Fuzzy C-Means clustering:

Fuzzy C-Means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters.

Example:

K-Mean clustering

Data Sets: (3, 4) (6, 7) (4, 5) (5, 7) (2, 6)

After clustering by K-Mean

Data	class	centroid
(3, 4) (4, 5) (2, 6)	I	(3, 5)
(6, 7) (5, 7)	II	(5.5, 7)

⇓

Data	class-I	class-II
(3, 4)	100%	0%
(4, 5)	100%	0%
(2, 6)	100%	0%
(6, 7)	0%	100%
(5, 7)	0%	100%

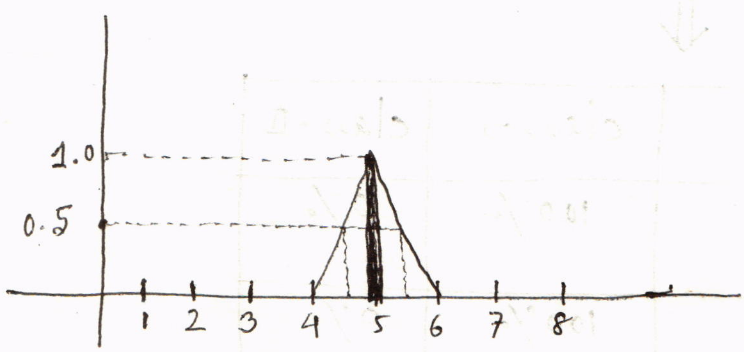
K-Mean clustering states that a data object belongs to a single class either in class-I or class-II in this case.

Fuzzy-C Mean clustering

Data	class-I	class-II
(3,4)	97%	3%
(4,5)	86%	14%
(2,6)	91%	9%
(6,7)	4%	96%
(5,7)	7%	93%

Fuzzy c-Mean clustering states that each data object belongs to two or more classes.

Fuzzy value :



Member	Degree of membership (u_{ij})
4	0
4.5	0.5
5	1.0
5.5	0.5
6	0

Fuzzy c-Mean clustering algorithm

step1: Initialize $U = [u_{ij}]$ matrix, $U^{(0)}$

step2: At k-step: calculate the centers vectors $C^{(k)} = [c_j]$ with $U^{(k)}$

$$c_j = \frac{\sum_{i=1}^{N} \left(u_{ij}^{m} \cdot x_i \right)}{\sum_{i=1}^{N} u_{ij}^{m}}$$

step3: Update $U^{(k)}, U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\| x_i - c_j \|}{\| x_i - c_k \|} \right)^{\frac{2}{m-1}}}$$

step4: If $\| U^{(k+1)} - U^{(k)} \| < \epsilon$ then stop; otherwise return to step 2.

Parameters:

Fuzzyness coefficient, $m = 2$ (m any real number greater than

Termination condition, $\max_{ij} \{ | u_{ij}^{(k+1)} - u_{ij}^{(k)} | \} < \epsilon$

where ϵ is a termination criterion between 0 and 1, whereas k are the iteration steps.

u_{ij} = degree of membership of x_i in the cluster j.

x_i = i th of d-dimensional measured data.

c_j = d-dimension center of the cluster.

$\| * \|$ = Any norm expressing the similarity between any measured data and the center.

Example :

cluster the data (3, 4) (6, 7) and (4, 6) into two clusters using the Fuzzy C-Means clustering algorithm.

Number of cluster 2 (C=2)

Ans:

step1:

$$U^{(0)} = \begin{array}{c|cc} & & \\ i=1 & 0.8 & 0.2 \\ i=2 & 0.3 & 0.7 \\ i=3 & 0.9 & 0.1 \\ & j=1 & j=2 \end{array}$$

Number of data set 3 (N=3)

[Total of each row must not greater than 1]

Here, $u_{11} = 0.8$ $u_{12} = 0.2$

$u_{21} = 0.3$ $u_{22} = 0.7$

$u_{31} = 0.9$ $u_{32} = 0.1$

Iteration-I

step2:

~~Since~~ c_1 and c_2 must be calculated because C=2

c_1 calculation:

(1st dim) $c_1 = \dfrac{u_{11} * 3 + u_{21} * 6 + u_{31} * 4}{u_{11} + u_{21} + u_{31}}$

$= \dfrac{0.8 * 3 + 0.3 * 6 + 0.9 * 4}{0.8 + 0.3 + 0.9} = \dfrac{2.4 + 1.8 + 3.6}{2} = \dfrac{7.8}{2}$

$= 3.9$

(2nd dim) $c_1 = \dfrac{u_{11} * 4 + u_{21} * 7 + u_{31} * 6}{u_{11} + u_{21} + u_{31}}$

$= \dfrac{0.8 * 4 + 0.3 * 7 + 0.9 * 6}{0.8 + 0.3 + 0.9} = \dfrac{10.7}{2} = 5.35$

$\therefore c_1 \equiv (3.9, 5.35)$

c_2 calculation :

(Ist dim) $c_2 = \dfrac{u_{12} * 3 + u_{22} * 6 + u_{32} * 4}{u_{12} + u_{22} + u_{32}}$

$= \dfrac{0.2 * 3 + 0.7 * 6 + 0.1 * 4}{0.2 + 0.7 + 0.1}$

$= \dfrac{5.2}{1} = 5.2$

(2nd dim) $c_2 = \dfrac{u_{12} * 4 + u_{22} * 7 + u_{32} * 6}{u_{12} + u_{22} + u_{32}}$

$= \dfrac{0.2 * 4 + 0.7 * 7 + 0.1 * 6}{0.2 + 0.7 + 0.1}$

$= \dfrac{6.3}{1} = 6.3$

$c_2 \equiv (5.2, 6.3)$

step3:

$i=1, \ j=1 \quad u_{ij} = \dfrac{1}{\sum_{k=1}^{c} \left(\dfrac{\| x_i - c_j \|}{\| x_i - c_k \|} \right)^{\frac{2}{m-1}}}$

$\Rightarrow u_{11} = \dfrac{1}{\sum_{k=1}^{2} \left(\dfrac{\| x_1 - c_1 \|}{\| x_1 - c_k \|} \right)^{2}}$ [if m=2]

$= 1 \div \left\{ \left(\dfrac{\| x_1 - c_1 \|}{\| x_1 - c_1 \|} \right)^2 + \left(\dfrac{\| x_1 - c_1 \|}{\| x_1 - c_2 \|} \right)^2 \right\}$

$= 1 \div \left\{ 1 + \left(\dfrac{\| (3,4) - (3.9, 5.35) \|}{\| (3,4) - (5.2, 6.3) \|} \right)^2 \right\}$

$= 1 \div \left\{ 1 + \dfrac{(3-3.9)^2 + (4-5.35)^2}{(3-5.2)^2 + (4-6.3)^2} \right\}$

$= 1 \div \left\{ 1 + \dfrac{(0.9)^2 + (1.35)^2}{(2.2)^2 + (2.3)^2} \right\} = 1 \div \left\{ 1 + \dfrac{2.6325}{10.13} \right\}$

$= 1 \div 1.26 = 0.794$

$i=1 \quad j=2 \quad u_{ij} = \dfrac{1}{\sum_{k=1}^{c} \left(\dfrac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$

$$\Rightarrow u_{12} = \frac{1}{\sum_{k=1}^{2} \left(\frac{\|x_1 - c_2\|}{\|x_1 - c_k\|} \right)^2}$$

$$= \frac{1}{\left(\frac{\|x_1 - c_2\|}{\|x_1 - c_1\|} \right)^2 + \left(\frac{\|x_1 - c_2\|}{\|x_1 - c_2\|} \right)^2}$$

$$= \frac{1}{\left(\frac{\|x_1 - c_2\|}{\|x_1 - c_1\|} \right)^2 + 1}$$

$$= \frac{1}{\left\{ \frac{\|(3,4) - (5.2, 6.3)\|}{\|(3,4) - (3.9, 5.35)\|} \right\}^2 + 1}$$

$$= \frac{1}{\frac{(3-5.2)^2 + (4-6.3)^2}{(3-3.9)^2 + (4-5.35)^2} + 1}$$

$$= \frac{1}{1 + \frac{2.2^2 + 2.3^2}{0.9^2 + 1.35^2}} = \frac{1}{1 + \frac{10.13}{2.6325}} = 0.206$$

$i=2 \quad j=1 \quad u_{ij} = \dfrac{1}{\sum\limits_{K=1}^{c} \left(\dfrac{\| x_i - c_j \|}{\| x_i - c_K \|} \right)^{\frac{2}{m-1}}}$

$$\Rightarrow u_{21} = \frac{1}{\sum\limits_{K=1}^{2} \left(\frac{\| x_2 - c_1 \|}{\| x_2 - c_K \|} \right)^{2}}$$

$$= \frac{1}{\left(\frac{\| x_2 - c_1 \|}{\| x_2 - c_1 \|} \right)^2 + \left(\frac{\| x_2 - c_1 \|}{\| x_2 - c_2 \|} \right)^2}$$

$$= \frac{1}{1 + \left(\frac{\| (6,7) - (3.9, 5.35) \|}{\| (6,7) - (5.2, 6.3) \|} \right)^2}$$

$$= \frac{1}{1 + \frac{(6-3.9)^2 + (7-5.35)^2}{(6-5.2)^2 + (7-6.3)^2}}$$

$$= \frac{1}{1 + \frac{2.1^2 + 1.65^2}{0.8^2 + 0.7^2}} = \frac{1}{1 + \frac{7.1325}{1.13}} = \frac{1}{7.312}$$

$$= 0.137$$

$i=2 \quad j=2 \quad u_{ij} = \dfrac{1}{\sum\limits_{K=1}^{c} \left(\dfrac{\| x_i - c_j \|}{\| x_i - c_K \|} \right)^{\frac{2}{m-1}}}$

$$\Rightarrow u_{22} = \frac{1}{\left(\frac{\| x_2 - c_2 \|}{\| x_2 - c_1 \|} \right)^2 + 1}$$

$$= \frac{1}{\left(\frac{\| (6,7) - (5.2, 6.3) \|}{\| (6,7) - (3.9, 5.35) \|} \right)^2 + 1} = \frac{1}{1 + \frac{(6-5.2)^2 + (7-6.3)^2}{(6-3.9)^2 + (7-5.35)^2}}$$

$$= 0.863 \qquad = \frac{1}{1 + 1.13/7.1325} = \frac{1}{1 + \frac{.8^2 + .7^2}{2.1^2 + 1.65^2}}$$

$i=3 \quad j=1 \quad u_{ij} = \dfrac{1}{\sum_{k=1}^{c}\left(\dfrac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^2}$

$= \dfrac{1}{\sum_{k=1}^{2}\left(\dfrac{\|x_3 - c_1\|}{\|x_3 - c_k\|}\right)^2}$

$= \dfrac{1}{1+\left(\dfrac{\|x_3 - c_1\|}{\|x_3 - c_2\|}\right)^2} = \dfrac{1}{1+\dfrac{\|(4,6)-(3.9,5.35)\|}{\|(4,6)-(5.2,6.3)\|}}$

$= \dfrac{1}{1+\dfrac{(4-3.9)^2+(6-5.35)^2}{(4-5.2)^2+(6-6.3)^2}}$

$= 0.780$

$i=3 \quad j=2 \quad u_{ij} = \dfrac{1}{\sum_{k=1}^{c}\left(\dfrac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^2}$

$= \dfrac{1}{\sum_{k=1}^{2}\left(\dfrac{\|x_3 - c_2\|}{\|x_3 - c_k\|}\right)^2} = \dfrac{1}{\left(\dfrac{\|x_3 - c_2\|}{\|x_3 - c_1\|}\right)^2+1}$

$= \dfrac{1}{\left(\dfrac{\|(4,6)-(5.2,6.3)\|}{\|(4,6)-(3.9,5.35)\|}\right)^2+1}$

$= \dfrac{1}{\dfrac{(4-5.2)^2+(6-6.3)^2}{(4-3.9)^2+(6-5.35)^2}+1} = 0.22$

$U^{(1)} = \begin{bmatrix} 0.794 & 0.206 \\ 0.137 & 0.863 \\ 0.780 & 0.220 \end{bmatrix}$

step 4:

$|u_{11}^{(1)} - u_{11}^{(0)}| = |0.794 - 0.8| = 0.006$

$|u_{12}^{(1)} - u_{12}^{(0)}| = |0.206 - 0.2| = 0.006$

$|u_{21}^{(1)} - u_{21}^{(0)}| = |0.137 - 0.3| = 0.163$

$|u_{22}^{(1)} - u_{22}^{(0)}| = |0.863 - 0.7| = 0.163$

$|u_{31}^{(1)} - u_{31}^{(0)}| = |0.780 - 0.9| = 0.12$

$|u_{32}^{(1)} - u_{32}^{(0)}| = |0.220 - 0.1| = 0.12$

$\max(0.006, 0.163, 0.12) = 0.163$

$0.163 < \epsilon \Rightarrow 0.163 < 0.3$ [if $\epsilon = 0.3$]

Algorithm stop

Data	class-I	class-II
(3,4)	79.4%	20.6%
(6,7)	13.7%	86.3%
(4,6)	78%	22%