$ is defined as $d_f^2(A,B) = \frac{1}{4}\left(g_-^2 + g_+^2 + (g_- - (a_3 - b_3))^2 + (g_+ + (a_4 - b_4))^2\right)$ where $g_- = 2(a_1 - b_1) - (a_2 - b_2)$ and $g_+ = 2(a_1 - b_1) + (a_2 - b_2)$.

Then, the dissimilarity function for both symbolic and fuzzy features is

$$d^2(X_j,A_i) = \sum_{\substack{k' \text{ of} \\ \text{symbolic}}} \left(\sum_p d^2\left(X_{jk'}, A_{k'p|i}\right) \cdot e_{k'p|i} \right) + \sum_{\substack{k \text{ of} \\ \text{fuzzy}}} d_f^2\left(X_{jk}, A_{ik}\right) \quad \textbf{(1)}$$

Let $\{X_1,...,X_n\}$ be a data set of mixed feature types. The MVFCM objective function is defined as $J(\mu,e,a) = \sum_{i=1}^{c}\sum_{j=1}^{n} \mu_{ij}^m d^2(X_j,A_i)$, where $d^2(X_j,A_i)$ is (1).

The equation to evaluate the membership degree is

$$\mu_{ij} = \left(\sum_{q=1}^{c} \frac{\left(d^2(X_j,A_i)\right)^{1/(m-1)}}{\left(d^2(X_j,A_q)\right)^{1/(m-1)}} \right)^{-1} \quad i = 1,...,c,\ j = 1,...,n \quad \textbf{(2)}$$

This algorithm evaluates two cluster centers, one for symbolic features as $A_{ik'} = \left[\left(A_{k'1|i}, e_{k'1|i}\right),...,\left(A_{k'p|i}, e_{k'p|i}\right)\right]$ where $A_{k'p|i}$ is the pth event of symbolic feature k' in cluster i and $e_{k'p|i}$ is the membership degree of association of the pth event $A_{k'p|i}$ to the feature k' in the cluster i.

The equation for $e_{k'p|i}$ is

$$e_{k'p|i} = \frac{\sum_{j=1}^{n} \mu_{ij}^m \cdot \theta}{\sum_{j=1}^{n} \mu_{ij}^m} \quad \textbf{(3)}$$

where $\theta \in \{0,1\}$ and $\theta = 1$ if the kth feature of the jth datum X_j consists of the pth event, otherwise $\theta = 0$. The membership of X_j in the cluster i is $\mu_{ij} = \mu_i(X_j)$.

For fuzzy features the center is calculated considering A_{ik} as the kth fuzzy feature of the ith cluster center with parametric form $A_{ik} = m(a_{ik1}, a_{ik2}, a_{ik3}, a_{ik4})$ where

$$a_{ik1} = \frac{\sum_{j=1}^{n} \mu_{ij}^m \left(8x_{jk1} - x_{jk3} + x_{jk4} + a_{ik3} - a_{ik4}\right)}{8\sum_{j=1}^{n} \mu_{ij}^m}, \quad \textbf{(4)}$$

$$a_{ik2} = \frac{\sum_{j=1}^{n} \mu_{ij}^{m} \left(4x_{jk2} + x_{jk3} + x_{jk4} - a_{ik3} - a_{ik4}\right)}{4\sum_{j=1}^{n} \mu_{ij}^{m}},$$

$$a_{ik3} = \frac{\sum_{j=1}^{n} \mu_{ij}^{m} \left(-2x_{jk1} + x_{jk2} + x_{jk3} + 2a_{ik1} - a_{ik2}\right)}{\sum_{j=1}^{n} \mu_{ij}^{m}},$$

$$a_{ik4} = \frac{\sum_{j=1}^{n} \mu_{ij}^{m} \left(2x_{jk1} + x_{jk2} + x_{jk4} - 2a_{ik1} - a_{ik2}\right)}{\sum_{j=1}^{n} \mu_{ij}^{m}}$$

MVFCM Algorithm

Step 1: Fix m and c. Give $\varepsilon>0$. Initialize a fuzzy c-partition $\mu^{(0)} = \left\{\mu_1^{(0)},...,\mu_c^{(0)}\right\}$. Set l=0.

Step 2: For symbolic feature k', compute i^{th} cluster center $A_{ik'}^{(l)} = \left[\left(A_{k'1|i}^{(l)}, e_{k'1|i}^{(l)}\right),...,\left(A_{k'p|i}^{(l)}, e_{k'p|i}^{(l)}\right)\right]$ using (3). For fuzzy feature k, compute i^{th} cluster center $A_{ik}^{(l)} = \left(a_{ik1}^{(l)}, a_{ik2}^{(l)}, a_{ik3}^{(l)}, a_{ik4}^{(l)}\right)$ using (4).

Step 3: Update $\mu^{(l+1)}$ using (2)

Step 4: Compare $\mu^{(l+1)}$ with $\mu^{(l)}$ in a convenient matrix norm.

IF $\left\|\mu^{(l+1)} - \mu^{(l)}\right\| < \varepsilon$, THEN STOP

ELSE $l = l$+1 and GOTO Step 2.

In this algorithm a dissimilarity function defined as the sum of the dissimilarity between symbolic features and the dissimilarity between fuzzy features is used to solve the mixed data problem. However, the dissimilarity for symbolic and fuzzy features is always computed using the expressions d_p, d_s, d_c and d_f respectively. In the practice, the manner for evaluating the similarity between feature values is not only in dependence of the nature of features. Also, the context or the problem must be considered. When d_p, d_s, d_c and d_f are used we are forcing to evaluate the dissimilarity always in the same form independently of the context or nature of the problem. A fixed function does not allow representing the criterion used by the specialist to compare these features in a determined context. Therefore if two features are of the same type, the manner of comparing them not necessarily must be the same.

In other hand, this algorithm evaluates two cluster centers, one for symbolic features and other one for fuzzy features. These cluster centers are fictitious elements, i.e. the cluster centers cannot be represented in the same space of the objects however they are used to classify the objects.

3 Fuzzy C-Means Algorithm Using Dissimilarity Functions (FCMD)

In this section, the fuzzy c-means algorithm using dissimilarity functions (FCMD) of Ayaquica and Martínez (2001) is presented.

Let us consider a clustering problem where a data set of n objects $\{O_1,O_2,...,O_n\}$ should be classified into c clusters. Each object is described by a set $R = \{x_1,x_2,...,x_m\}$ of features. The features take values in a set of admissible values D_i, $x_i(O_j) \in D_i$, i=1,2,...,m. We assume that in D_i there exists a symbol "?" to denote missing data.

Thus, the features can be of any nature (qualitative: Boolean, multi-valued, etc. or quantitative: integer, real) and incomplete descriptions of the objects can be considered.

For each feature a comparison criterion $C_i : D_i \times D_i \to L_i$ i=1,2,...,m is defined, where L_i is a totally ordered set. This function allows to evaluate the similarity between two values of a feature. In the practice, this function is defined in basis of the manner to compare or evaluate the similarity between two values of the feature. When features are numeric, it is usually used a norm or distance, but it cannot be the unique way to evaluate the similarity between values. Therefore, in the formulation proposed here, this function is a parameter that the user can define according to the problem.

Some examples of comparison criteria are:

1. $C_s\left(x_s(O_i), x_s(O_j)\right) = \begin{cases} 0 & \text{if } x_s(O_i) = x_s(O_j) \vee x_s(O_i) = ? \vee x_s(O_j) = ? \\ 1 & \text{otherwise} \end{cases}$

where $x_s(O)$ is the value of the feature x_s in the object O; 0 means that the values are coincident and 1 means that the values are different; "?" denote missing data. This is a Boolean comparison criterion.

2. $C_s\left(x_s(O_i), x_s(O_j)\right) = \begin{cases} 0 & \text{if } x_s(O_i) = ? \vee x_s(O_j) = ? \\ 1 & \text{if } x_s(O_i), x_s(O_j) \in A_{s1} \\ 2 & \text{if } x_s(O_i), x_s(O_j) \in A_{s2} \\ \vdots & \\ k-1 & \text{if } x_s(O_i), x_s(O_j) \in A_{sk-1} \end{cases}$

where $A_{s1} \cup,...,\cup A_{sk-1} = D_s$. This is a k-value comparison criterion.

3. $C_s\left(x_s(O_i), x_s(O_j)\right) = \left|x_s(O_i) - x_s(O_j)\right|$ this is a comparison criterion that takes real values.

In this way, it is not fixed a unique comparison criterion for all problems to solve, but fairly we give the liberty of using the comparison criterion, which more reflects the manner that the objects are compare in the practice. Note that the dissimilarity functions defined by Yang et al. (2003) for quantitative and qualitative features may be used too.

In addition, let $\Psi : (D_1 \times ... \times D_n)^2 \to [0,1]$ be a dissimilarity function. This function allows evaluating the dissimilarity between object descriptions. Thus, this function is given in dependence of comparison criteria.

Let

$$\Psi(O_i, O_j) = \sum_{x_s \in R} C_s\left(x_s(O_i), x_s(O_j)\right) \Big/ |R| \qquad \textbf{(5)}$$

be the dissimilarity between the objects O_j and O_k. The value $\Psi(O_j, O_k)$ satisfies the following three conditions:

1. $\Psi(O_j, O_k) \in [0,1]$ *for* $1 \le j \le n,\ 1 \le k \le n$
2. $\Psi(O_j, O_j) = 0$ *for* $1 \le j \le n$
3. $\Psi(O_j, O_k) = \Psi(O_k, O_j)$ *for* $1 \le j \le n,\ 1 \le k \le n$

Let u_{ik} the degree of membership of the object O_k in the cluster K_i, and let $R^{c \times n}$ be the set of all real $c \times n$ matrices. Any fuzzy c-partition of the data set is represented by a matrix $U=[u_{ik}] \in R^{c \times n}$. The fuzzy c-partition matrix U is determined from

minimization of the objective function given by $J_m(U,\vartheta)=\sum_{k=1}^{n}\sum_{i=1}^{c}u_{ik}\Psi(O_k,O_i^*)$ where ϑ is a set of representative objects, one for each K_i, and $\Psi(O_k,O_i^*)$ is the dissimilarity between the object O_k and the representative object O_i^* of K_i. In the case of classical fuzzy c-means ϑ are the centers for the clusters and Ψ is the Euclidean distance.

Since in our algorithm the objects descriptions are not only in terms of quantitative features the mean cannot be computed. Then instead of use a center (or centroid) for a cluster we will use an object in the sample as representative for this cluster. In order to determine a *representative object* O_i^* for the cluster K_i, i=1,...,c we proceed as follows:

We consider the crisp subset K_i' of objects that have their maximum degree of membership in this cluster K_i. Then the representative object of cluster K_i is determined as the object O_i^* that satisfies

$$O_i^* = \min_{q\in K_i'}\left\{\sum_{k=1}^{r}u_{ik}\Psi(O_k,O_q)\right\} \quad \textbf{(6)}$$

The object O_i^* may be not unique, then we take the first object found.

Note that our algorithm considers as representative object one object of the sample instead of one fictitious element as occurs in the MVFCM algorithm.

In order to determine the degree of membership of the object O_k to the cluster K_i, we define for each object O_k the sets $I_k=\{i/1\le i\le c;\ \Psi(O_k,O_i^*)=0\}$ This set contains the indexes of the clusters such that the dissimilarity between the object to classify O_k and the representative objects O_i^*, i=1,..,c, is zero. And $\bar{I}_k=\{1,2,...,c\}-I_k$ in this set are those indexes of the clusters such that the dissimilarity between O_k and O_i^*, i=1,..,c, is greater than zero.

Thus, the degree of membership of O_k to K_i is computed via (7a) or (7b).

$$I_k=\varnothing \ \Rightarrow u_{ik}=\frac{1}{\sum_{j=1}^{c}\left(\frac{\Psi(O_k,O_i^*)}{\Psi(O_k,O_j^*)}\right)^2} \quad \textbf{(7a)}$$

We can see that the degree of membership u_{ik} increases if simultaneously the dissimilarity between O_k and O_i^* for K_i decreases and the dissimilarity between O_k and O_j^* for K_j, j=1,..,c, increases (and vice versa).

$$I_k\neq\varnothing \ \Rightarrow u_{ik}=0\ \forall i\in\bar{I}_k \text{ and } \sum_{i\in I_k}u_{ik}=1 \quad \textbf{(7b)}$$

The equation (7b) is the alternative form for O_k when $\exists i\in I_k$ so that $\Psi(O_k,O_i^*)$=0. The membership of O_k to the clusters K_i (u_{ik}) $i\in I_k$ will be $\frac{1}{|I_k|}$, i.e., the degree of membership is distributed among the clusters K_i, $i\in I_k$. In addition, for the clusters $i\in\bar{I}_k$ we assign zero as degree of membership.

FCMD Algorithm

Step 1. Fix c, $2\le c\le n$. l = 0

Step 2. Select c objects in the data as representative objects, $\vartheta^{(l)}$.

Step 3. Calculate the c-partition $U^{(l)}$ using (7a) and (7b).

Step 4. Determine the representative objects for each fuzzy cluster using (6), $\Im^{(l+1)}$.
Step 5. If $\Im^{(l)} = \Im^{(l+1)}$ then STOP

Otherwise $l=l+1$ and go to step 3.

An important point to highlight is that this algorithm has the flexibility that uses a dissimilarity function, which is defined in terms of comparison criteria. The comparison criteria allow expressing the way in which features values are compared depending of the problem context to solve.

This algorithm, unlike the MVFCM algorithm, evaluates a unique "cluster center" called the *representative object*, which is an object of the sample instead of one fictitious element as occurs in the MVFCM algorithm. It is more reasonable consider an object of the data set to classify as representative object instead of using an element that cannot be represented in the same space of the objects.

4 Analysis

In this section, the *c*-partitions generated by both algorithms are analyzed. The analysis is based on the manner to calculate the membership degrees and the way to calculate the cluster centers (representative objects).

In order to make the analysis, the data set shown in Table 1 was used. There are 10 brands of automobiles from four companies: Ford, Toyota, China-Motor and Yulon-Motor in Taiwan. In each brand, there are six feature components –company, exhaust, price, color, comfort and safety. In the color feature, the notations W=white, S=silver, D=dark, R=red, B=blue, G=green, P=purple, Gr=grey and Go=golden are used. The features: company, exhaust and color are symbolic, the feature price is real data and the features comfort and safety are fuzzy. The obtained results are shown in Table 2.

In this experimentation for FCMD Ψ was used as dissimilarity function and as comparison criteria the functions $d(A_k,B_k)$ defined by Yang for quantitative and qualitative features were used. In other words, the same criteria for features were used in both algorithms.

The results shown in table 2 for MVFCM were taken from Yang et al. (2003).

Table 1. Data set of automobiles

No.	Brands	Company	Exhaust (L)	Price (NT$10000)	Color	Comfort	Safetiness
1	Virage	China-Motor	1.8	63.9	W,S,D,R,B	[10,0,2,2]	[9,0,3,3]
2	New Lancer	China-Motor	1.8	51.9	W,S,D,R,G	[6,0,2,2]	[6,0,3,3]
3	Galant	China-Motor	2.0	71.8	W,S,R,G,P,Gr	[12,4,2,0]	[15,5,3,0]
4	Tierra Activa	Ford	1.6	46.9	W,S,D,R,G,Go	[6,0,2,2]	[6,0,3,3]
5	M2000	Ford	2.0	64.6	W,S,D,G,Go	[8,0,2,2]	[9,0,3,3]
6	Tercel	Toyota	1.5	45.8	W,S,R,G	[4,4,0,2]	[6,0,3,3]
7	Corolla	Toyota	1.8	74.3	W,S,D,R,G	[12,4,2,0]	[12,0,3,3]
8	Premio G2.0	Toyota	2.0	72.9	W,S,D,G	[10,0,2,2]	[15,5,3,0]
9	Cerfiro	Yulon-Motor	2.0	69.9	W,S,D	[8,0,2,2]	[12,0,3,3]
10	March	Yulon-Motor	1.3	39.9	W,R,G,P	[4,4,0,2]	[3,5,0,3]

As the dissimilarities matrices are symmetric then triangular matrices are shown in a unique matrix in the Table 3, where a) is calculated using the expression (1) and b) is calculated using the expression (5). The values in b) are values of Ψ normalized in [0,1] .

Table 2. Clusters obtained with MVFCM and FCMD for mixed data.

Data	MVFCM		FCMD	
	μ_{1j}	μ_{2j}	μ_{1j}	μ_{2j}
1	**0.9633**	0.0367	**0.9215**	0.0785
2	**0.9633**	0.0367	0.0000	**1.0000**
3	**0.9951**	0.0049	**0.9959**	0.0041
4	0.0966	**0.9034**	0.0020	**0.9980**
5	**0.9951**	0.0049	**0.9561**	0.0439
6	0.0135	**0.9865**	0.0047	**0.9953**
7	**0.9633**	0.0367	**0.9959**	0.0041
8	**0.9951**	0.0049	**0.9978**	0.0022
9	**0.9951**	0.0049	**1.0000**	0.0000
10	0.0185	**0.9815**	0.0258	**0.9742**

Table 3. Dissimilarities matrices.

a)

0.0	0.0213	0.0132	0.0397	0.0006	0.0471	0.0156	0.0148	0.0062	0.0820
676.1	0.0	0.0646	0.0031	0.0220	0.0055	0.0725	0.0677	0.0460	0.0205
420.8	2046.0	0.0	0.0929	0.0133	0.1025	0.0023	0.0009	0.0041	0.1517
1257.2	101.1	2943.3	0.0	0.0412	0.0009	0.1039	0.0974	0.0719	0.0085
19.3	698.3	423.5	1305.5	0.0	0.0480	0.0152	0.0138	0.0047	0.0839
1492.9	175.1	3248.4	31.0	1520.5	0.0	0.1142	0.1073	0.0801	0.0059
494.7	2297.0	73.0	3293.1	482.5	3619.2	0.0	0.0025	0.0046	0.1666
471.1	2145.9	31.2	3086.2	438.3	3400.0	79.7	0.0	0.0031	0.1582
197.4	1457.4	131.8	2277.8	149.6	2538.1	147.8	99.8	0.0	0.1259
2596.7	649.5	4807.1	269.3	2657.6	187.2	5277.9	5011.9	3987.4	0.0

b)

The fuzzy c-means algorithm has as main characteristic that builds clusters where objects with low dissimilarity obtain high membership degree into the same cluster while objects that are relatively distinct obtain high membership degree into different clusters.

The object 2 obtains high membership degree into the cluster 1, but it has low dissimilarity with objects having high membership degree to the cluster 2 (see Table 3), i.e. the description of the object 2 is more similar with the description of the objects 4, 6 and 10 (see Table 2). Therefore, the object 2 should have high membership degree to the cluster 2. So the MVFCM algorithm does not build clusters with the characteristic above mentioned. The membership degrees for MVFCM are calculated using the expression (2). This expression is in function of the dissimilarity between the object to be classified and the cluster centers. In the example, the object 2 obtains high membership to the cluster 1 because it is less dissimilar with the center of cluster 1 than the center of the cluster 2. So that the cluster centers play a determinant role in these dissimilarity values; therefore the obtained c-partitions depend of these centers.

The FCMD, unlike the MVFCM, builds clusters, which satisfy the characteristic above mentioned. So the object 2 obtains high membership degree to the cluster 2. The membership degrees for FCMD are evaluated using the expression (7a) and (7b). These expressions also are in function of the dissimilarity between the object to be classified and the representative objects. But in this case, the representative objects are objects in the sample. So, if two objects are low dissimilar with the representative object, then they must be low dissimilar between them. In this example, the object 2 is just the representative object of the cluster 2.

In addition, the objects 1, 2 and 7 obtain the same membership degree to the cluster 1, and then according to the fuzzy c-means algorithm classification strategy, the

descriptions of these objects must be similar or equal. However, the dissimilarities between the objects 1, 2 and 7 are very different; the dissimilarity between 1 and 2 is 676.1, the dissimilarity between 2 and 7 is 2297.0 and the dissimilarity between 1 and 7 is 494.7 (see Table 3). This shows that, the manner in which the cluster centers are calculated in the MVFCM algorithm determines that objects having low dissimilarity with the cluster center can be very dissimilar among them. In the case of FCMD algorithm, the object 2 has a high membership degree to the cluster 2 and the objects 1 and 7 both have different membership degree to the cluster 1 (see Table 2).

When objects have the same membership degree to a cluster for FCMD algorithm, for example, objects 3 and 7 in the cluster 1; the dissimilarity between them is 73.0357, very low (see Table 3). Again this situation occurs because the representative object is an object in the sample.

5 Conclusions

The FCMD algorithm allows using comparison criteria defined by the specialist according to the specific context of a practical problem. In addition, this algorithm evaluates "cluster centers" called representative objects, which are objects in the sample instead of a fictitious element as occurs in the MVFCM algorithm. Also, as we can observe in the definition of comparison criteria, the symbol "?" was introduced to denote missing data, then the FCMD algorithm allows working with databases that contain incomplete descriptions of objects.

We can observe that the MVFCM algorithm builds clusters containing objects which have high membership degree to a cluster but with low dissimilarity with objects belonging with high membership degree to other clusters. On the other hand, the FCMD algorithm builds clusters where the objects with high membership degree to a cluster have low dissimilarity among them.

Based on the analysis made we can say that the FCMD algorithm is a more flexible alternative in the solution of fuzzy unsupervised classification problems where mixed and missing data appear.

Acknowledgement. This work was financially supported by CONACyT (Mexico) through project J38707-A.

References

1. Schalkoff R. J. (1992): Pattern Recognition: Statistical, Structural and Neural approaches, John Wiley & Sons, Inc. USA.
2. Yang M. S., Hwang P. Y. and Chen D. H. (2003): Fuzzy Clustering algorithms for mixed feature variables, Fuzzy Sets and Systems, in Press.
3. Ayaquica M. I. and Martínez T. J. F. (2001): Fuzzy c-means algorithm to analyze mixed data. VI Iberamerican Symp. on Pattern Recognition. Florianopolis, Brazil. pp. 27-33.
4. Ayaquica M. I. (2002): Fuzzy c-means algorithm using dissimilarity functions. Thesis to obtain the Master Degree. Center for Computing Research, IPN, Mexico. In Spanish.
5. Gowda K. C. and Diday E. (1991): Symbolic clustering using a new dissimilarity measure, Pattern Recognition 24 (6) pp. 567-578.