Fuzzy C-Means

MACHINE LEARNING

Pakarat Musikawan

Crisp Logic/Set vs Fuzzy Logic/Set

Crisp Logic/Set

- An element either completely belongs to a set or it does not belong to the set.
- The membership degree is strictly binary—either 0 or 1.

Fuzzy Logic/Set

- An element can belong to a set to a certain degree, rather than being strictly in or out.
- The membership degree ranges between 0 and 1.

Crisp Logic/Set vs Fuzzy Logic/Set

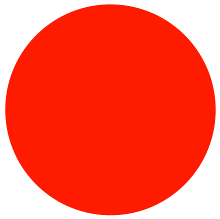
Crisp Logic/Set

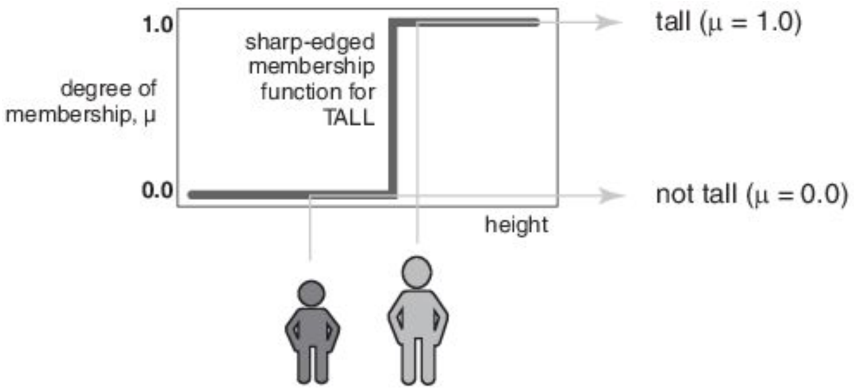
- **Clustering Example**: In crisp clustering (e.g., k-means), each data point is assigned exactly to one cluster, without any ambiguity. A data point either fully belongs to a cluster or does not.
- **Logic Example**: In crisp (Boolean) logic, statements are either true (1) or false (0), with no intermediate states.

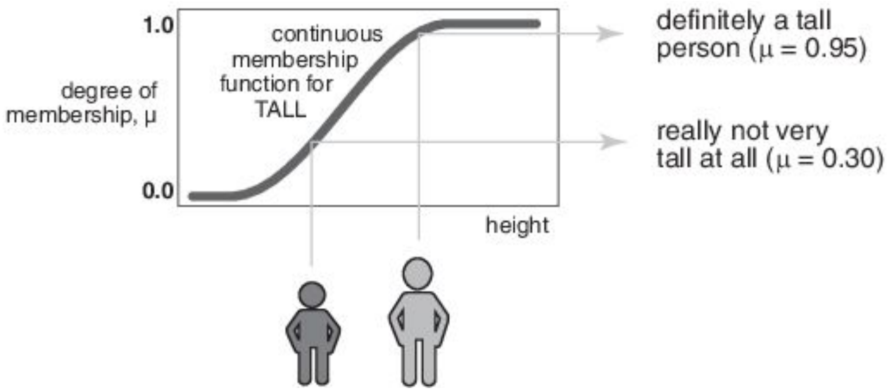
Fuzzy Logic/Set

- **Clustering Example**: In fuzzy clustering (e.g., fuzzy c-means), each data point can belong to multiple clusters with varying degrees of membership. A data point could have 0.6 membership in one cluster and 0.4 membership in another.
- **Logic Example**: In fuzzy logic, truth values are not strictly true or false. Instead, they can take any value between 0 and 1. For example, the truth of a statement like "the weather is hot" might be 0.7 (partially true) if it's warm but not extremely hot.

Crisp Logic/Set vs Fuzzy Logic/Set

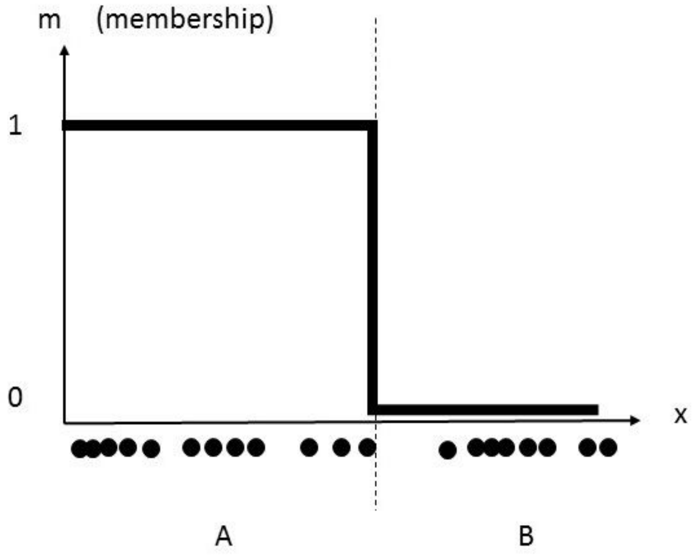


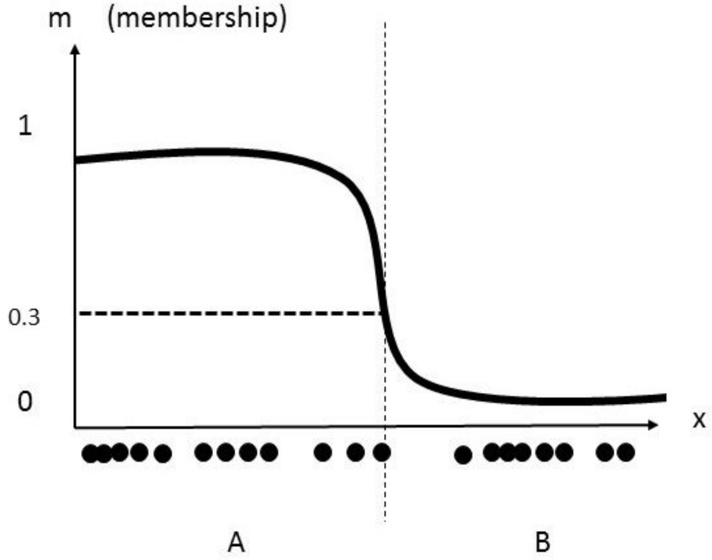




Fuzzy C-Means

Fuzzy C-Means (FCM) is a clustering algorithm that extends the concept of K-means clustering by allowing each data point to belong to multiple clusters to varying degrees.





Fuzzy C-Means

Key Characteristics of Fuzzy C-Means:

- Fuzzy Membership: Each data point has a membership value for each cluster, indicating the degree of belonging.
- Cost Function: A weighted sum of squared distances between data points and cluster centroids.

$$\mathcal{J} = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{i,j}^{m} \left\| \boldsymbol{x}_i - \boldsymbol{v}_j \right\|^2 ; \quad \text{s.t.} \sum_{j=1}^{C} u_{i,j} = 1$$

- Fuzziness Parameter: The parameter *m* controls the degree of fuzziness in the clustering.

Steps of Fuzzy C-Means

1. Initialization
2. Calculate Cluster Centroids
3. Update Membership Values
4. Check for Convergence
5. Repeat Steps 2-4

Steps of Fuzzy C-Means – Initialization

Choose the number of clusters *C* and the fuzziness parameter *m* (typically *m* > 1).

Initialize the membership matrix U randomly. The membership values for each data point should sum to 1 for each cluster.

$$\mathbf{U} = \begin{bmatrix} u_{1,1} & \cdots & u_{1,C} \\ \vdots & \ddots & \vdots \\ u_{N,1} & \cdots & u_{N1,C} \end{bmatrix} ; \sum_{j=1}^{C} u_{i,j} = 1$$

Steps of Fuzzy C-Means – Centroid Calculation

Update the centroid of each cluster based on the current membership values.

$$\boldsymbol{v}_j = \frac{\sum_{i=1}^{N} u_{i,j}^{m} \boldsymbol{x}_i}{\sum_{i=1}^{N} u_{i,j}^{m}}$$

Steps of Fuzzy C-Means – Membership Update

Update the membership values based on the distances from each data point to each centroid.

$$u_{i,j} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|\boldsymbol{x}_i - \boldsymbol{v}_j\|}{\|\boldsymbol{x}_i - \boldsymbol{v}_k\|} \right)^{\frac{2}{m-1}}}$$

Steps of Fuzzy C-Means – Stopping Criteria

Change in Membership Values

$$\left\|\mathbf{U}^{t+1} - \mathbf{U}^{t}\right\| \leq \epsilon$$

Change in Centroids

$$\left\|\mathbf{V}^{t+1} - \mathbf{V}^{t}\right\| \leq \epsilon$$

Maximum Number of Iterations

$$t > T_{\max}$$

Example

Let's consider a dataset with 4 data points in a 2D space

$$\mathbf{X} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \\ \boldsymbol{x}_4 \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 \\ 0.2 & 0.4 \\ 0.3 & 0.1 \\ 0.8 & 0.5 \end{bmatrix}$$

Example

Let's randomly initialize the membership matrix, where *C*=2

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{32} \\ u_{41} & u_{42} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \\ 0.7 & 0.3 \\ 0.2 & 0.8 \end{bmatrix}$$

Example

Calculate the centroids for each cluster, where *m*=2

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{32} \\ u_{41} & u_{42} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \\ 0.7 & 0.3 \\ 0.2 & 0.8 \end{bmatrix}$$

$$\boldsymbol{v}_1 = \frac{(0.5^2 \cdot (0.1, 0.2)) + (0.4^2 \cdot (0.2, 0.4)) + (0.7^2 \cdot (0.3, 0.1)) + (0.2^2 \cdot (0.8, 0.5))}{(0.5^2 + 0.4^2 + 0.7^2 + 0.2^2)}$$

$$\approx (0.251, 0.194)$$

Example

Calculate the centroids for each cluster, where *m*=2

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{32} \\ u_{41} & u_{42} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \\ 0.7 & 0.3 \\ 0.2 & 0.8 \end{bmatrix}$$

$$\boldsymbol{v}_2 = \frac{(0.5^2 \cdot (0.1, 0.2)) + (0.6^2 \cdot (0.2, 0.4)) + (0.3^2 \cdot (0.3, 0.1)) + (0.8^2 \cdot (0.8, 0.5))}{(0.5^2 + 0.6^2 + 0.3^2 + 0.8^2)}$$

$$\approx (0.474, 0.390)$$

Example

Calculate the centroids for each cluster, where *m*=2

$$\boldsymbol{v}_1 = \frac{(0.5^2 \cdot (0.1, 0.2)) + (0.4^2 \cdot (0.2, 0.4)) + (0.7^2 \cdot (0.3, 0.1)) + (0.2^2 \cdot (0.8, 0.5))}{(0.5^2 + 0.4^2 + 0.7^2 + 0.2^2)}$$

$$\approx (0.251, 0.194)$$

$$\boldsymbol{v}_2 = \frac{(0.5^2 \cdot (0.1, 0.2)) + (0.6^2 \cdot (0.2, 0.4)) + (0.3^2 \cdot (0.3, 0.1)) + (0.8^2 \cdot (0.8, 0.5))}{(0.5^2 + 0.6^2 + 0.3^2 + 0.8^2)}$$

$$\approx (0.474, 0.390)$$

Example

Calculate the Euclidean distance for each data point to the centroids

$$\|\boldsymbol{x}_1 - \boldsymbol{v}_1\| = \|(0.1, 0.2) - (0.251, 0.194)\| \approx 0.151$$
$$\|\boldsymbol{x}_1 - \boldsymbol{v}_2\| = \|(0.1, 0.2) - (0.474, 0.390)\| \approx 0.420$$

$$\|\boldsymbol{x}_2 - \boldsymbol{v}_1\| = \|(0.2, 0.4) - (0.251, 0.194)\| \approx 0.211$$
$$\|\boldsymbol{x}_2 - \boldsymbol{v}_2\| = \|(0.2, 0.4) - (0.474, 0.390)\| \approx 0.274$$

$$\|\boldsymbol{x}_3 - \boldsymbol{v}_1\| = \|(0.3, 0.1) - (0.251, 0.194)\| \approx 0.106$$
$$\|\boldsymbol{x}_3 - \boldsymbol{v}_2\| = \|(0.3, 0.1) - (0.474, 0.390)\| \approx 0.338$$

$$\|\boldsymbol{x}_4 - \boldsymbol{v}_1\| = \|(0.8, 0.5) - (0.251, 0.194)\| \approx 0.628$$
$$\|\boldsymbol{x}_4 - \boldsymbol{v}_2\| = \|(0.8, 0.5) - (0.474, 0.390)\| \approx 0.343$$

Example

Calculate the new membership values using the distances

$$u_{11} = \frac{1}{\left(\frac{0.151}{0.151}\right)^{\frac{2}{1}} + \left(\frac{0.151}{0.420}\right)^{\frac{2}{1}}} \approx 0.885, \quad u_{12} = \frac{1}{\left(\frac{0.420}{0.151}\right)^{\frac{2}{1}} + \left(\frac{0.420}{0.420}\right)^{\frac{2}{1}}} \approx 0.114$$

$$u_{21} = \frac{1}{\left(\frac{0.211}{0.211}\right)^{\frac{2}{1}} + \left(\frac{0.211}{0.274}\right)^{\frac{2}{1}}} \approx 0.627, \quad u_{22} = \frac{1}{\left(\frac{0.274}{0.211}\right)^{\frac{2}{1}} + \left(\frac{0.274}{0.274}\right)^{\frac{2}{1}}} \approx 0.372$$

$$u_{31} = \frac{1}{\left(\frac{0.106}{0.106}\right)^{\frac{2}{1}} + \left(\frac{0.106}{0.338}\right)^{\frac{2}{1}}} \approx 0.910, \quad u_{32} = \frac{1}{\left(\frac{0.338}{0.106}\right)^{\frac{2}{1}} + \left(\frac{0.338}{0.338}\right)^{\frac{2}{1}}} \approx 0.089$$

$$u_{41} = \frac{1}{\left(\frac{0.628}{0.628}\right)^{\frac{2}{1}} + \left(\frac{0.628}{0.343}\right)^{\frac{2}{1}}} \approx 0.229, \quad u_{42} = \frac{1}{\left(\frac{0.343}{0.628}\right)^{\frac{2}{1}} + \left(\frac{0.343}{0.343}\right)^{\frac{2}{1}}} \approx 0.770$$

Workshop

$$\mathbf{X} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \\ \boldsymbol{x}_4 \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 \\ 0.2 & 0.4 \\ 0.3 & 0.1 \\ 0.8 & 0.5 \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \\ u_{41} & u_{42} & u_{43} \end{bmatrix} = \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.5 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.4 & 0.2 \end{bmatrix}$$

$$m = 2$$