

NPTEL ONLINE CERTIFICATION COURSES

Course Name: Fuzzy logic and Neural Networks

Faculty Name: Prof. D. K. Pratihar

Department : Mechanical Engineering

Week 3

NPTEL



NPTEL ONLINE CERTIFICATION COURSES

Course Name: Fuzzy logic and Neural Networks

Faculty Name: Prof. D. K. Pratihar

Department : Mechanical Engineering

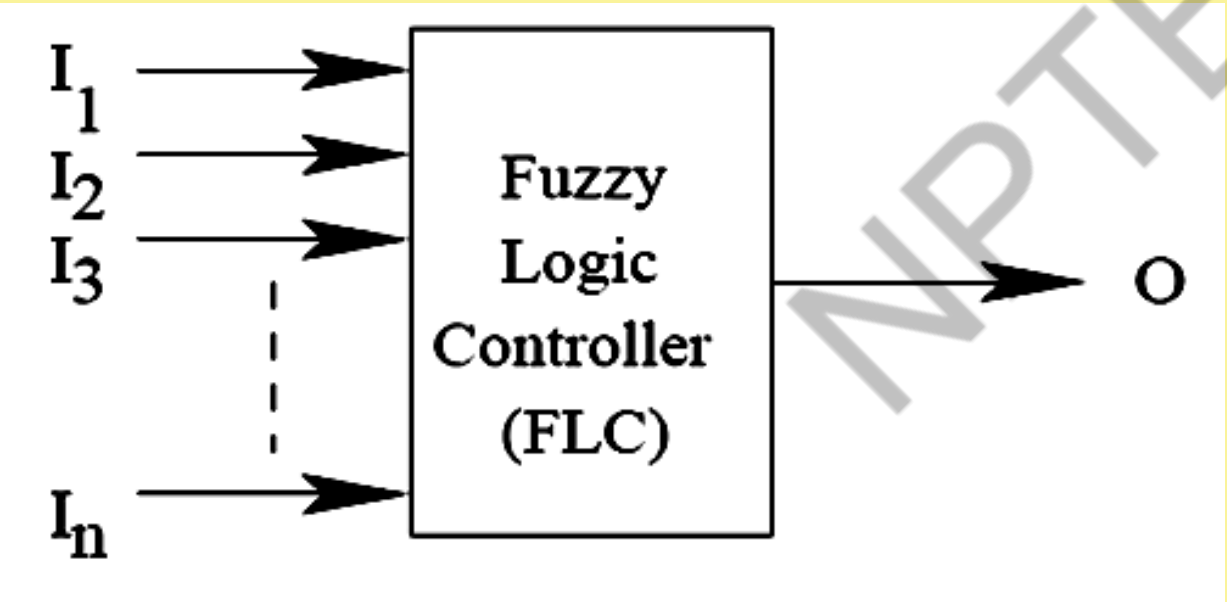
Topic

Lecture 03: Applications of Fuzzy Sets (Contd..)

NPTEL

Hierarchical Fuzzy Logic Controller

Let us assume that a process having n inputs ($I_1, I_2, \ldots\ldots I_n$) and one output (that is, O) is to be modeled using an FLC



NPTEL

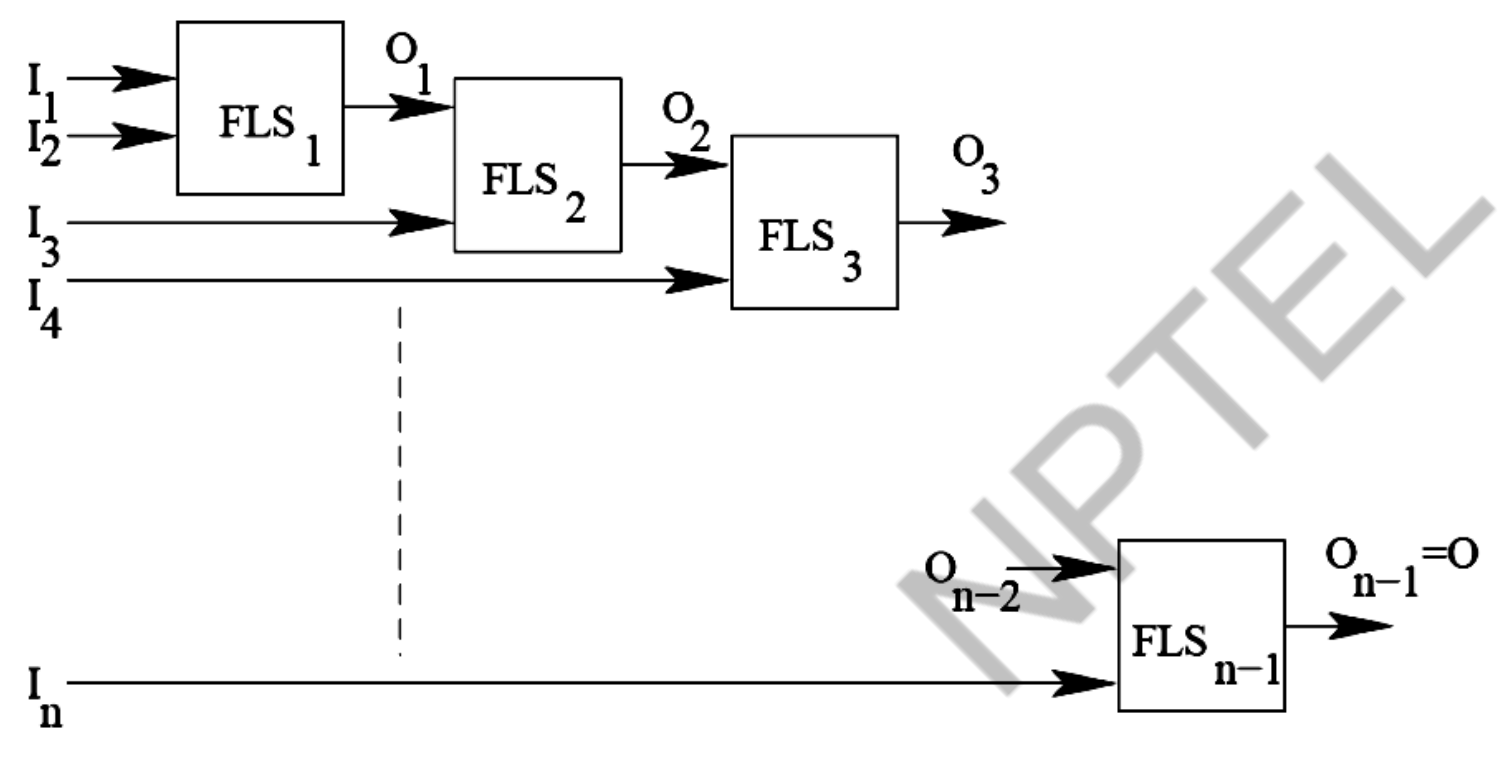
Let us also assume that m linguistic terms (fuzzy sets) are used to represent each variable.

Total no. of rules = m^n

Problem : curse of dimensionality

NPTEL

To overcome this problem, we go for Hierarchical Fuzzy Logic Controller



NPTEL

Total no. of rules = $m^2(n-1)$

Sensitivity Analysis

- To determine the importance of input variable towards the output

$$\text{Sensitivity } S = \delta O / \delta I$$

Methods to Determining Sensitivity

1. We carry out experiment by varying the input variables through different amounts, say 0.1%, 1.0%, 10.0% etc. and recording the outputs of the controller

NPTEL

2.$O = f(I_1, I_2)$

We change the input I_1 by a small amount δI_1 and determine the change in output using Mean-Value theorem

$$\delta O = f(I_1 + \delta I_1, I_2) - f(I_1, I_2) = \frac{\partial f}{\partial I_1}_{\overline{I_1}, I_2} \delta I_1$$

NPTEL

FREE ONLINE EDUCATION swayam

Where $\overline{I_1}$ lies between I_1 and $I_1+\delta I_1$

If $\left|\frac{\partial f}{\partial I_1}\right| > 1$, then $\delta O > \delta I_1$

Thus, the contribution of I_1 on the output O can be determined.

NPTEL

Advantages of FLC

- Potential tool for dealing with imprecision and uncertainty

- Does not require an extensive mathematical formulation of the problem

NPTEL

Disadvantages of FLC

- Designing a proper KB of an FLC is a difficult task

- May not be suitable for modeling a process involving many variables

NPTEL

FREE ONLINE EDUCATION swayam

Fuzzy Clustering

- Clustering is a powerful tool of data mining

- In clustering, grouping of data into several clusters is done based on their similarity among them

- Clusters could be either crisp or fuzzy in nature

- Several methods of fuzzy clustering: Fuzzy C-Means Algorithm, Potential-based Clustering, Entropy-based Clustering, and others

NPTEL

Fuzzy C-Means Clustering

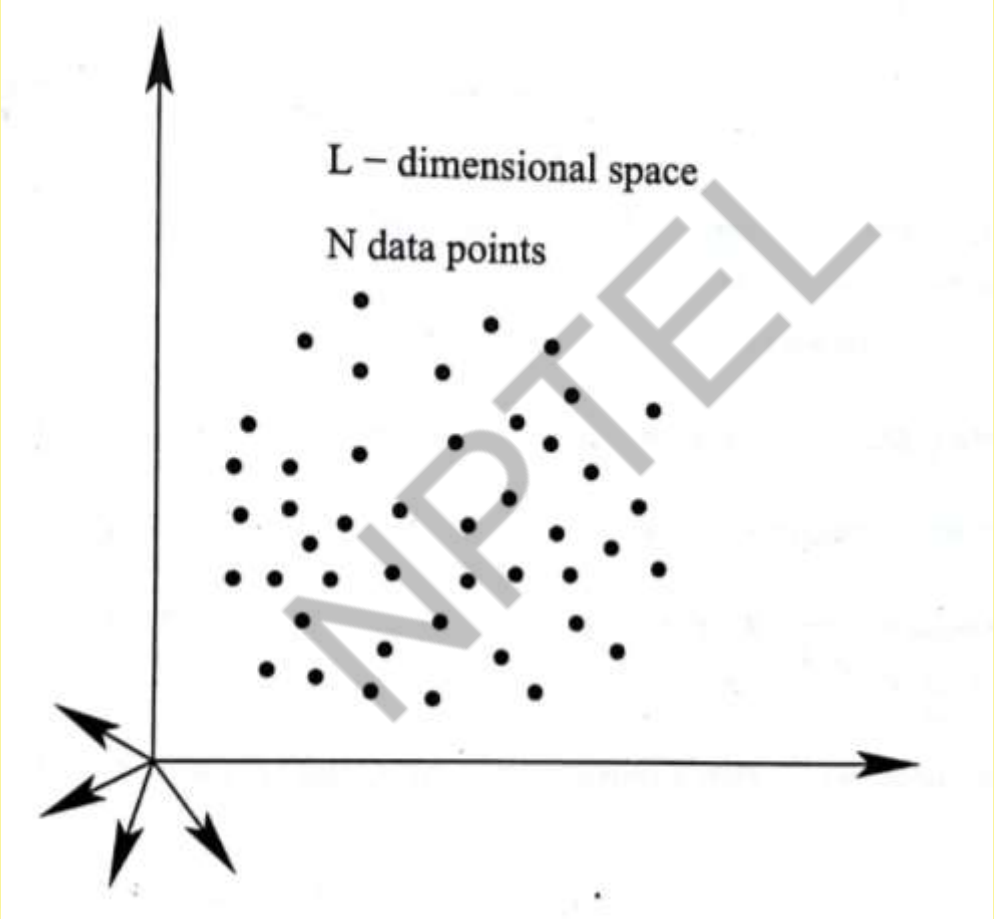
- Proposed by Bezdek in 1973

- A particular data of the set may be the member of several clusters with different values of membership

- An attempt is made to minimize dissimilarity measure (expressed in terms of Euclidean distance) of the data points with the pre-defined clusters

NPTEL

Fuzzy C-Means Clustering



NPTEL

Fuzzy C-Means Clustering

$$MinimizeDissimilarity$$

$$F(\mu, C) = \sum_{j=1}^{C} \sum_{i=1}^{N} \mu_{ij}^{g} d_{ij}^{2}$$

$$s.t. \sum_{j=1}^{C} \mu_{ij} = 1.0$$

$$LagrangeMultiplier$$

$$\bar{F}(\mu, C, \lambda_{1,} ....., \lambda_{N}) = \sum_{j=1}^{C} \sum_{i=1}^{N} \mu_{ij}^{g} d_{ij}^{2} + \sum_{i=1}^{N} \lambda_{i} (\sum_{j=1}^{C} \mu_{ij} - 1.0)$$

$$\frac{d\bar{F}}{d\mu} = 0; \frac{d\bar{F}}{dC} = 0; \frac{d\bar{F}}{d\lambda_{1}} = 0; .....; \frac{d\bar{F}}{d\lambda_{N}} = 0$$

NPTEL

Fuzzy C-Means Clustering

$$CC_j = \frac{\sum_{i=1}^{N} \mu_{ij}^{g} x_i}{\sum_{i=1}^{N} \mu_{ij}^{g}}$$

$$\mu_{ij} = \frac{1}{\sum_{m=1}^{C} \left(\frac{d_{ij}}{d_{im}} \right)^{\frac{2}{g-1}}}$$

NPTEL

•An iterative algorithm, in which the cluster centers and membership values of the data points with the clusters are updated to minimize the dissimilarity measure

NPTEL

Steps of FCM Algorithm

- **Step 1:** Assume the number of clusters to be made, that is, C

  $2 \leq C \leq N$

- **Step 2:** Choose an appropriate level of cluster fuzziness $g > 1$

- **Step 3:** μ_{ij}: Membership value of *i*-th data point with *j*- th cluster. Initialize the N×C sized membership matrix [μ] at random, such that $\mu_{ij} \in [0.0, 1.0]$ and $\sum_{j=1}^{C} \mu_{ij} = 1.0$, for each *i*

NPTEL

- **Step 4:** Calculate k-th dimension of *j*-th cluster center CC_{jk} as follows:

$$CC_{jk} = \frac{\sum_{i=1}^{N} \mu_{ij}^{g} x_{ik}}{\sum_{i=1}^{N} \mu_{ij}^{g}}$$

- **Step 5:** Calculate the Euclidean distance between *i*-th data point and *j*-th cluster center, that is, d_{ij}

NPTEL

- **Step 6:** Update fuzzy membership matrix [μ] according to d_{ij}

$$\mu_{ij} = \frac{1}{\sum_{m=1}^{C} \left(\frac{d_{ij}}{d_{im}} \right)^{\frac{2}{g-1}}}, \quad \text{if } d_{ij} > 0$$

$$= 1.0, \quad \text{if } d_{ij} = 0$$

- **Step 7:** Repeat from step 4 to step 6 until the changes in [μ]

come out to be less than some pre-specified values

NPTEL

FREE ONLINE EDUCATION swayam

Numerical Example

Ten points lying on a free-form surface are shown below.

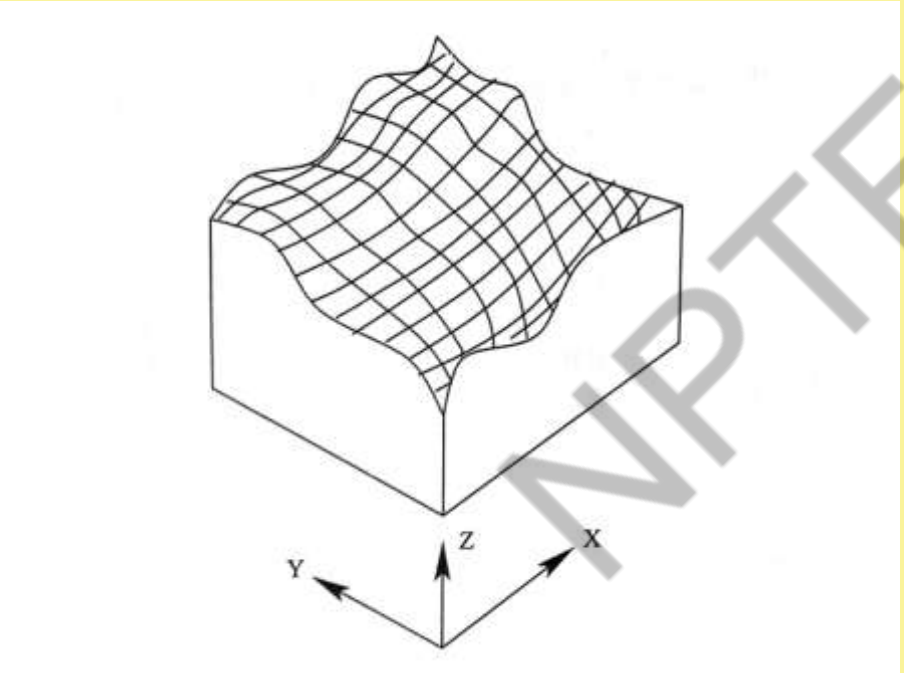


Figure: Free-form surface

NPTEL

Points	X-coordinate	Y-coordinate	Z-coordinate
1	0.2	0.4	0.6
2	0.4	0.3	0.8
3	0.8	0.2	0.5
4	0.9	0.5	0.4
5	0.6	0.6	0.6
6	0.3	0.4	0.5
7	0.7	0.6	0.5
8	0.2	0.5	0.3
9	0.3	0.6	0.8
10	0.8	0.3	0.1

NPTEL

Carry out clustering using the fuzzy C-means algorithm. Assume g = 1.25, termination criterion ε = 0.01.

Solution:

Number of data points to be considered N = 10

Dimensions of the data L = 3

Level of cluster fuzziness g = 1.25

Let us assume that the number of clusters to be made C = 2 .

NPTEL

FREE ONLINE EDUCATION swayam

Let us also suppose that a membership matrix [μ] of size 10×2 is generated at random, as given below.

$$[\mu] = \begin{bmatrix} 0.680551 & 0.319449 \\ 0.495150 & 0.504850 \\ 0.821897 & 0.178103 \\ 0.303795 & 0.696205 \\ 0.333966 & 0.666034 \\ 0.431538 & 0.568462 \\ 0.415384 & 0.584616 \\ 0.509643 & 0.490357 \\ 0.469850 & 0.530150 \\ 0.189164 & 0.810836 \end{bmatrix}$$

NPTEL

The first dimension of first cluster center CC_{11} is determined as follows:

$$CC_{11} = \frac{\sum_{i=1}^{N} \mu_{i1}^{g} x_{i1}}{\sum_{i=1}^{N} \mu_{i1}^{g}} = \frac{A}{B},$$

where

$$\begin{aligned} A = & \left(0.680551^{1.25} \times 0.2\right) + \left(0.495150^{1.25} \times 0.4\right) + \left(0.821897^{1.25} \times 0.8\right) \\ & + \left(0.303795^{1.25} \times 0.9\right) + \left(0.333966^{1.25} \times 0.6\right) + \left(0.431538^{1.25} \times 0.3\right) \\ & + \left(0.415384^{1.25} \times 0.7\right) + \left(0.509643^{1.25} \times 0.2\right) + \left(0.469850^{1.25} \times 0.3\right) \\ & + \left(0.189164^{1.25} \times 0.8\right) \\ = & 1.912120 \end{aligned}$$

NPTEL

FREE ONLINE EDUCATION swayam

$$B = \left(0.680551^{1.25}\right) + \left(0.495150^{1.25}\right) + \left(0.821897^{1.25}\right) + \left(0.303795^{1.25}\right) + \left(0.333966^{1.25}\right) + \left(0.431538^{1.25}\right) + \left(0.415384^{1.25}\right) + \left(0.509643^{1.25}\right) + \left(0.469850^{1.25}\right) + \left(0.189164^{1.25}\right) = 3.923066$$

Therefore, $CC_{11} = \frac{A}{B} = \frac{1.912120}{3.923066} = 0.487404$

Similarly, we detemine $CC_{12} = 0.412837$; $CC_{13} = 0.543316$

$CC_{21} = 0.554860$; $CC_{22} = 0.459073$; $CC_{23} = 0.477262$

NPTEL

Coordinates of two cluster centers are found to be as follows:

$C_1 : (0.487404, 0.412837, 0.543316)$

$C_2 : (0.554860, 0.459073, 0.477262)$

We determine the modified values of membership, as follows:

$$\mu_{11} = \frac{1}{1 + \left(\frac{d_{11}}{d_{12}}\right)^{\frac{2}{g-1}}},$$

NPTEL

FREE ONLINE EDUCATION swayam शिक्षित भारत, उन्नत भारत

where $d_{11} = \sqrt{(0.487404-0.2)^2 + (0.412837-0.4)^2 + (0.543316-0.6)^2}$

$= 0.293221526$

$d_{12} = \sqrt{(0.554860-0.2)^2 + (0.459073-0.4)^2 + (0.477262-0.6)^2}$

$= 0.380105056$

$\therefore \mu_{11} = \dfrac{1}{1+\left(\dfrac{d_{11}}{d_{12}}\right)^8} = 0.888564,$ as $g = 1.25$

NPTEL

Other elements of $[\mu]$ can be determined as follows:

$$[\mu] = \begin{bmatrix} 0.888564 & 0.111436 \\ 0.909207 & 0.090793 \\ 0.376533 & 0.623467 \\ 0.142926 & 0.857074 \\ 0.217307 & 0.782693 \\ 0.922229 & 0.077771 \\ 0.060962 & 0.939038 \\ 0.562085 & 0.437915 \\ 0.788303 & 0.211697 \\ 0.232446 & 0.767554 \end{bmatrix}$$

It completes one iteration of the algorithm.

NPTEL

After few iterations, we get Cluster-centers

$C_1 : (0.296189, 0.447977, 0.600323)$

$C_2 : (0.773939, 0.425782, 0.404034)$

NPTEL

Values of the membership grades are determined as follows:

$$[\mu] = \begin{bmatrix} 0.999999 & 0.000001 \\ 0.997102 & 0.002898 \\ 0.001223 & 0.998777 \\ 0.000008 & 0.999992 \\ 0.351332 & 0.648668 \\ 0.999992 & 0.000008 \\ 0.002764 & 0.997236 \\ 0.992437 & 0.007563 \\ 0.999452 & 0.000548 \\ 0.001836 & 0.998164 \end{bmatrix}$$

NPTEL

First cluster: 1-st , 2-nd, 6-th, 8-th and 9-th points

Second cluster: 3-rd, 4-th, 5-th, 7-th and 10-th points

NPTEL

Entropy-Based Fuzzy Clustering

- Entropy is an index determined based on a similarity measure, which depends on Euclidean distance

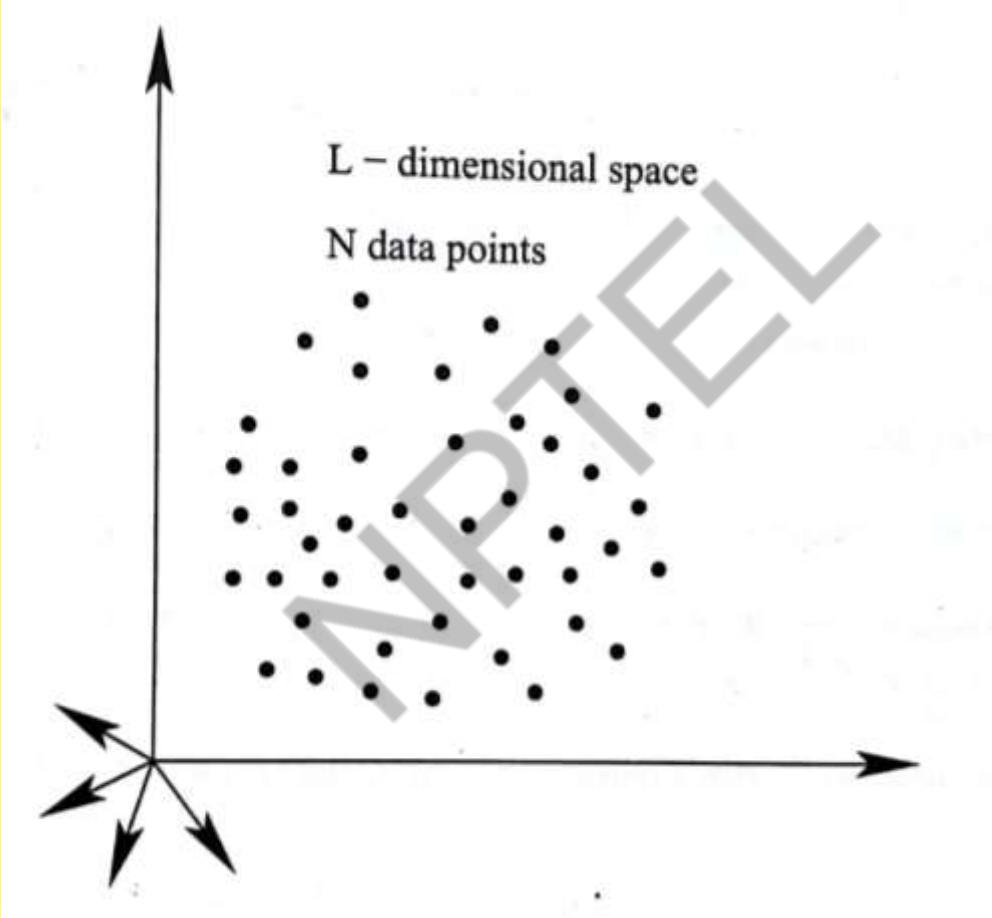
- The data point having the minimum entropy value is selected as the cluster center

- The data points having similarity with this cluster center more than a pre-specified value will form a cluster

NPTEL

FREE ONLINE EDUCATION swayam

Entropy-based Fuzzy Clustering



NPTEL

•Let us consider N data points in L-D hyperspace

•<u>Step 1:</u> Arrange the data set in N rows and L columns

•<u>Step 2:</u> Calculate Euclidean distance between the points i and j as follows:

$$d_{ij} = \sqrt{\sum_{k=1}^{L} \left(x_{ik} - x_{jk} \right)^2}$$

NPTEL

•<u>Step 3:</u> Determine the similarity S_{ij} between two data points i and j

$$S_{ij} = e^{-\alpha d_{ij}},$$

where α is a constant to be determined

NPTEL

Let us assume a similarity of 0.5, when the distance between two data points d_{ij} becomes equal to mean distance of all pairs of data points.

$$d_{ij} = \bar{d} = \frac{1}{{}^{N}C_{2}} \sum_{i=1}^{N} \sum_{j>i}^{N} d_{ij}$$

We get $\alpha = \ln 2 / \bar{d}$

Note: S_{ij} varies in the range of 0.0 to 1.0

NPTEL

- **Step 4:** Calculate the entropy E of all N data points

$$E = -S \log_2 S - (1 - S) \log_2 (1 - S)$$

- **Step 5:** Determine total entropy value at a data point x_i with respect to all other data points

$$E_i = -\sum_{j \in x}^{j \neq i} \left(S_{ij} \log_2 S_{ij} + (1 - S_{ij}) \log_2 (1 - S_{ij}) \right)$$

NPTEL

Clustering Algorithm

- **Step 1:** Calculate entropy E_i for each data point x_i lying in [T]

- **Step 2:** Identify x_i that has the minimum E_i value and select it $(x_{i,min})$ as the cluster center

NPTEL

- **Step 3:** Put $x_{i,min}$ and the data points having similarity with $x_{i,min}$ greater than β (threshold value of similarity) in a cluster and remove them from [T]

- **Step 4:** Check if [T] is empty. If yes, terminate the program, else go to step 2.

NPTEL

Concept of Outliers

After the clustering is done, we count the number of data points lying in each cluster and if this number becomes greater than or equal to γ% of total number of data points, we declare this cluster to be a valid one. Otherwise, these data points will be declared as the outliers.

NPTEL

FREE ONLINE EDUCATION swayam

Numerical Example

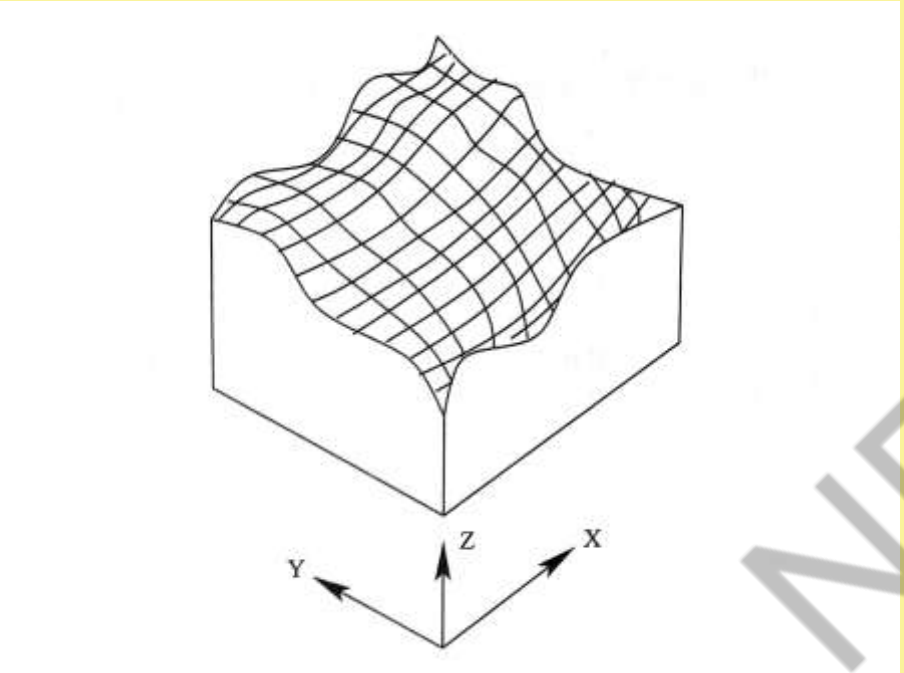


Figure: Free-form surface

NPTEL

- **Ten points lying on a free-form surface are shown in the table.**

Points	X-coordinate	Y-coordinate	Z-coordinate
1	0.2	0.4	0.6
2	0.4	0.3	0.8
3	0.8	0.2	0.5
4	0.9	0.5	0.4
5	0.6	0.6	0.6
6	0.3	0.4	0.5
7	0.7	0.6	0.5
8	0.2	0.5	0.3
9	0.3	0.6	0.8
10	0.8	0.3	0.1

NPTEL

Carry out fuzzy clustering based on their similarity and entropy values. Assume β=0.5, γ=10%.

Solution:

Calculate Euclidean distance between two points: *i* and *j*

$$d_{ij} = \sqrt{\sum_{k=1}^{L} (x_{ik} - x_{jk})^2}$$

Here, number of data points N=10

NPTEL

There will be $10_{C_2} = 45$ distance values (d_{ij})

$$d_{ji}=d_{ij}$$

and

$$d_{00} = d_{11} = d_{22} = d_{33} = \text{------} = d_{99} = 0.0$$

NPTEL

Mean of distance values

$$\bar{d} = \frac{\sum d_{ij}}{45} = 0.518373$$

$$\alpha = \frac{ln\ 2}{\bar{d}} = 1.337160$$

Similarity between the points *i* and *j*

$$S_{ij} = e^{-\alpha d_{ij}}$$

NPTEL

Euclidean distance and similarity values are calculated, as shown in the table.

Combination of data points	Euclidean distance	Similarity
0,1	0.3000000	0.669551
0,2	0.640312	0.424773
0,3	0.734847	0.374334
0,4	0.447214	0.549912
0,5	0.141421	0.827701
0,6	0.547723	0.480757
0,7	0.316228	0.655179
0,8	0.300000	0.669551
0,9	0.787401	0.348931

NPTEL

1,2	0.509902	0.505695
1,3	0.670820	0.407793
1,4	0.412311	0.576186
1,5	0.331662	0.641795
1,6	0.519615	0.499170
1,7	0.574456	0.463875
1,8	0.316228	0.655179
1,9	0.806226	0.340257

NPTEL

FREE ONLINE EDUCATION swayam शिक्षित भारत, उन्नत भारत

2,3	0.331662	0.641795
2,4	0.458258	0.541851
2,5	0.538516	0.486712
2,6	0.412311	0.576186
2,7	0.700000	0.392189
2,8	0.707107	0.388479
2,9	0.412311	0.576186
3,4	0.374166	0.606337
3,5	0.616441	0.438550
3,6	0.244949	0.720697
3,7	0.707107	0.388479
3,8	0.728011	0.377771
3,9	0.374166	0.606337

NPTEL

4,5	0.374166	0.606337
4,6	0.141421	0.827701
4,7	0.509902	0.505695
4,8	0.360555	0.617473
4,9	0.616441	0.438550
5,6	0.447214	0.549912
5,7	0.244949	0.720697
5,8	0.360555	0.617473
5,9	0.648074	0.420387
6,7	0.547723	0.480757
6,8	0.500000	0.512436
6,9	0.509902	0.505695
7,8	0.519615	0.499170
7,9	0.663325	0.411901
8,9	0.911043	0.295759

NPTEL

Entropy of *i*-th data point

$$E_i = -\sum_{j \in x}^{j \neq i} (S_{ij} log_2 S_{ij} + (1 - S_{ij}) log_2 (1 - S_{ij}))$$

Entropy values are found to be as follows:

$$E_0 = 8.285456; E_1 = 8.665640$$
$$E_2 = 8.815512; E_3 = 8.568537$$

NPTEL

$$E_4 = 8.516900; E_5 = 8.348550$$
$$E_6 = 8.490743; E_7 = 8.686420$$
$$E_8 = 8.560217; E_9 = 8.632634$$

0-th point has the least entropy value, and it is to be the first cluster center.

Let the threshold value β=0.5

1-st, 4-th, 5-th, 7-th and 8-th are found to make cluster with the 0-th point.

NPTEL

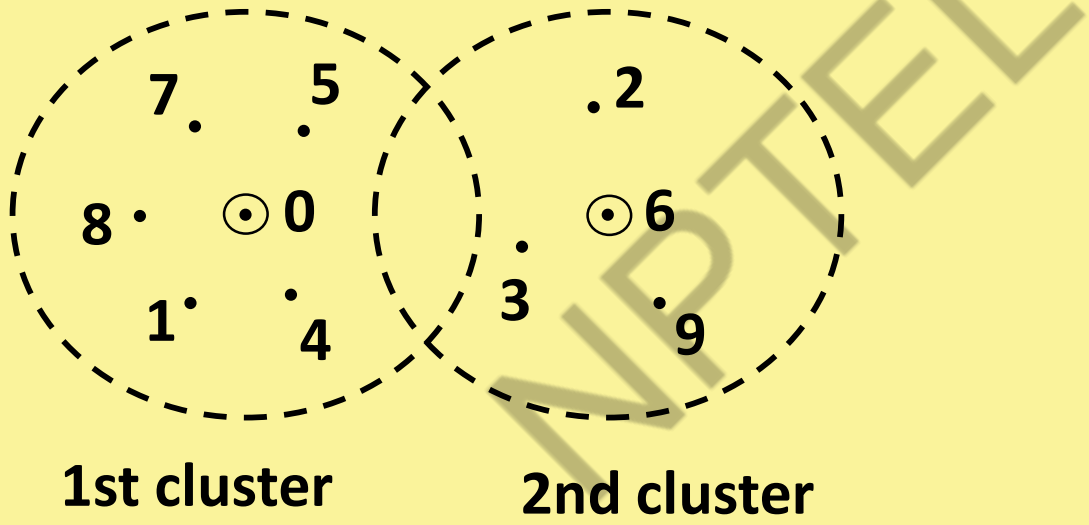
Remaining points are : 2-nd, 3-rd, 6-th, 9-th

Out of these four points, the 6-th point is seen to have the least values of entropy.

6-th point: center of second cluster

NPTEL

As β=0.5, 2-nd, 3-rd and 9-th points are seen to make cluster with the 6-th point.



NPTEL

Entropy-Based Fuzzy Clustering

- Number and nature/quality of clusters depend on threshold value of similarity

- More flexible compared to FCM algorithm

- Yields more distinct but less compact clusters

- We want more distinct and more compact clusters, for which Entropy-based Fuzzy C-Means Clustering has been proposed
- Performances of clustering algorithms are data dependent

NPTEL

FREE ONLINE EDUCATION swayam

References

References:

- Soft Computing: Fundamentals and Applications by D.K. Pratihar, Narosa Publishing House, New-Delhi, 2014
- Fuzzy Sets and Fuzzy Logic: Theory and Applications by G.J. Klir, B. Yuan, Prentice Hall, 1995
- V. Dey, D.K. Pratihar, G.L. Datta, “Genetic algorithm-tuned entropy-based fuzzy c-means algorithm for obtaining distinct and compact clusters,” Fuzzy Optimization and Decision Making, 10, pp. 153-166, 2011.

NPTEL

Conclusion

Conclusion:

- Working principles of Fuzzy Reasoning Tool (Fuzzy Logic Controller) and Fuzzy Clustering have been discussed with suitable numerical examples

NPTEL



FREE ONLINE EDUCATION
swayam
शिक्षित भारत, उन्नत भारत

NPTEL ONLINE CERTIFICATION COURSES

Thank you!

NPTEL