

**

*

A Suggested Algorithm for the Selection of Initial Centers of Clusters

ABSTRACT

In this research, we suggest an algorithm to select the initial centers of clusters; this algorithm includes the selection of features of different attributes as an initial center of each cluster. As a fruit of this algorithm it gives a unique cluster when processes any group of data many times. This result leads to save time and effort. Simulation experiments performed to test the proposed algorithm. Also, a real data deals with the food security of Arabic World is applied.

كلية علوم الحاسبات والرياضيات/جامعة الموصل / *

كلية علوم الحاسبات والرياضيات/جامعة الموصل / **

تاريخ التسلم: 2008/12/3 تاريخ القبول: 2009/ 5/ 26

: .1

Data

Grouping

.(Dunham,2004 ; Kumar and Rishan 2002)

(Jantzen,2004 ; Jermyn

.and et al., 2001)

:
(2006) .1

(2008)

.2

: .2

Subsets

p

[p/c]

[p/c]

c

.c

p

:

D

$$D = \begin{bmatrix} x_{11} & x_{12} & \cdot & \cdot & \cdot & x_{1p} \\ x_{21} & x_{22} & \cdot & \cdot & \cdot & x_{2p} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{n1} & x_{n2} & \cdot & \cdot & \cdot & x_{np} \end{bmatrix}$$

$$\begin{array}{ccc}
 c & .n > p & n \\
 c & 2 \leq c \leq p & c \\
 s = [p/c] & (&) \\
 .n * s & & .
 \end{array}$$

$$j = 1, 2, \dots, c \quad j$$

:

$\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_s$

$$S = \begin{bmatrix}
 x_{11} & x_{12} & \dots & \dots & x_{1s} \\
 x_{21} & x_{22} & \dots & \dots & x_{2s} \\
 \dots & \dots & \dots & \dots & \dots \\
 x_{n1} & x_{n2} & \dots & \dots & x_{ns}
 \end{bmatrix}$$

i

$$\mathbf{x}_i = [x_{i1} \quad x_{i2} \quad \dots \quad x_{is}]; \quad i = 1, 2, \dots, n$$

S

$$\begin{array}{ccc}
 . x^{(i1)} < x^{(i2)} < \dots < x^{(is)} & \mathbf{x}^*_{i.} = [x_1^{(i1)} & x_2^{(i2)} & \dots & x_s^{(is)}] \\
 : & c_j & j
 \end{array}$$

$$c_j = \begin{cases}
 x^{\left(\frac{s+1}{2}\right)} & \text{if } s \text{ is odd} \\
 \frac{x^{\left(\frac{s}{2}\right)} + x^{\left(\frac{s+2}{2}\right)}}{2} & \text{if } s \text{ is even}
 \end{cases} \quad ; \text{for } j = 1, 2, \dots, c \quad (1)$$

j

c_j

:

[p/c]

:
:
:
:(1)

.3

Pentium

MATLAB 7.5

X_1, X_2, \dots, X_9

9 -1

X_2, X_4, X_6, X_8

Normal Distribution

$X_i = \text{normrnd}(a,b)$; $i=2,4,6,8$

.1 20

.a=20,b=1

X_1, X_3, X_5

Gamma Distribution

$X_i = 7 + \text{gamrnd}(a,b)$; $i = 1,3,5$

10

.a=10,b=1/√a

.1

Beta Distribution

X_7, X_9

100

HCM

.(Hard C-Means)(Albayrak and Amasyali,2003 ;Jantzen

X_1, X_2, \dots, X_9

:

$C_1 = \{ X_2, X_4, X_6, X_8 \}$

$C_2 = \{ X_1, X_3, X_5 \}$

$C_3 = \{ X_7, X_9 \}$

(1)

100 HCM

.(n=1000) (n=500) (n=100)

(%) :(1)

معدل زمن التنفيذ (ثانية)	معدل عدد التكرارات	النسب المئوية لعدد حالات النجاح (%)	خوارزمية HCM
0.0192	1.5300	56	n = 100
0.0192	1.5100	63	n = 500
0.0257	1.6100	67	n = 1000

HCM

100

,0.2350 ,0.2030

.1

n=1000,n=500,n=100

0.2550

:

(2006

) 2004-1998

)

.(1

) (2006

(2008)

.5

:

C1={ X12 }

C2={ X13}

C3={X3 , X4 , X7 , X 9 , X 17}

C4={X5 , X16}

C5={X1 , X2 , X6 , X 8 , X 10 , X11 , X14 , X15 , X 18 }

.(1)

-

X12 : C1

%0.00007

.((1))

X13 : C2

%0.02

.((1))

X3 , X4 , X7 , X 9 , X 17 : C3

%82

%360

%60.5

%72

%48.7

)

.((2) (1)

: X5 , X16 : C4

.1

.2

.3

%25

%14

%100

%41.5

X1, X2, X6, X 8 , X 10, X11,

: C5

:

X14, X15, X18

.0.2350

7

100

5.1900

%5

.0.0375

.6 :

:

-

-

:

.1

.2

%5

100

.3

.1 " (2006)

.2 " (2008)

1 ."

35 8-1

3. Albayrak, S. and Amasyali, F. (2003), "Fuzzy C-Means Clustering on Medical Diagnostic Systems". International XII. Turkish Symposium on Artificial Intelligence and Neural Networks(TAINN).
4. Dunham, M.H. (2004), "Data Mining Introductory and Advanced Topics". PEARSON Education Inc.
5. Kumar, S. and Rishan, P.R. (2002), "Mining Web Data Using Clustering Technique for Web Personalization". International Journal of Computational Intelligence and Applications, Vol. 2, No. 3, pp. 255-265.
6. Jantzen, J(2004), "A Physical Analogy to Fuzzy Clustering". Fourth European Symposium on Intelligent Technologies and their Implementation on Smart Adaptive Systems (EUNITE), Aachen, Germany, pp. 521-525 .
7. Jermyn, A. , Dixon, P. M. and Read, B.J, (2001), "Preparing Clean Views of Data for Data Mining". European Research Consortium for Information and Mathematics Workshop Proceedings.
8. www.aoad.org \ الكتاب السنوي \ (2006) \ منظمة التنمية الزراعية العربية \ المجلد 26 \ الإحصاءات الزراعية العربية

(1): الانتاج النباتي لمحصول القمح في الدول العربية للفترة 1998-2004.

: / :

2004			2003			متوسط الفترة 2002-98			الدولة
الانتاج	الانتاجية	المساحة	الانتاج	الانتاجية	المساحة	الانتاج	الانتاجية	المساحة	
Prod.	Yield	Area	Prod.	Yield	Area	Prod.	Yield	Area	
13.17	1227	10.73	42.53	1428	29.78	32.73	1250	26.18	الأردن
0.02	2000	0.01	0.02	2000	0.01	0.24	5304	0.05	الإمارات
1722.00	1663	1035.20	1983.40	2141	926.60	1026.52	1132	907.00	تونس
2730.70	1358	2010.50	2964.85	1448	2047.57	1760.02	929	1894.53	الجزائر
2775.68	5307.2275	523.00	2524.00	4882	517.00	2011.80	4477	449.32	السعودية
364.00	2287	159.18	398.00	2311	172.20	255.60	2131	119.95	السودان
4537.46	2478	1831.23	4913.00	2736	1796.00	3885.50	2322	1673.06	سوريا
0.24	706	0.34	0.24	706	0.34	0.25	633	0.40	الصومال A
1832.00	1190	1539.75	2329.00	1343	1734.00	1577.79	1072	1471.61	العراق
1.20	3084.8329	0.39	1.21	3270	0.37	1.54	4042	0.38	عمان
46.36	2117	21.90	44.23	2042	21.66	47.16	2230	21.14	فلسطين
0.02	2316	0.01	0.03	2333	0.01	0.07	2205	0.03	قطر
5.09	17551.724	0.29	5.09	23136	0.22	2.31	11687	0.20	الكويت
136.80	2934.6777	46.62	116.30	2539	45.81	90.33	2158	41.87	لبنان
46.40	1600	29.00 A	40.00 A	1379	29.00	55.29	1300	42.55	ليبيا
7177.86	6559	1094.30	6844.69	6503	1052.59	6376.66	6297	1012.71	مصر
5539.70	1808	3063.70	5146.80	1722	2988.90	2917.14	1041	2801.22	المغرب
105.27	1256	83.80	103.79	1256	82.63	146.66	1565	93.70	اليمن
27033.97	2361.06	11449.95	27457.18	2399.12	11444.69	20187.61	1912.45	10555.90	الجملة

(2)

البيانات العامة
إستخدام الاراضى لعام 2004
المساحة 100 هكتار

مساحة المراعي	مساحة الغابات	المساحة المتروكة(1)	مساحة المحاصيل الموسمية SEASONAL CROPS		مساحة المحاصيل المستديمة PERM.CROPS AREA		الدولة
			المرويه	المطريه	المرويه	المطريه	
			IRRIGATED	RAINFED	IRRIGATED	RAINFED	
7000.00	130.91	153.25	45.65	32.20	55.11	114.00	الأردن
305.00	31.10	-	56.97	-	190.86	-	الإمارات
4.00	-	0.74	0.71	-	2.88	-	البحرين
4885.38	658.32	711.74	185.46	1990.56	170.07	2106.33	تونس
32824.40	4260.00	3382.92	414.16	3695.81	344.16	359.77	الجزائر
200.00	6.00	-	-	0.41	-	-	جيبوتي A
170000.00	2700.00(2)	3185.00	964.00	-	208.00	-	السعودية
117180.00	64359.96	3990.84	1522.08	14028.00	136.50	0.84	السودان
8278.60	593.30	796.20	1273.00	2587.80	166.40	702.20	سوريا
42000.00	9040.00	350.00	150.00	1218.00	-	32.00	الصومال
4000.00	478.50	3149.75	2097.50	1313.75	244.50	24.00	العراق
1000.00	0.16	-	30.60	-	51.05	-	عمان
150.00	26.00	-	17.01	50.00	7.19	107.90	فلسطين
50.00	0.40	17.40	7.10	-	2.60	-	قطر
136.22	-	-	5.77	-	3.04	-	الكويت A
16.00	4.00	-	54.00	72.00	55.00	87.00	لبنان A
13300.00	600.00	850.00	99.00	767.00	159.00	769.00	ليبيا A
-	35.00	-	2704.66	101.52	585.84	85.01	مصر
21000.00	12246.00	1624.30	681.10	5798.60	623.50	207.80	المغرب
39250.00	4400.00	-	22.29	190.00	20.00	50.00	موريتانيا يا
7000.00	1600.00	420.68	396.50	551.30	157.20	83.80	اليمن
468579.60	101169.65	18632.82	10727.56	32396.95	3182.90	4729.65	المجموع

(1)

(2)