Asian Journal of Infomation Technology 4 (11): 981-985, 2005

The Fuzzy Possibilistic C-Means Classifier

[1]H. Boudouda, [1,2]H. Seridi and [2,3]H.Akdag
[1]Laig, Université 08 mai 1945, Guelma 24000, Algérie
[2]Crestic-leri, Université de Reims Champagne Ardenne, France
[3]Lip 6 Université P. & M. Curie, Paris, France

Abstract: In front of the mass of information which does not cease growing in an exponential way, the human expert is often confronted to data classification problems in the pattern recognition domain. The methods of classification are generally the result of a formalism based on an artificial reasoning, which is at least close to that of a human reasoning. The various approaches suggested in literature, differ the ones from the others by the membership concept of an object to a class; however the initialization method remains ambiguous. In this same study present a new approach of unsupervised automatic classification under the C-Means family. This new approach based on the fusion of fuzzy and the possibility theory, allows on the one hand to solve, simultaneously the problem of coincidence and the noise and on the other hand to accelerate classification. The initialization methodology used in this study is based on probabilistic membership matrix. To show the performances of this new approach, tests were carried out on the Iris data basis.

Key Words: Classification, unsupervised training, pattern recognition, fuzzy logic, approximate reasoning

INTRODUCTION

The goal of an automatic classification is to find structures within a set of objects to be classified by formulation of groups. Each object is represented in the attributes space by a vector called attributes vector. The unsupervised classification by C-Means is a procedure, which consists in identifying and gathering similar objects in the same class using a distance measurement between each object and a prototype characterizing each class. Indeed, the

C-Means methods seek to find representative prototypes of a set of objects. There exist several algorithms of unsupervised classification by C-Means. Each algorithm has a characteristic of membership, which distinguishes it from the others, Hard[1], Fuzzy[2,3]or Possibilistic[4,6]. An algorithm of classification is optimal when it converges to a partition, which verifies a good property of clustering. This stake requires an adequate initialization of the classifier. One distinguishes two methods of initialization: Initialization by gravity centers, or by membership matrix. To have an optimal classification, we need on the one hand a robust concept of membership which corrects simultaneously the problem of overlapping, noise and coincidence and on the other hand an adequate initialization, which may not generate a set of poor final centers such as dead centers, redundancy centers or local minima. In the literature there exists several methods of initialization and there is no simple or universally good solution to this problem, for example: The mean shift algorithm[7], the capture affect neural network algorithm[8], the strong forms method or the randomly method. Taking into account the improvements made by the various methods of classification and initialization, we propose an original approach based on the fusion of the two concepts fuzzy and possibilistic called the FPCM[9] with a new initialization method by probabilistic membership matrix[10].

CLASSIFICATION METHODS BY C-MEANS

In what follows three methods of unsupervised automatic classification will be presented:

Fuzzy C-MEANS algorithm: FCM: The fuzzy C-Means algorithms are based on the concept of fuzzy sub-set within the meaning of Zadeh[2], to solve the problems of the classes, which are badly definite at the borders. A fuzzy partition is defined by:

$$\begin{cases} \forall i,k \quad 1 \le i \le c, \ 1 \le k \le n \quad u_{ik} \in [0,1] \\ \forall i \quad 1 \le i \le c \quad 0 < \sum_{k=1}^{n} u_{ik} < n \\ \forall k \quad 1 \le k \le n \quad \sum_{i=1}^{c} u_{ik} = 1 \end{cases}$$

Corresponding Author: H. Boudouda, Laig, Université 08 mai 1945, Guelma 24000, Algérie

In this study we present the FCM algorithm [3], which minimizes the least squares function (WGSS) defined by:

$$J^f_m(U,V;X)=\sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m d_{ik}^2 \quad (1)$$

Where $m\in]1,\infty[$ is the value, which characterizes the fuzzy in the partition.

$J_m(U,V;X)$ is a global minimum if and only if :

$$u_{ik}=\frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}} \quad (2)$$

And $$V_i=\frac{\sum_{k=1}^{n}(u_{ik})^m x_k}{\sum_{k=1}^{n}(u_{ik})^m} \quad (3)$$

With $$\sum_{i=1}^{c}u_{ik}=1 \quad (4)$$

Possibilistic C-MEANS algorithm: PCM: The possibilistic approach was proposed by Krishnapuram and Keller[6], to surmount limited algorithms of fuzzy classification in the presence of noise, they released the constraint of probabilistic inspiration and proposed a new partition, which is defined in the following manner:

$$\begin{cases}\forall i,k\ \ 1\le i\le c,\ 1\le k\le n & u_{ik}\in[0,1]\\ \forall i\ \ 1\le i\le c\ \ 0<\sum_{k=1}^{n}u_{ik}\le n & \\ \forall k\ \ 1\le k\le n\ \ \max u_{ik}\succ 0 & \end{cases}$$

The objective function to be minimized is the following one:

$$J^p_m(U,V;X)=\sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m d_{ik}^2+\sum_{i=1}^{c}\eta_i\sum_{k=1}^{n}(1-u_{ik})^m \quad (5)$$

Where η_i is a homogeneous positive number at a square distance. The minimization of J_m makes a modification in the calculation of the degrees membership:

$$u_{ik}=\frac{1}{1+\left(\frac{d_{ik}^2}{\eta_i}\right)^{\frac{1}{m-1}}} \quad (6)$$

With $$\eta_i=\frac{k\left(\sum_{k=1}^{n}(u_{ik})^m d_{ik}^2\right)}{\sum_{k=1}^{n}(u_{ik}^m)} \quad (7)$$

K: is often chosen equal to 1.

Fuzzy- possibilistic C-MEANS algorithm FPCM

Fuzzy–possibilistic memberships: The suggested approach[9], called Fuzzy-possibilistic C-Means FPCM, is based on the fusion of fuzzy and the possibility concepts, so that an object which is likely to be in a class with the possibility sense is assigned in the latter with a fuzzy degree, while its memberships to the remaining classes is evaluated by possibilistic degrees. The FPCM approach allows to solve, simultaneously the overlapping[1], the coincidence[9,11] and the noise[9,12] problems. The objective function to be minimized in this case is based on the fusion of the fuzzy and the possibilistic memberships concept:

$$J^{fp}_m(U,V;X)=J^f_m(U,V;X)+J^p_m(U,V;X) \quad (8)$$

The FPCM membership concept is defined by the following three rules:

R$_1$: If an object belongs to the influence zone of a class then one assigns it to the latter with a fuzzy degree and to the others with Possibilistics degrees.

R$_2$: If an object belongs to the intersection of two or several influences zones, one assigns it with a fuzzy degree to each overlapping class.

R$_3$: If an object does not belong to any influence zone, one assigns it with a possibilistic degree to each class.

Initialization methodology: Contrary to the initialization methods which are based on the research of the gravity centers, the proposed initialization[10] method seeks to find a membership matrix by respecting the probabilistic inspiration constraint. The membership matrix degrees are computed by the following equations:

$$u(i,k)=\begin{cases}u(i,k) & \forall i=1:c/i\ne j,\forall k=1:n\\ u_1(i,k) & \forall j\in\{1,2,\ldots,c\},\forall k=1:n\end{cases} \quad (9)$$

Where: $u(i,k)=1/(c+\beta(i,k))$ represente the intial membership matrix;

$\beta(i,k)=i+k-2$ the regularization term

$$u_1(j,k)=1-\sum_{i=1}^{c}u(i,k)+u(j,k),\ \forall j\in\{1,2,\ldots,c\},\ i=1:c,\ k=1:n$$

the reinforced memberships.

FPCM algorithm: The general form of fuzzy-possibilistic clustering algorithm is presented as follows:

$X=\{x_1,x_2,...x_n\}\subset R^p$ is the set of features vectors to be classified.

Fix the classes number $c\in]1,n]$, the fuzzifier $m\in]1,\infty[$ and $\varepsilon>0$;
Set iteration counter l=0;

Estimate $u_{ik(0)}$ using Eq. 9

- Update the C-prototypes v_i using Eq . 3

Estimate u_{ik} using Eq . 2
Compare $u_{i,k(l)}$ and $U_{i,k(l-1)}$ if $\|u_{i,k(l)}-u_{i,k(l-1)}\|<\varepsilon$ then STOP
Else $l\leftarrow l+1$; GO TO 1:
End if
Set iteration counter l=0;
$u_{i,k(0)}\leftarrow u_{i,k}$, $v_{(0)}\leftarrow v_i$;
Estimate $\eta_{i(0)}$ using $u_{i,k(0)}$;

- Compute $u_{i,k}$ using Eq. 2 and Eq. 6;

Estimate η_i using Eq. 7;
Update the C-prototypes v_i using Eq. 3;
Compare $u_{i,k(l)}$ and $\|u_{i,k(l)}-u_{i,k(l-1)}\|<\varepsilon$ ifthen STOP

Else $l\leftarrow l+1$; GO TO 2:
End if

EXPERIMENTATION

To appreciate the performances of algorithms FCM, PCM and FPCM in both cases of initialization (initialization by membership matrix or by gravity centers), we carried out tests on the Iris data basis, which is composed of 150 flowers described by 4 variables (length and width of sepals, and petals). The number of classes is equal to 3, the objects are uniformly divided into three classes, classes 2 and 3 are easily separable from class 1, but not easily separable between them. The results obtained by three algorithms FCM, PCM and FPCM change according to the choice of the algorithm initialization.

Initialization by centers of gravity: In order to initialize the algorithm by centers of gravity, the applied method consists in carrying out several times the algorithm based on different initial centers. If stable centers are obtained, in a repeated way, from one repetition to another, they allowed us to consider them as reliable. By applying this procedure, one obtains the following centers:

$$V_{FCM}=\begin{pmatrix}5.11 & 5.22 & 1.71 & 1.41\\ 5.53 & 6.51 & 4.63 & 2.15\\ 4.80 & 3.53 & 5.40 & 7.50\end{pmatrix}$$

The tests carried out on the Iris data basis with the three algorithms of classification are given in table1.

The classification rate obtained by the FPCM is equal to 97.33% with an iterations number equal to 365 and minimal cost of $2.40.10^{-1}$, however the PCM reaches 94.00% after 403 iterations with a satisfaction index of 9.42.10, while the FCM reaches a success rate of 92.00 % after 387 iterations with a satisfaction index of $4.32.10^{-6}$.

Table 2 gives the confusion matrix corresponding to the results of classification for each algorithm.

Table 2 shows that the three algorithms recognize 100% the first class (C1) and make errors during the classification of objects of the two remaining classes (C2 and C3). The total error which corresponds respectively to FCM, PCM, and the FPCM is 11, 9 and 4.

Initialization by membership matrix: One initializing the various algorithms by the proposed membership matrix, the classification results of the Iris data basis with the three classification algorithms are given in Table 3.

By analyzing the results of table 3, one notices that the iterations number has largely decreased and the rates of classification and the satisfaction index remain unchangeable compared to the case of initialization by centers of gravity.

In Table 4, gives the final centers generated by the FCM, PCM and FPCM in the case of initialization by membership matrix.

Table 4 shows that the PCM generates classes having the same centers of gravity (i.e. $V_2=V_3$); while the FPCM and FCM lead to a good separation of the centers.

The curves represented by Fig.e 1, illustrate the variation of the classification success rate and the satisfaction index according to the fuzzification factor m for the FCM, PCM and FPCM in the case of initialization by membership matrix. For a better data classification, one chooses the parameter m, which controls the fuzzy into classification so that the success rate is maximum and the quadratic error is minimum. By analysing the curves,

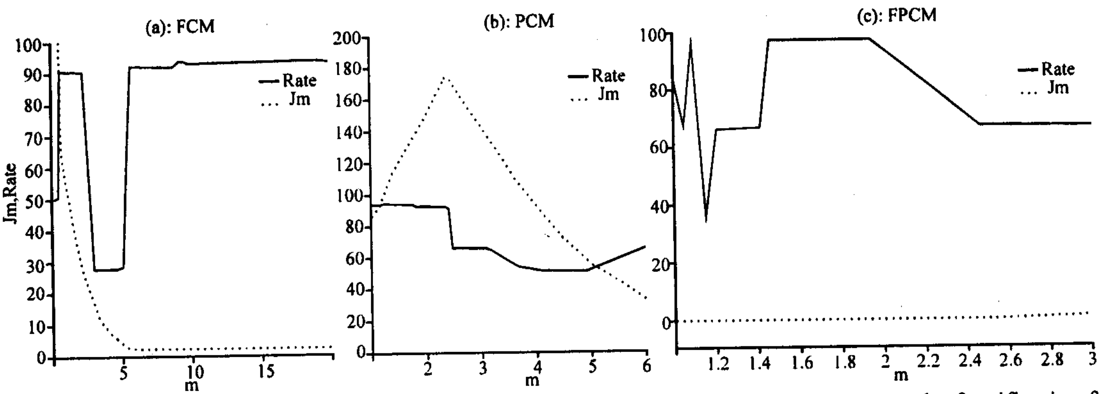

Fig.1 a,b,c: Variation of the classification success rate and the satisfaction index according to the fuzzification factor for the FCM (Fig.1a), PCM (Fig.1b) and PCM (Fig.1c) using initialization by membership matrix.

Table1 : comparison between various results of classification obtained by the algorithms FCM, PCM, FPCM

Algorithms of Classification	Success rate (%)	Satisfaction index (J^m)	Iterations numbers for the research of The initial partition	Iterations numbers for the convergance	Total iterations No
FCM	92.66	$4.32.10^{-6}$	00	34	34
PCM	94.00	9.42.10	00	51	51
GPCM	97.33	$2.40.10^{-1}$	00	13	13

Table 2: Confusion matrix

Classes	C_1	C_2	C_3	C_1	C_2	C_3	C_1	C_2	C_3
C_1	50	0	0	50	0	0	50	0	0
C_2	0	45	8	0	42	8	0	47	3
C_3	0	6	44	0	1	49	0	1	49

Table 3: Comparison between various results of classifications obtained by the three algorithms FCM, PCM and FPCM using initialization by centers of gravity

Algorithms of classification	Success rate (%)	Satisfaction index (J_m)	Iterations numbers for the research of Initial partition	IterationsNumbers for the convergence	Total iterations No
FCM	92.66	$4.32.10^{-6}$	352	35	387
PCM	94	$9.42.10^{0}$	352	51	403
FPCM	97.33	$2.40.10^{-1}$	352	13	365

Table 4: Comparison between various gravity centers obtained by the three algorithms FCM, PCM and FPCM

Centers	V1	V2	V3
FCM	[5.02 3.39 1.51 0.25]	[6.02 2.87 4.48 1.46]	[6.49 2.99 5.24 1.89]
PCM	[5.06 3.43 1.46 0.24]	[6.17 2.88 4.76 1.60]	[6.17 2.88 4.76 1.60]
FPCM	[5.10 3.45 1.45 0.20]	[6.10 2.95 4.65 1.40]	[6.05 3.00 4.85 1.80]

Table 5: Optimal values of m for the FCM, PCM and the FPCM

Algorithms	m	Rate (%)	J^m
FCM	17.697	92.66	$4.32.10^{-6}$
PCM	1.823	94.00	9.42.10
FPCM	1.230	97.33	$2.40.10^{-1}$

3the values of m which correspond to a minimal satisfaction index (J_m) and a maximum rate of success for each algorithm of classification are given in Table 5.

CONCLUSIONS

The experimental tests carried out on the Iris data 1.0 basis, allowed us to conclude that the FPCM is definitely higher than FCM and PCM. The computation using classical methods which are based on the determination of a hard partition, are confronted to overlapping classes problem (in our case the second and the third) and thus lead to a bad classification. To resolve this difficulty, the FCM proposes the sharing concept of an object between the various classes by respecting the probabilistic constraint but remains nevertheless sensitive to the noises witch degrades the performances of the classifier. To overcome the noise problem, the PCM proposes the substitution of the sharing concept by that of typicality. Although the PCM has satisfactory results, this last leads to classes having the same centers of gravity. The FPCM which is based on the fusion of the FCM and PCM, is robust in the presence of the noise, generates separable centers, solves the problem of overlapping and converges quickly. Although initialization by centers of gravity gives satisfactory results, it remains a random method and it is very expensive, on the other hand the implementation of the suggested initialization is simple and permits to eliminate the search time of the initial partition and consequently accelerate the classification.

REFERANCES

1. Bezdek, J.C., 1993. A review of probabilistic, fuzzy, and neural models for pattern recognition, J. Intell. Fuzzy Systems,1: 1-25.
2. Zadeh, L.A., 1965. Fuzzy sets. Inform. Control, 8: 338-353.
3. Bezdek, J.C., 1981. Pattern recognition with Fuzzy objective function algorithms, plenum press, New - York.
4. Dubois, D., H. Prade, 1987. Théorie des possibilités. Application à la Représentation des connaissances en Informatique, Masson, Paris.
5. Zadeh, L.A., 1978. Fuzzy sets as basic for the theory of possibility, Fuzzy sets and Systems, 1: 3-28.
6. Krishnapuram, R. and J.M. Keller, 1996. The possibilistic C-means algorithm: Insights and recommendations, IEEE Trans. Fuzzy systems, 4: 385-396.
7. Cheng, Y. and M. Shift, 1995. Mode seeking and clustering, IEEE trans. Pattern analysis and Machine Vision, 17: 790-799.
8. Firenze, F. and P. Morasso, 1993. The capture effect model: A new approach to self-organizclustering, In Proceedings of the Sixth International Conference on Neural Networks and their Industrial and Cognitive Applications and Exhibition Catalog, NEURO-NIMES 93 Conference, France, pp: 45–54.
9. Boudouda, H. and H. Seridi, 2004. Une nouvelle approche De classification automatique non supervisée Par C-Means : Fusion Des algorithmes Flou Et Possibiliste ", dans proc. of 2ème conférence internationale : Sciences electronique, Technologies de l'information et des Télécommunications (IEEE- SETIT'2004), Tunisie, pp: 163.
10. Boudouda, H., M. Nemissi, H. Seridi and H . Akdag, 2005. Unsupervised automatic classification by c-means: Resolution of initialisation problem, Al-Azhar University Engineering Journal, AEUJ, 8: 72-79.
11. Bezdek, J.C., 1995, optimisation of clustering criteria by Reformulation, IEEE Transaction of Fuzzy Systems, 3: 241-245,
12. Dave, R.N., 1992. Characterization and detection of noise in clustering, pattern recognition Lett. , vol. 12, no. 11, pp: 657-664.