M.Sc. Mathematics

Bharat & Sandeep

2022

Image Segmentation

A DISSERATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE

OF
MASTER OF SCIENCE

IN

MATHEMATICS

Submitted by:

Bharat Varshney & Sandeep

2K20/MSCMAT/07 & 2K20/MSCMAT/27

Under the supervision of

Dr. Dhirendra Kumar

DEPARTMENT OF APPLIED MATHEMATICS

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

MAY, 2022

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

DECLARATION

We, Bharat Varshney & Sandeep, 2K20/MSCMAT/07 & 2K20/MSCMAT/27 student of M.Sc. Mathematics, hereby declare that the project Dissertation titled "Image Segmentation" which is submitted by us to the Department of Applied Mathematics Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Science, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi

BHARAT VARSHNEY

SANDEEP

Date:

DEPARTMENT OF APPLIED MATHEMATICS

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled " Image Segmentation " which is submitted by Bharat Varshney & Sandeep , Roll No. 2K20/MSCMAT/07 & 2K20/MSCMAT/27 [Department of Applied Mathematics], Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Science, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Date:

DR. Dhirendra Kumar
SUPERVISOR

DEPARTMENT OF APPLIED MATHEMATICS
DELHI TECHNOLOGICAL UNIVERSITY
BAWANA ROAD, DELHI-110042

Abstract

The Objective of this thesis is to talk about the usage of Fuzzy Logic in pattern recognition. There are different fuzzy approaches to recognize the pattern and the structure in data. The fuzzy approach that we choose to process the data is completely depends on the type of data.

Pattern reorganization as we know involves various mathematical transforms so as to render the pattern or structure with the desired properties such as the identification of a probabilistic model which provides the explaination of the process generating the data clarity seen and so on and so forth. With this basic school of thought we plunge into the world of Fuzzy Logic for the process of pattern recognition.

Fuzzy Logic like any other mathematical field has its own set of principles, types, representations, usage so on and so forth. Hence our job primarily would focus to venture the ways in which Fuzzy Logic is applied to pattern recognition and knowledge of the results. That is what will be said in topics to follow.

Pattern recognition is the collection of all approaches that understand, represent and process the data as segments and features by using fuzzy sets. The representation and processing depend on the selected fuzzy technique and on the problem to be solved.

In the broadest sense, pattern recognition is any form of information processing for which both the input and output are different kind of data, medical records, aerial photos, market trends, library catalogs, galactic positions, fingerprints, psychological profiles, cash flows, chemical constituents, demographic features, stock options, military decisions.. Most pattern recognition techniques involve treating the data as a variable and applying standard processing techniques to it.

ACKNOWLEDGEMENT

The satisfaction of successful completion of any task would be incomplete without mentioning the people who made it possible and whose constant guidance and encouragement crown all of my efforts with success. I express my deep sense of gratitude to my parents, my brother and my guide Dr. Dhirender Kumar, Assistant Professor, Department of Applied Mathematics, Delhi Technological University for her inspiration, constant assistance, valuable suggestions, sympathetic advice, fruitful conversation and unparalleled encouragement made throughout the course of study, without which this piece of work would not have taken its present shape.

Bharat Varshney

Sandeep

Table of Contents

Chapter 1

Technology & Trends

Fuzzy logic is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth - truth values between "completely true" and "completely false". It was introduced by Dr. Lofti Zadeh of U.C. Berkeley in the 1960's.

1.1 History

A pictorial object is a fuzzy set which is specified by some membership function defined on all picture points. From this point of view, each element participates in many memberships. Some of this uncertainty is due to degradation, but some of it is inherent…In fuzzy set terminology, making figure/ground distinctions is equivalent to transforming from membership functions to characteristic functions.“ 1970, J.M.B. Prewitt

Fuzzy logic is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth - truth values between "completely true" and "completely false". Fuzzy Logic is aimed at precision of approximate reasoning.

The use of fuzzy logic for creating decision-support and expert systems has grown in popularity among management and financial decision-modeling experts. Still others are putting it to work in pattern recognition, economics, data analysis, and other areas that involve a high level of uncertainty, complexity, or nonlinearity.

1.2 The Data

There are different types of data like qualitative, quantitative, numerical, pictorial, textural, and linguistic or in some cases can be indifferent combinations of these. Examples of the data sources are medical records, aerial photos, market trends, library catalogs, galactic positions, fingerprints, psychological profiles, cash flows, chemical constituents, demographic features, stock options, military decisions. The technique or method of search pattern is applicable to any of these data types and sources. The search option is used to know the techniques for data processing and usually the data set is denoted by X.

1.3 Cluster Analysis

Clustering in X has classically meant the identification of an integer $C, 2 \leq c \leq m$ and a partitioning of X by c mutually exclusive, collectively exhaustive subsets of X (the "clusters"). The member of each and every cluster have more similarity to one another than members of other clusters. It will be precisely applicable for the mathematical similarities between the $x_k's$ in some operational sense. Cluster structure in X tells associations among individuals of a population.

1.4 Classifications

X has been drawn from the data space which is denoted by S, i.e., $X \subset S$.
A classifer for S is a device where S itself is going to be partitioned into c decision regions. Clear representation of these regions rely on: • The nature of S

• The way in which the regions are formed

• The model we choose such as on the data, the search, and the structure.

The above factors have an effect on the role played by a sample data set X from data space S in classifier design. X is commonly used to delineate the decision regions in S. It is possible to search for structure in an entire data space during the process of classification. The structure may enable us to classify subsequent observations rapidly and automatically. The main purpose of classification is construction of taxonomy of the physical processes involved. Classification attempts to discover associations between subclasses of a population.

Chapter 2

Classical Clustering Analysis

The clustering analysis concerns to the partition of the data into the equallant subsets also known as clusters. The data in each equallant subset has the common behavior such as distance, density based algorithm. Density based algorithm is a major equallant subset which explains the number of unknown cluster in advance. It progressively increases the total strength or connectivity of the cluster set (equallant subset) by cumulative attraction of the nodes between the clusters. Major clust is the most pronounce and successful algorithm of unsupervised document clustering. Graph theory depends on the algorithm assigned to each document to that cluster the majority of its neighbors belongs to.

The node neighborhood is going to be calculated by using a particular similarity measure which has been assumed to be the weight of each edge between the nodes of the graph.

Major cluster tells the number of clusters and assigns each target document to precisely one cluster. A different approach should be required in order to determine how the document has to be assigned to more one than one category. The traditional major clust algorithm deals with the crisp data, the another version called F-major clust deals with fuzzy dat.

2.1 K-MEANS CLUSTERING

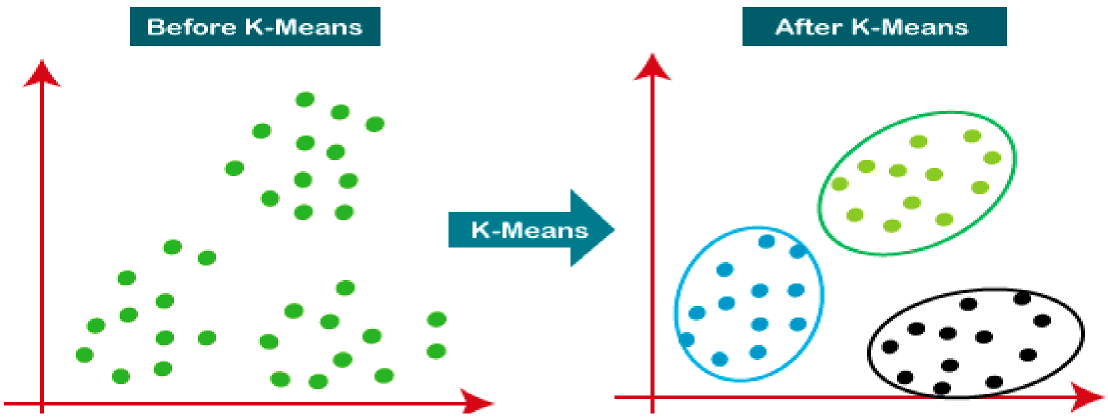
K-Means Clustering is an Unsupervised Learning algorithm which groups the unlabeled dataset into different clusters. Here K defines the number of pre-defined clusters that need to be created in the process, as if K=2, there will be two clusters, and for K=3, there will be three clusters, and so on.

It allows us to cluster the data into different groups and a convenient way to discover the categories of groups in the unlabeled dataset on its own without the need for any training.

It is a centroid-based algorithm, where each cluster is associated with a centroid. The main aim of this algorithm is to minimize the sum of distances between the data point and their corresponding clusters

- The algorithm takes the unlabeled dataset as input, divides the dataset into k-number clusters, and repeats the process until it does not find the best clusters. The value of k should be predetermined in this algorithm. The k-means clustering algorithm mainly performs two tasks: Determines the best value for K center points or centroids by an iterative process.
- Assigns each data point to its closest k-center. Those data points which are near to the particular k-center, create a cluster.

Hence each cluster has datapoints with some commonalities, and it is away from other clusters. The below diagram explains the working of the K-means Clustering Algorithm:



The way **k means** algorithm works is as follows:

1) Set K – To choose a number of desired clusters, K.

2) Initialization – To choose k starting points which are used as initial estimates of the cluster centroids. They are taken as the initial starting values.

3) Classification – To examine each point in the dataset and assign it to the cluster whose centroid is nearest to it.

4) Centroid calculation – When each point in the data set is assigned to a cluster, it is needed to recalculate the new k centroids.

5) Convergence criteria – The steps of (iii) and (iv) require to be repeated until no point changes its cluster assignment or until the centroids no longer move. Compute the sum of the squared distance between data points and all centroids.

Assign each data point to the closest cluster (centroid).

Compute the centroids for the clusters by taking the average of the all data points that belong to each cluster.

The approach kmeans follows to solve the problem is called **Expectation-Maximization**. The E-step is assigning the data points to the closest cluster.

The M-step is computing the centroid of each cluster. Below is a break down of how we can solve it mathematically

The objective function is:

$$J = \sum_{i=1}^{m} \sum_{k=1}^{K} w_{ik} \|x^i - \mu_k\|^2$$

where wik=1 for data point xi if it belongs to cluster *k*; otherwise, wik=0. Also, μk is the centroid of xi's cluster. It's a minimization problem of two parts. We first minimize J w.r.t. wik and treat μk fixed. Then we minimize J w.r.t. μk and treat wik fixed. Technically speaking, we differentiate J w.r.t. wik first and update cluster assignments (*E-step*). Then we differentiate J w.r.t. μk and recompute the centroids after the cluster assignments from previous step (*M-step*). Therefore, E-step:

$$\frac{\partial J}{\partial w_{ik}} = \sum_{i=1}^{m} \sum_{k=1}^{K} \|x^i - \mu_k\|^2$$

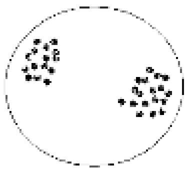
$$\Rightarrow w_{ik} = \begin{cases} 1 & \text{if } k = argmin_j \|x^i - \mu_j\|^2 \\ 0 & \text{otherwise.} \end{cases}$$

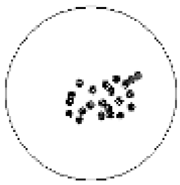
2.2 Drawbacks

K means algorithm is good in capturing structure of the data if clusters have a spherical-like shape. It always try to construct a nice spherical shape around the centroid. That means, the minute the clusters have a complicated geometric shapes, k means does a poor job in clustering the data. We'll illustrate three cases where k means will not perform well.

First, k means algorithm doesn't let data points that are far-away from each other share the same cluster even though they obviously belong to the same cluster.

Clustering models are generally interpreted geometrically by taking into consideration of their response on two-or more three dimensional examples. Some difficulties are intrinsic in trying to elaborate the successful clustering criterion for a wide spectrum of data structures. The ideal cases compact, well-separated, equally proportioned clusters are encountered in the real data of the applications.
Data sets are mixture of shapes as spherical, elliptical and sizes are intensities, unequal number of observations, and geometries are linear, angular and curved.





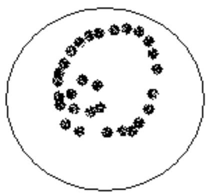
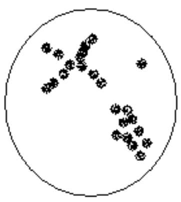
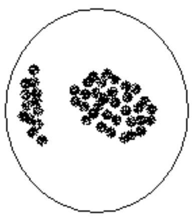


Figure2.0 Figure2.1 Figure2.2





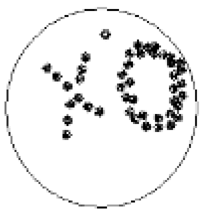


Figure2.3 Figure2.4

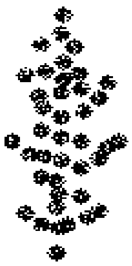
Figure2.5

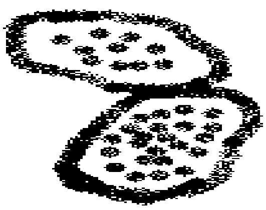
Typical clusters have been obtained by the processing of the two or more data sets with the methods as follows:

- Distance based objective function algorithm and
- Distance based graphic-theoretic method.

This kind of typical behavior could be occured, even though both algorithms use the same distance, it means the equal measure of the similarity between the points in X, but with the different clustering phenomenon in the process. We emphasize the similarity measures by building blocks for clustering criteria, and in the simplest models the measure of similarity can be served as a criterion of validity, but more generally the same measure can be used with various criteria to yield different models and results for the same data set.

Figure 2.6 Figure 2.7 Figure 2.8





2.3 Graph-Theoretic Methods

X should be regarded as the node set in this group. Edge weights between pairs of the nodes could be depended on the measure of similarity between pairs of nodes. The criterion for clustering is commonly some measure of connectivity or bonding between the groups of nodes-breaking edges in minimal spanning tree to form subgraphs. This is often used the clustering strategy. These sought of techniques are well acceptable to data with chains or pseudolinear structure.

Example: The single linkage technique is well suited to data of figures in the above examples.

Some difficulty may cause to the pure graph theoretic methods due to mixed data structures. Data with hyper elliptical clusters, noise, and bridges are usually badly distorted by graph-theoretic models because of their chaining tendencies

2.4 Example for Fuzzy Set

- As an example, we can regard the linguistic variable color which helps us to explain the conception of a fuzzy set. It can be described by a list of terms yellow, orange, red, violet and blue , where each term is a name of the corresponding fuzzy set.

  Color = {yellow, orange, red, violet, blue}

- The non-crisp boundaries between the colors can be represented much better. A soft computing becomes possible (see Figure below).

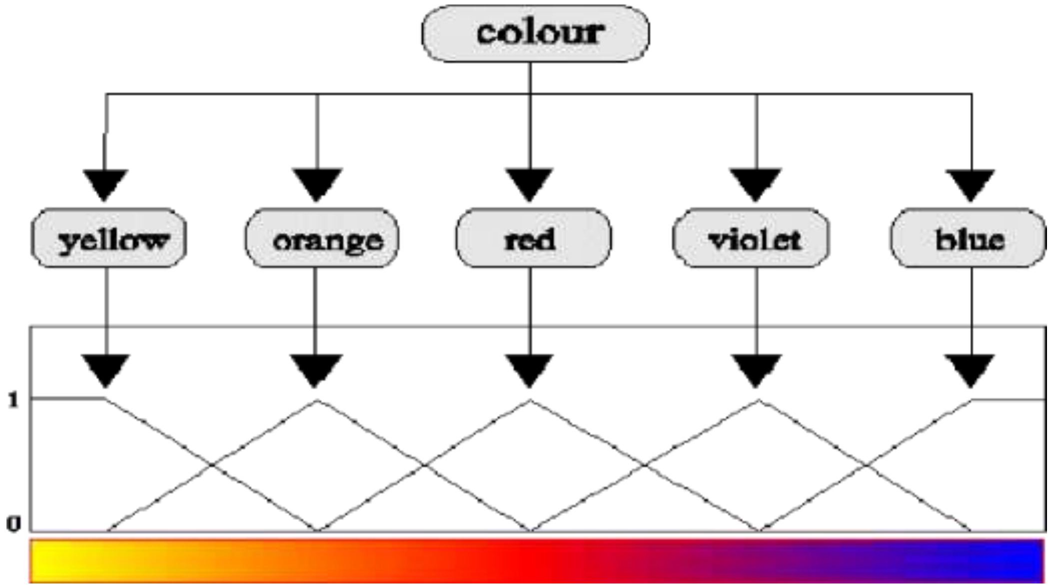
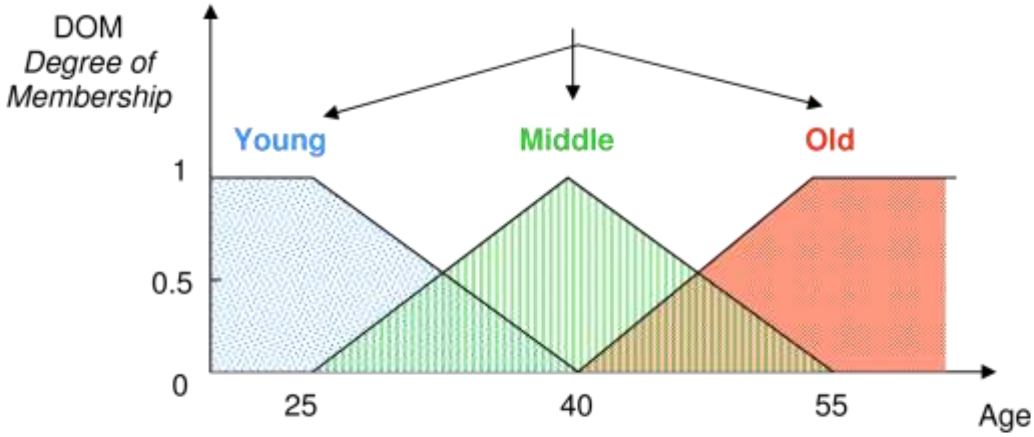


Figure 2-3: Example of Fuzzy sets from the list “color”

Fuzzy values



Elements of "age" have associated degrees of membership in the different set.

Figure 2-4: Membership functions of fuzzy sets forming the linguistic variable “age”

Membership functions representation

Ann is 28, 0.8 in set “Young” Bob is 35, 0.1 in set “Young”
Charlie is 23, 1.0 in set “Young” Don is 54, 0.0 in set “Young”

2.5 Fuzzy and Probability concepts:

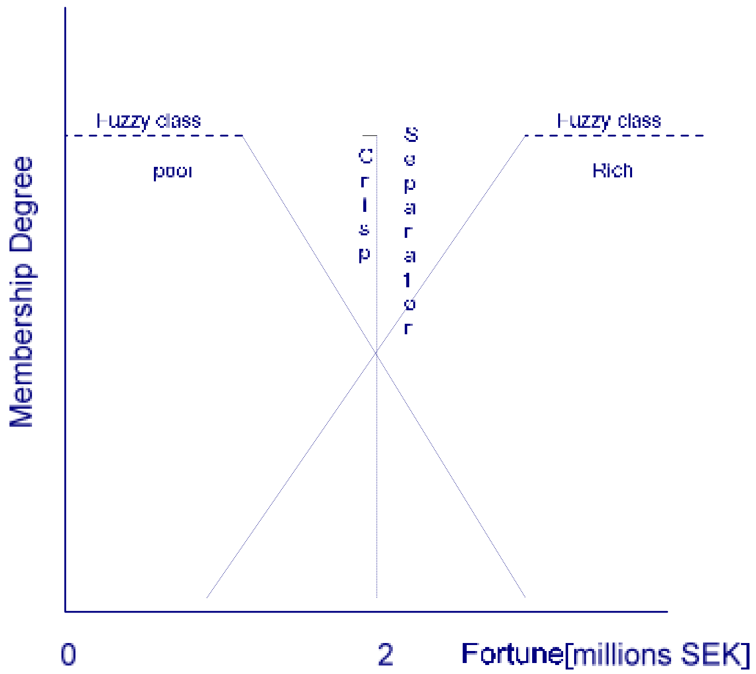
The relationship of fuzzy sets and probability has been intensively discussed since zadeh introduced fuzzy sets in 1965. Recently it has become accepted that they can be considered as independent and complementary to each other. Fuzzy sets are indicators for similarities or neighbourhood relations while probability is related to probabilistic uncertainity. Fuzzy set is defined in terms of the membership function which is a mapping from the universal set U to the interval [0,1]. Membership function could be defined as the graphical representation of the magnitude of participation of each input. It associates a weighting with each of the inputs that are processed, defines functional overlap between inputs, and ultimately determines an output response. Fuzziness is also considered as one form of the uncertaintiy and it is related to the linguistic variables. Linguistic variables are described as membership functions. To avoid this confusing this with fuzzy uncertainty, we explicitly refer to probabilistic uncertainity when we are in the field of probability theory.

2.6 Fuzzy Concept:

Let's consider an example, a bank wants to classify their customers into two groups as rich and poor customers. Obviously there is no crisp separation between rich and poor in a way that customers own less than 2 million SEK are poor while they own a fortune of 2million SEK or more are rich. So according to the above explanation a person with a fortune of 2.1 million SEK is considered as reasonably rich but still a little bit poor. The indicator for similarity in fuzzy sets is called membership degree μ={0,.....,1}. A membership degreeμ =1 indicates that an object completely belongs to set and a membership degree μ=0 shows a total dissimilarity between an object and a set. In our example, the customer with 2.1million SEK may have membership

degrees of $\mu_{rich}(2.1millionSEK) = 0.65$ to the set rich and $\mu_{poor}(2.1millionSEK) = 0.35$ to the set poor. This indicates that the customer is rich but not extremely wealthy. However a customer having 30billion SEK would surely have memberships of $\mu_{rich}(1billionSEK) = 1.0$and $\mu_{poor}(1billionSEK) = 0.0$. There is probabilistic uncertainity neither about the fortune of the customer who is having 2.1million SEK nor about the rules for how to classify him into one or other of the two sets which are determined by the functions given in figure2.2.

So the membership degrees do not indicate any probability of belonging to the sets, but similarities of values to those sets.



2-5: Fuzzy Concept

2.7 Probability Concept:

The same bank may face the probabilistic uncertainity about the wealth of the customer. For example, a new customer driving up with an old bicycle might be considered of having a fortune of say, 5 SEK while it might be assumed that a customer comes to the bank in a big limousine could have a millions of swedish kronars. However these are only the guesses of the bank employees. The vehicles of the customers are indicators for the wealth but no proof. So the bank clerks have to act under probabilistic uncertainity. The biker could be a billionaire while other customer who comes to the bank in a limousine might be a debt-ridden comman.

The biker (BI) might be having the fortune of 5 SEK with a probability of P_{BI}(5 SEK)=0.9 and a fortune of one million SEK with a probability of P_{BI}(1MILLION SEK)=0.1 while the limousine customer (LI) has the probabilities as follows: P_{LI}(5sek)=0.2 and P_{LI}(1million sek)=0.8.

In this example only probabilistic uncertainty is taken into consideration. In contrast to the fuzzy concept as shown in the previous section, the amounts of money (5SEK and 1 million SEK) are not examined with respect to their similarity to the sets poor and rich.

2.8 Combined Fuzzy and Probability Concept:

The fuzzy and probability concepts are independent and they can be combined. Lets consider the example of the bike rider. First the bank clerks estimate the fortune of the new customes: the bike might be having the fortune of 5 SEK with a probability of P_{BI}(5 SEK)=0.9 and a fortune of one million SEK with a probability of P_{BI}(1MILLION SEK)=0.1. Now the given amounts of money are examined with respect to their similarity to the sets rich and poor. 5 SEK may be classified with the following membership degrees: $\mu_{POOR}(5SEK) = 0.95$and $\mu_{RICH}(5SEK) = 0.05$. For 1 million SEK we may get the membership degree as follows:

$\mu_{POOR}(1MILLIONSEK) = 0.02$ and $\mu_{RICH}(1millionsek) = 0.98$.

Combining the probability and fuzziness we finally get: The biker does belong with a probability of P_{BI}=0.9 and to a membership degree of $\mu_{POOR} = 0.95$to the set poor as well as to the set rich with $\mu_{RICH} = 0.05$.With a probability of $P_{BI} = 0.1$ and with a membership degree of $\mu_{RICH} = 0.98$ he belongs to the set rich and as well as to the set poor with $\mu_{POOR} = 0.02$.

Features of the function:

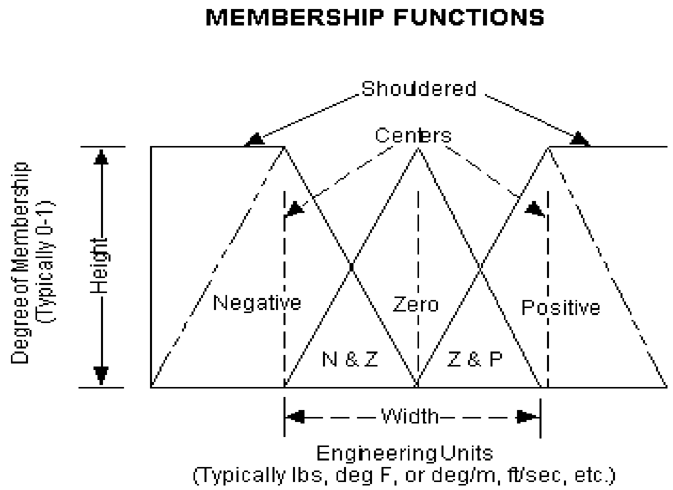


Figure 2-6: Triangular membership functions

Chapter 3:

Fuzzy Logic and Fuzzy Clustering

3.1 Fuzzy logic:

Fuzzy logic is the convenient way to map an input space to an output space.
Example: How good your service was at your resturant, and we will decide what the tip should be.

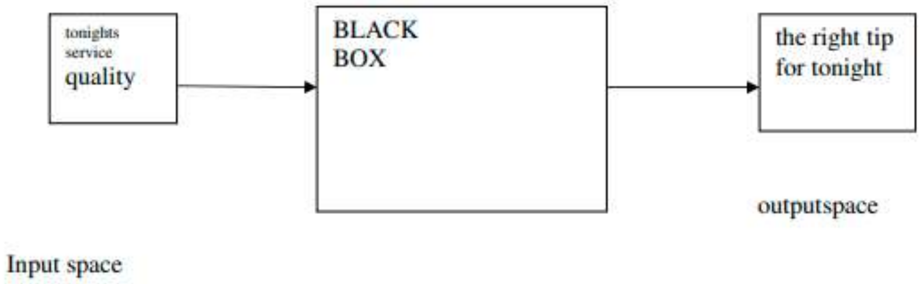


Figure 3-1: An input-output map for the tipping problem

It is all the matter of mapping inputs to the appropriate outputs. We have a black box between input and the output and the black box contains any number of things such as fuzzy systems, linear systems, expert systems neural networks, differential equations, interpolated multi-dimensional lookup tables and etc... There are many ways to make the black box work, it turns out the fuzzy is often the best way.

3.2 K-Means Clustering Numerical Example and Algorithm:

What is K-means clustering?

It is an algorithm to classify or to group the objects depending on attributes/features into K number of group. K is a positive integer number. The grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid. Thus, the purpose of K-mean clustering is to classify the data.

Example:

Let us consider we have 4 objects as the training data point and each object has 2 attributes. Each attribute represents coordinate of the object.

Object	Attribute 1(X) :Weig	Attribute 2(Y): pH

Medicine A	1	1
Medicine B	2	1
Medicine C	4	3
Medicine D	5	4

These objects belong to two groups of medicine(cluster 1 and cluster 2). Now we have to determine which medicines belong to cluster 1 and which medicines belong to the other cluster. Each medicine represents one point with two coordinates.The attribute 1(X) and atrribute 2(Y) values are assigned to the objects which give the cluster identification (which medicine belongs to cluster 1 and which medicine belongs to cluster 2).

Numerical Example:

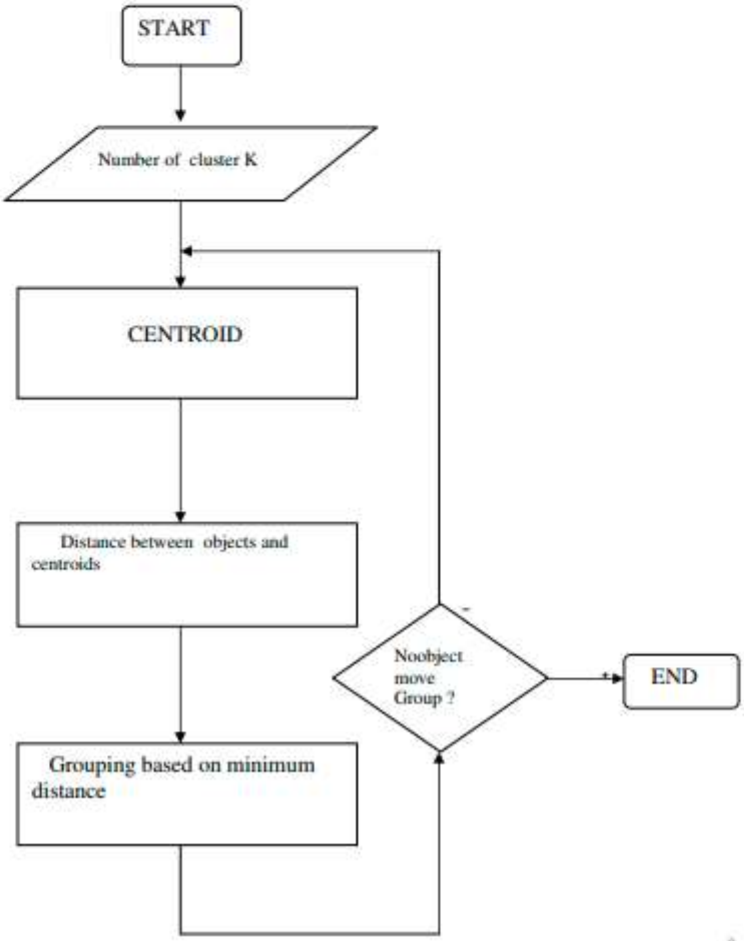
In the begining of the K-means clustering , we determine a number of clusters K and we assume the existence of the centroids or centers of these clusters. We can take any random objects as the initial centroids or the first K objects can also serve as the initial centroids.

Then the K-means algorithm will do the three steps:

Iterate until stable (No object move group):

1.Determine the centroid coordinates
2. Determine the distance of each object to the centroids
3.Group the objects based on minimum distance (we have to find the closest centroid)

Figure 3-9: K-Means Algorithm Flow Chart



The following numerical example is given to understand this simple iteration Suppose we have several objects (4 types of medicines) and each object have two attributes or features as shown in the table below. Our aim is to group these objects into K=2 group of medicine depending on the two features (pH and weight index).

Object	Feature 1 (X): weig	Feature 2(Y): pH
Medicine A	1	1
Medicine B	2	1
Medicine C	4	3
Medicine D	5	4

Each medicine represents one point with two features as (X ,Y).

3.3 Initial value of the centroids:

In this we use medicine A and B as the first centroids. Let c_1 and c_2 are the coordinate of the centroid, then c_1= (1,1) and c_2=(2,1).

Objects-centroids distance:

We calculate the distance between cluster centroid and each object. We use Euclidean distance, then the distance matrix at iteration 0 is:

$$D^o = \begin{bmatrix} 0 & 1 & 3.61 & 5 \\ 1 & 0 & 2.83 & 4.24 \end{bmatrix} \begin{matrix} c_1 = (1,1) group1 \\ c_2 = (2,1) group2 \end{matrix}$$

$$\begin{matrix} & A & B & C & D \\ X & 1 & 2 & 4 & 5 \\ Y & 1 & 1 & 3 & 4 \end{matrix}$$

Each column in the distance matrix represents the object. The first row of the distance matrix corresponds to the distance between each object and the first centroid. The second row is the distance between each object and the second centroid. For example the distance between the third object C(4,3) and the first centroid is 3.61 and its distance to the second centroid is 2.83.

Object Clustering:

We assign each object to cluster 1 or cluster 2 respectively when basing on the minimum distance. So, medicine A is assigned to group 1, medicine B is assigned to group 2, medicine C to group 2 and medicine D to group 2. The element of the group matrix below is 1 if and only if the object is assigned to that group

$$G^o = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{matrix} group & 1 \\ group & 2 \end{matrix}$$

$$\begin{matrix} A & B & C & D \end{matrix}$$

Iteration 1, determine centroids:

Group 1 has only one member so the centroid remains in c_1= (1,1). Group 2 now has three members, so the centroid is the average coordinate among the three members C_2= (11/3, 8/3).

Iteration 1, Objects-Centroids distances: The next step is to find out the distance of all objects to the new centroids.

$$D^1 = \begin{bmatrix} 0 & 1 & 3.61 & 5 \\ 3.14 & 2.36 & 0.47 & 1.87 \end{bmatrix} \begin{matrix} c_1 = (1,1) group1 \\ c_2 = (11/3, 8/3) group2 \end{matrix}$$

$$\begin{matrix} & A & B & C & D \\ X & 1 & 2 & 4 & 5 \\ Y & 1 & 1 & 3 & 4 \end{matrix}$$

Iteration 1, Objects clustering:

Similar to step 3, we assign each object based on the minimum distance and the group matrix is shown below.

$$G^1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{matrix} group1 \\ group2 \end{matrix}$$

$$\begin{matrix} A & B & C & D \end{matrix}$$

Iteration 2, determine centroids:

Group 1 and Group 2 both have two members, so the new centroids are c_1=(3/2 , 1) and c_2=(9/2 , 7/2).

Iteration 2, Objects-Centroids distances:

We have new distance matrix at iteration 2 as

$$D^2 = \begin{bmatrix} 0.5 & 0.5 & 3.20 & 4.61 \\ 4.30 & 3.54 & 0.71 & 0.71 \end{bmatrix} \begin{matrix} c_1 = (3/2, 1) group1 \\ c_2 = (9/, 7/2) group2 \end{matrix}$$

$$\begin{matrix} & A & B & C & D \\ X & 1 & 2 & 4 & 5 \\ Y & 1 & 1 & 3 & 4 \end{matrix}$$

Iteration 2, Objects-Clustering: We assign each object based on minimum distance

$$G^2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\begin{matrix} A & B & C & D \end{matrix}$$

We obtained the result as G^2=G^1 from the grouping of previous iteration and this iteration we conclude that the objects do not move group anymore. So the computation of the K-means clustering has reached its stability and no more iteration is needed. We conclude the final grouping as results.

Object	Feature 1 (X): ex	Feature 2(Y):	Group (result)
Medicine A	1	1	1
Medicine B	2	1	1
Medicine C	4	3	2
Medicine D	5	4	2

3.4 K-Means Algorithm Clustering Plot:

3.4.1 K-Means Algorithm:

The K-means Clustering Algorithm

The K-means algorithm can be considered the workhorse of clustering algorithms. It is a popular clustering method that minimizes the clustering error criterium. Here follows the K-means algorithm.

1. **Initialization:** Choose K vectors from the training vector set **X** at random. These vectors will be the initial centroids μk.

2. **Recursion:** For each vector x_n in the training set, let every vector belong to a cluster k. This is done by choosing the cluster centroid μk closest to the training vector x_n.

kn =

argmin k _

$(\mathbf{x}n - \mu k)T\ (\mathbf{x}n - \mu k)$ (3)

The function chooses the cluster which minimizes the Euclidean distance between the centroid and the training vector.

3. **Test:** Recompute the centroids μk by taking the mean of the vectors that belong to this cluster. This is done for all K centroids. If no vector belongs to μk, create a new centroid μk by assigning it a random vector from the training set. If none of the centroids μk changed from previous iteration, the algorithm terminates. Otherwise, go back to step 2.

4. **Termination:** From the clustering, the following parameters are found.

- The cluster centroids μk.
- The index k_n that indicates which centroid training vector **x**n belongs to. **Task 4-A**: Write a Matlab function for K-means clustering, kmeans. Use the function prototype listed below.

Task 4-B: Write a Matlab function to calculate the smallest distance from a feature vector to a cluster centroid, k means d. Use the function prototype listed below.

Task 4-C: In the file task4.mat there is a matrix of feature vectors, X. Find the centroids using the

K-means clustering algorithm, using K = 3 clusters. The feature vectors are two dimensional and the clustering can therefore be visualized and verified easily.

3.5 The FUZZY C-means Algorithm

Clustering is the process of grouping feature vectors into classes in the self organized mode. Choosing the cluster centers is crucial to the clustering. Fuzzy clustering plays an important role in solving problems in the areas of pattern recognition and fuzzy model identification. The FCM algorithm is more suited to data that is more or less evenly distributed around the cluster centers. The FCM algorithm lumps the two clusters with natural shapes but close boundaries into a large cluster.

Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by Dunn in 1973 and improved by Bezdek in 1981) is frequently used in pattern recognition. It is based on minimization of the following objective function.

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{i,j}^m \left\| x_i - c_j \right\|^2, 1 \le m \le \infty$$

where m is any real number greater than 1, $u_{i,j}$ is the degree of membership of x_i in the cluster j, x_i is the i th of d-dimensional measured data, c_j is the d-dimensional

center of the cluster, and ||*|| is any norm expressing the similarity between any measured data and the center.

Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership $u_{i,j}$ and the cluster centers c_j by:

$$u_{i,j} = 1/\sum_{k=1}^{C}\left(\left\|x_i - c_j\right\| / \left\|x_i - c_k\right\|\right)^{2/(m-1)}, c_j = \sum_{i=1}^{N}(u_{i,j}^m . x_i)/\sum_{i=1}^{N} u_{i,j}^m$$

CN *N*

This iteration will stop when $\max_{i,j}\{u_{i,j}(k+1) - u_{i,j}(k)\} < \varepsilon$, where ε is a termination criterion between 0 and 1, whereas k are the iteration steps. This procedure converges to a local minimum or a saddle point of J_m. A saddle point is a point in the domain of a function which is a stationary point but not a local extremum. The algorithm is described with the following steps:

1. Initialize $U = [u_{i,j}]$matrix,$U^{(0)}$
2. At k-step: calculae vectors $C^k = [c_j]$ with $U^{(k)}$
3. If $\|U^{(k+1)} - U^k\| < \varepsilon$ then stop; otherwise return to step 2.

3.6 Example-C Means Algorithm-Implementations

Clustering analysis is the process of grouping the unity according with the similar characters in the data concentration. It is widely applied in the field of data

$U^{(k)}, U^{(K+1)}$ drawing, image partition, model identification and signal compression.

The problem of first class clustering can be treated as restriction optimal problem. Its target is to search the optimal classification of sample set, to make clustering function based on error of intra class and inter-class. (The main difference between intra class and inter-class is the data are pooled to estimate the mean and variance)

C-Mean Algorithm (CMA) is the common way to solve this kind of problem. This algorithm is easy and its convergence speed is very fast but it is sensitive to the initialization condition and it has different clustering result for different initialization value. Clustering method based on genetics algorithm can solve the problems of initialization sensitivity of CMA and has a lot of chance to get the optimal solution. This example in virtue of optimum mechanism vertebrate immune system combining CMA puts forward one kind of hybrid clustering algorithm which keeps the mechanism of individual variety. And also put forward new immune selection strategy. Using this strategy we can overcome the immature convergence phenomenon.

Chapter 4:

Implementation

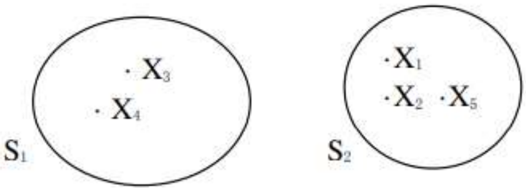
In this chapter we have discussed two basic examples along with two other practical examples. In the basic examples all the calculations are given so that readers can understand the practical examples properly.

4.1 Example 1

Let X={X_1 X_2 X_3 X_4 },there are three elements divided in two subsets which are non-empty by matrices.

$$U1 = \begin{matrix} & X_1 & X_2 & X_3 & X_4 & X_5 \\ S_1 & 0 & 0 & 1 & 1 & 0 \\ S_2 & 1 & 1 & 0 & 0 & 1 \end{matrix}$$

$$S_1 = \{X_3, X_4\},\ S_2 = \{X_1, X_2, X_5\}$$



$$U2 = \begin{matrix} & X_1 & X_2 & X_3 & X_4 & X_5 \\ S_1 & 1 & 0 & 0 & 1 & 1 \\ S_2 & 0 & 1 & 1 & 0 & 0 \end{matrix}$$

$$S_1 = \{X_1, X_4, X_5\},\ S_2 = \{X_2, X_3\}$$

Each element belongs only to one clustering

$X = \{X_1, X_2, X_3, X_4, X_5\}$

$$\tilde{u} = \begin{matrix} & X_1 & X_2 & X_3 & X_4 & X_5 \\ S_1 & 0 & 0.25 & 0.5 & 0.75 & 1 \\ S_2 & 1 & 0.75 & 0.5 & 0.25 & 0 \\ \sum = & 1 & 1 & 1 & 1 & 1 \end{matrix}$$

∧

Where $S_1 = 0/X_1 + 0.25/X_2 + 0.5/X_3 + 0.75/X_4 + 1/X_5$

$\wedge$

$S_2 = 1/X_1 + 0.75/X_2 + 0.5/X_3 + 0.25/X_4 + 0/X_5$

We use distance between two objects X_k and X_l as the function $d : X \times X \to R$ such that

$d(X_k, X_l) = d_{kl} \geq 0$

$d(X_k, X_l) = 0$ if $X_k = X_l$

$d(X_k, X_l) = d(X_l, X_k)$

Let $X = \{X_1, \ldots\ldots, X_n\}$, V_{cn} is the set of all real $c \times n$ matrices, where $l \leq c \leq n$ is an integer.

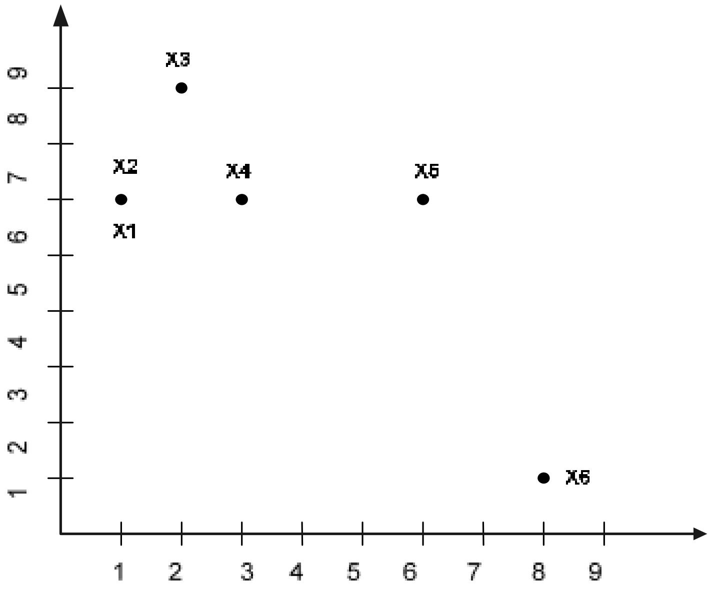
The matrix $u = (u_{ik}) \subset V_{cn}$ is called a crisp c-partition if

1. $u_{ik} \in \{0, 1\} \quad 1 \leq i \leq c$
2. $\sum_{i=1}^{c} u_{ik} = 1$
3. $0 < \sum_{k=1}^{n} u_{ik} < n$

4.2 Example 2

Here we have discussed the second basic example along with the whole algorithm. But in this example we have taken different parameters into account.

$X = \{(1, 6), (1, 6), (2, 8), (3, 6), (6,6), (8, 1)\}$



k=6, c=2, m=1

$$\begin{array}{c} \\ \tilde{u} = \begin{array}{c}\tilde{S}_1 \\ \tilde{S}_2\end{array} \end{array}\begin{array}{c} \begin{array}{cccccc} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 \end{array} \\ \begin{bmatrix} 0.6 & 0.6 & 0.5 & 0.4 & 0.4 & 0.4 \\ 0.4 & 0.4 & 0.5 & 0.6 & 0.6 & 0.6 \end{bmatrix} \end{array}$$

$$V_1 = \frac{0.6*(1,6)+0.6*(1,6)+0.5*(2,8)+0.4*(3,6)+0.4*(6,6)+0.4*(8,1)}{0.6+0.6+0.5+0.4+0.4+0.4}$$

$$= \frac{(9,\ 16.4)}{2.9} = (3.103,\ 5.655)$$

$$V_2 = \frac{0.4*(1,6)+0.4*(1,6)+0.5*(2,8)+0.6*(3,6)+0.6*(6,6)+0.6*(8,1)}{0.6+0.6+0.5+0.4+0.4+0.4}$$

$$= \frac{(12,\ 16.6)}{2.9} = (3.871,\ 5.355)$$

$d_{11}=d(x_1, v_1)=\sqrt{(3.103-1)^2+(5.655-6)^2}=2.131$

$d_{12}=d(x_2, v_1)=\sqrt{(3.103-1)^2+(5.655-6)^2}=2.131$

$d_{13}=d(x_3, v_1)=\sqrt{(3.103-2)^2+(5.655-8)^2}=2.591$

$d_{14}=d(x_4, v_1)=\sqrt{(3.103-3)^2+(5.655-6)^2}=0.360$

$d_{15}=d(x_5, v_1)=\sqrt{(3.103-6)^2+(5.655-6)^2}=2.917$

$d_{16}=d(x_6, v_1)=\sqrt{(3.103-8)^2+(5.655-1)^2}=6.756$

$d_{21}=d(x_2, v_1)=\sqrt{(3.871-1)^2+(5.355-6)^2}=2.943$

$d_{22}=d(x_2, v_2)=\sqrt{(3.871-1)^2+(5.355-6)^2}=2.943$

$d_{23}=d(x_2, v_3)=\sqrt{(3.871-2)^2+(5.355-8)^2}=3.240$

$d_{24}=d(x_2, v_4)=\sqrt{(3.871-3)^2+(5.355-6)^2}=1.084$

$d_{25}=d(x_2, v_5)=\sqrt{(3.871-6)^2+(5.355-6)^2}=2.225$

$d_{26}=d(x_2, v_6)=\sqrt{(3.871-8)^2+(5.355-1)^2}=6.001$

$$\tilde{u}^1_{22}=\frac{\frac{1}{d(X2,V2)}}{\frac{1}{d(X2,V1)}+\frac{1}{d(X2,V2)}}=\frac{\frac{1}{2.943}}{\frac{1}{2.131}+\frac{1}{2.943}}=\frac{2.131}{5.074}=0.420$$

$$\tilde{u}^1_{23}=\frac{\frac{1}{d(X3,V2)}}{\frac{1}{d(X3,V1)}+\frac{1}{d(X3,V2)}}=\frac{\frac{1}{3.240}}{\frac{1}{2.591}+\frac{1}{3.240}}=\frac{2.591}{5.831}=0.444$$

$$\tilde{u}^1_{11}=\frac{\frac{1}{d(X1,V1)}}{\frac{1}{d(X1,V1)}+\frac{1}{d(X1,V2)}}=\frac{\frac{1}{2.131}}{\frac{1}{2.131}+\frac{1}{2.943}}=\frac{2.943}{5.074}=0.580$$

$$\tilde{u}^1_{12}=\frac{\frac{1}{d(X2,V1)}}{\frac{1}{d(X2,V1)}+\frac{1}{d(X2,V2)}}=\frac{\frac{1}{2.131}}{\frac{1}{2.131}+\frac{1}{2.943}}=\frac{2.943}{5.074}=0.580$$

$$\tilde{u}^1_{13}=\frac{\frac{1}{d(X3,V1)}}{\frac{1}{d(X3,V1)}+\frac{1}{d(X3,V2)}}=\frac{\frac{1}{2.591}}{\frac{1}{2.591}+\frac{1}{3.240}}=\frac{3.240}{5.831}=0.556$$

$$\tilde{u}^1_{14}=\frac{\frac{1}{d(X4,V1)}}{\frac{1}{d(X4,V1)}+\frac{1}{d(X4,V2)}}=\frac{\frac{1}{0,360}}{\frac{1}{0.360}+\frac{1}{1.084}}=\frac{1.084}{1.444}=0.751$$

$$\tilde{u}^1_{15}=\frac{\frac{1}{d(X5,V1)}}{\frac{1}{d(X5,V1)}+\frac{1}{d(X5,V2)}}=\frac{\frac{1}{2.917}}{\frac{1}{2.917}+\frac{1}{2.225}}=\frac{2.225}{5.142}=0.433$$

$$\tilde{u}^1_{16}=\frac{\frac{1}{d(X6,V1)}}{\frac{1}{d(X6,V1)}+\frac{1}{d(X6,V2)}}=\frac{\frac{1}{6.756}}{\frac{1}{6.756}+\frac{1}{6.001}}=\frac{6.001}{12.757}=0.470$$

$$\tilde{u}^1_{21}=\frac{\frac{1}{d(X1,V2)}}{\frac{1}{d(X1,V1)}+\frac{1}{d(X1,V2)}}=\frac{\frac{1}{2.943}}{\frac{1}{2.131}+\frac{1}{2.943}}=\frac{2.131}{5.074}=0.420$$

$$\tilde{u}^1_{22}=\frac{\frac{1}{d(X2,V2)}}{\frac{1}{d(X2,V1)}+\frac{1}{d(X2,V2)}}=\frac{\frac{1}{2.943}}{\frac{1}{2.131}+\frac{1}{2.943}}=\frac{2.131}{5.074}=0.420$$

$$\tilde{u}^1_{23}=\frac{\frac{1}{d(X3,V2)}}{\frac{1}{d(X3,V1)}+\frac{1}{d(X3,V2)}}=\frac{\frac{1}{3.240}}{\frac{1}{2.591}+\frac{1}{3.240}}=\frac{2.591}{5.831}=0.444$$

$$\tilde{u}^1_{24}=\frac{\frac{1}{d(X4,V2)}}{\frac{1}{d(X4,V1)}+\frac{1}{d(X4,V2)}}=\frac{\frac{1}{1.084}}{\frac{1}{0.360}+\frac{1}{1.084}}=\frac{0.360}{1.444}=0.249$$

$$\tilde{u}^1_{25}=\frac{\frac{1}{d(X5,V2)}}{\frac{1}{d(X5,V1)}+\frac{1}{d(X5,V2)}}=\frac{\frac{1}{2.225}}{\frac{1}{2.917}+\frac{1}{2.225}}=\frac{2.917}{5.142}=0.567$$

$$\tilde{u}^1_{26}=\frac{\frac{1}{d(X6,V2)}}{\frac{1}{d(X6,V1)}+\frac{1}{d(X6,V2)}}=\frac{\frac{1}{6.001}}{\frac{1}{6.756}+\frac{1}{6.001}}=\frac{6.756}{12.757}=0.530$$

$$\tilde{u}^1 = \begin{matrix} \tilde{S}_1 \\ \tilde{S}_2 \end{matrix} \overset{\begin{matrix} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 \end{matrix}}{\begin{bmatrix} 0.580 & 0.580 & 0.556 & 0.751 & 0.433 & 0.470 \\ 0.420 & 0.420 & 0.444 & 0.249 & 0.467 & 0.530 \end{bmatrix}}$$

$$V_1^1=\frac{0.580*(1,6)+0.580*(1,6)+0.556*(2,8)+0.751*(3,6)+0.433*(6,6)+0.47*(8,1)}{0.580+0.580+0.556+0.751+0.433+0.47}$$
$$=\frac{(10.883,\ 18.982)}{3.37}=(3.229,\ 5.633)$$

$$V_2^1=\frac{0.420*(1,6)+0.420*(1,6)+0.444*(2,8)+0.249*(3,6)+0.567*(6,6)+0.53*(8,1)}{0.420+0.420+0.444+0.249+0.567+0.53}$$
$$=\frac{(10.117,\ 14.081)}{2.63}=(3.847,\ 5.330)$$

$d_{11}=d(x_1,\ v^1{}_1)=\sqrt{(3.229-1)^2+(5.633-6)^2}=2.259$

$d_{12}=d(x_2, v^1{}_1)=\sqrt{(3.229-1)^2+(5.633-6)^2}=2.259$

$d_{13}=d(x_3, v^1{}_1)=\sqrt{(3.229-2)^2+(5.633-8)^2}=2.667$

$d_{14}=d(x_4, v^1{}_1)=\sqrt{(3.229-3)^2+(5.655-6)^2}=0.433$

$d_{15}=d(x_5, v^1{}_1)=\sqrt{(3.229-6)^2+(5.633-6)^2}=2.795$

$d_{16}=d(x_6, v^1{}_1)=\sqrt{(3.229-8)^2+(5.633-1)^2}=6.650$

$d_{21}=d(x_1, v^1{}_2)=\sqrt{(3.847-1)^2+(5.330-6)^2}=2.925$

$d_{22}=d(x_2, v^1{}_2)=\sqrt{(3.847-1)^2+(5.330-6)^2}=2.925$

$d_{23}=d(x_3, v^1{}_2)=\sqrt{(3.847-2)^2+(5.330-8)^2}=3.247$

$d_{24}=d(x_4, v^1{}_2)=\sqrt{(3.847-3)^2+(5.330-6)^2}=1.080$

$d_{25}=d(x_5, v^1{}_2)=\sqrt{(3.847-6)^2+(5.330-6)^2}=2.255$

$d_{26}=d(x_6, v^1{}_2)=\sqrt{(3.847-8)^2+(5.330-1)^2}=6.000$

$$\tilde{u}^2{}_{11}=\frac{\frac{1}{d(X1,V^11)}}{\frac{1}{d(X1,V^11)}+\frac{1}{d(X1,V^12)}}=\frac{\frac{1}{2.259}}{\frac{1}{2.259}+\frac{1}{2.925}}=\frac{2.925}{5.184}=0.564$$

$$\tilde{u}^2{}_{12}=\frac{\frac{1}{d(X2,V^11)}}{\frac{1}{d(X2,V^11)}+\frac{1}{d(X2,V^12)}}=\frac{\frac{1}{2.259}}{\frac{1}{2.259}+\frac{1}{2.925}}=0.564$$

$$\tilde{u}^2{}_{13}=\frac{\frac{1}{d(X3,V^11)}}{\frac{1}{d(X3,V^11)}+\frac{1}{d(X3,V^12)}}=\frac{\frac{1}{2.667}}{\frac{1}{2.667}+\frac{1}{3.247}}=\frac{3.247}{5.914}=0.549$$

$$\tilde{u}^2{}_{14}=\frac{\frac{1}{d(X4,V^11)}}{\frac{1}{d(X4,V^11)}+\frac{1}{d(X4,V^12)}}=\frac{\frac{1}{0.433}}{\frac{1}{0.433}+\frac{1}{1.080}}=\frac{1.080}{1.513}=0.714$$

$$\tilde{u}^2{}_{15}=\frac{\frac{1}{d(X5,V^11)}}{\frac{1}{d(X5,V^11)}+\frac{1}{d(X5,V^12)}}=\frac{\frac{1}{2.795}}{\frac{1}{2.795}+\frac{1}{2.255}}=\frac{2.255}{5.05}=0.447$$

1

$$\tilde{u}^2_{16}=\frac{\frac{1}{d(X6,V^11)}}{\frac{1}{d(X6,V^11)}+\frac{1}{d(X6,V^12)}}=\frac{\frac{1}{6.650}}{\frac{1}{6.650}+\frac{1}{6.000}}=\frac{6.000}{12.650}=0.474$$

$$\tilde{u}^2_{21}=\frac{\frac{1}{d(X1,V^12)}}{\frac{1}{d(X1,V^11)}+\frac{1}{d(X1,V^12)}}=\frac{\frac{1}{2.925}}{\frac{1}{2.259}+\frac{1}{2.925}}=\frac{2.259}{5.184}=0.436$$

$$\tilde{u}^2_{22}=\frac{\frac{1}{d(X2,V^12)}}{\frac{1}{d(X2,V^11)}+\frac{1}{d(X2,V^12)}}=\frac{\frac{1}{2.925}}{\frac{1}{2.259}+\frac{1}{2.925}}=0.436$$

$$\tilde{u}^2_{23}=\frac{\frac{1}{d(X3,V^12)}}{\frac{1}{d(X3,V^11)}+\frac{1}{d(X3,V^12)}}=\frac{\frac{1}{3.247}}{\frac{1}{2.667}+\frac{1}{3.247}}=\frac{2.667}{5.914}=0.451$$

$$\tilde{u}^2_{24}=\frac{\frac{1}{d(X4,V^12)}}{\frac{1}{d(X4,V^11)}+\frac{1}{d(X4,V^12)}}=\frac{\frac{1}{1.080}}{\frac{1}{0.433}+\frac{1}{1.080}}=\frac{0,443}{1.513}=0.286$$

$$\tilde{u}^2_{25}=\frac{\frac{1}{d(X5,V^12)}}{\frac{1}{d(X5,V^11)}+\frac{1}{d(X5,V^12)}}=\frac{\frac{1}{2.255}}{\frac{1}{2.795}+\frac{1}{2.255}}=\frac{2.795}{5.05}=0.553$$

$$\tilde{u}^2_{26}=\frac{\frac{1}{d(X6,V^12)}}{\frac{1}{d(X6,V^11)}+\frac{1}{d(X6,V^12)}}=\frac{\frac{1}{6.000}}{\frac{1}{6.650}+\frac{1}{6.000}}=\frac{6.650}{12.650}=0.526$$

$$\tilde{u}^2 = \begin{array}{c} \\ \tilde{S}_1 \\ \tilde{S}_2 \end{array} \begin{array}{cccccc} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 \\ \end{array}\begin{bmatrix} 0.564 & 0.564 & 0.549 & 0.714 & 0.447 & 0.474 \\ 0.436 & 0.436 & 0.451 & 0.286 & 0.553 & 0.526 \end{bmatrix}$$

$$V_1^2 = \frac{0.564*(1,6)+0.564*(1,6)+0.549*(2,8)+0.714*(3,6)+0.447*(6,6)+0.474*(8,1)}{0.564+0.564+0.549+0.714+0.447+0.474}$$

$$= \frac{(10.842,\ 18.6)}{3.312} = (3.274,\ 5.616)$$

$$V_2^2 = \frac{0.436*(1,6)+0.436*(1,6)+0.451*(2,8)+0.286*(3,6)+0.553*(6,6)+0.526*(8,1)}{0.436+0.436+0.451+0.286+0.553+0.526}$$

$$= \frac{(10.158,\ 14.4)}{2.688} = (3.779,\ 5.357)$$

$d_{11}=d(x_1, v^3_1)=\sqrt{(3.274-1)^2+(5.616-6)^2}$ =2.306

$d_{12}=d(x_2, v^3_1)=\sqrt{(3.274-1)^2+(5.616-6)^2}$ =2.306

$d_{13}=d(x_3, v^3_1)=\sqrt{(3.274-2)^2+(5.616-8)^2}$ =2.703

$d_{14}=d(x_4, v^3_1)=\sqrt{(3.274-3)^2+(5.616-6)^2}$ =0.472

$d_{15}=d(x_5, v^3_1)=\sqrt{(3.274-6)^2+(5.616-6)^2}$ =2.753

$d_{16}=d(x_6, v^3_1)=\sqrt{(3.229-8)^2+(5.616-1)^2}$ =6.606

$d_{21}=d(x_1, v^3_1)=\sqrt{(3.779-1)^2+(5.357-6)^2}$ =2.852

$d_{22}=d(x_2, v^2_2)=\sqrt{(3.779-1)^2+(5.357-6)^2}$ =2.852

$d_{23}=d(x_3, v^2_2)=\sqrt{(3.779-2)^2+(5.357-8)^2}$ =3.186

$d_{24}=d(x_4, v^2_2)=\sqrt{(3.779-3)^2+(5.357-6)^2}$ =1.01

$d_{25}=d(x_5, v^2_2)=\sqrt{(3.779-6)^2+(5.357-6)^2}$ =2.312

$d_{26}=d(x_6, v^2_2)=\sqrt{(3.779-8)^2+(5.357-1)^2}$ =6.066

$$\tilde{u}^3_{11}=\frac{\frac{1}{d(X1,V^21)}}{\frac{1}{d(X1,V^21)}+\frac{1}{d(X1,V^22)}}=\frac{\frac{1}{2.306}}{\frac{1}{2.306}+\frac{1}{2.852}}=\frac{2.852}{5.158}=0.553$$

$$\tilde{u}^3_{12}=\frac{\frac{1}{d(X2,V^21)}}{\frac{1}{d(X2,V^21)}+\frac{1}{d(X2,V^22)}}=\frac{\frac{1}{2.306}}{\frac{1}{2.306}+\frac{1}{2.852}}=0.553$$

$$\tilde{u}^3_{13}=\frac{\frac{1}{d(X3,V^21)}}{\frac{1}{d(X3,V^21)}+\frac{1}{d(X3,V^22)}}=\frac{\frac{1}{2.703}}{\frac{1}{2.703}+\frac{1}{3.186}}=\frac{3.186}{5.889}=0.541$$

$\sqrt{(3.274-1)^2+(5.616-6)^2}$

$\sqrt{(3.274-1)^2+(5.616-6)^2}$

$\sqrt{(3.274-2)^2+(5.616-8)^2}$

$\sqrt{(3.274-3)^2+(5.616-6)^2}$

$\sqrt{(3.274-6)^2+(5.616-6)^2}$

$\sqrt{(3.229-8)^2+(5.616-1)^2}$

$\sqrt{(3.779-1)^2+(5.357-6)^2}$

$\sqrt{(3.779-1)^2+(5.357-6)^2}$

$\sqrt{(3.779-2)^2+(5.357-8)^2}$

$) = \sqrt{(3.779-3)^2+(5.357-6)^2} \quad =1.01$

$\sqrt{(3.779-6)^2+(5.357-6)^2}$

$\sqrt{(3.779-8)^2+(5.357-1)^2}$

$$\tilde{u}^3{}_{14}=\frac{\frac{1}{d(X4,V^21)}}{\frac{1}{d(X4,V^21)}+\frac{1}{d(X4,V^22)}}=\frac{\frac{1}{0.472}}{\frac{1}{0.472}+\frac{1}{1.01}}=0.682$$

$$\tilde{u}^3{}_{15}=\frac{\frac{1}{d(X5,V^21)}}{\frac{1}{d(X5,V^21)}+\frac{1}{d(X5,V^22)}}=\frac{\frac{1}{2.753}}{\frac{1}{2.753}+\frac{1}{2.312}}=0.456$$

$$\tilde{u}^3{}_{16}=\frac{\frac{1}{d(X6,V^21)}}{\frac{1}{d(X6,V^21)}+\frac{1}{d(X6,V^22)}}=\frac{\frac{1}{6.606}}{\frac{1}{6.606}+\frac{1}{6.066}}=0.479$$

$$\tilde{u}^3{}_{21}=\frac{\frac{1}{d(X1,V^22)}}{\frac{1}{d(X1,V^21)}+\frac{1}{d(X1,V^22)}}=\frac{\frac{1}{2.852}}{\frac{1}{2.306}+\frac{1}{2.852}}=0.447$$

$$\tilde{u}^3{}_{22}=\frac{\frac{1}{d(X2,V^22)}}{\frac{1}{d(X2,V^21)}+\frac{1}{d(X2,V^22)}}=\frac{\frac{1}{2.852}}{\frac{1}{2.306}+\frac{1}{2.852}}=0.447$$

$$\tilde{u}^3{}_{23}=\frac{\frac{1}{d(X3,V^22)}}{\frac{1}{d(X3,V^21)}+\frac{1}{d(X3,V^22)}}=\frac{\frac{1}{3.186}}{\frac{1}{2.703}+\frac{1}{3.186}}=\frac{2.703}{5.889}=0.459$$

$$\tilde{u}^3{}_{24}=\frac{\frac{1}{d(X4,V^22)}}{\frac{1}{d(X4,V^21)}+\frac{1}{d(X4,V^22)}}=\frac{\frac{1}{1.01}}{\frac{1}{0.472}+\frac{1}{1.01}}=\frac{0.472}{1.482}=0.318$$

$$\tilde{u}^3{}_{25}=\frac{\frac{1}{d(X5,V^22)}}{\frac{1}{d(X5,V^21)}+\frac{1}{d(X5,V^22)}}=\frac{\frac{1}{2.312}}{\frac{1}{2.312}+\frac{1}{2.753}}=\frac{2.753}{5.065}=0.544$$

$$\tilde{u}^3{}_{26}=\frac{\frac{1}{d(X6,V^22)}}{\frac{1}{d(X6,V^21)}+\frac{1}{d(X6,V^22)}}=\frac{\frac{1}{6.066}}{\frac{1}{6.066}+\frac{1}{6.606}}=\frac{6.606}{12.672}=0.521$$

$d_{11}=d(x_1,v^3{}_1)=$

=2.306

$$\tilde{u}^3 = \begin{matrix} & X_1 & X_2 & X_3 & X_4 & X_5 & X_6 \\ \tilde{S}_1 & 0.553 & 0.553 & 0.541 & 0.682 & 0.456 & 0.470 \\ \tilde{S}_2 & 0.447 & 0.447 & 0.459 & 0.318 & 0.544 & 0.520 \end{matrix}$$

4.3 Fuzzy clustering analysis and Fuzzy C-means algorithm Implementations

The impact of hydrological forecasting results mainly performance uncertain factor such as hydrology, meteorology, geography and geology. These factors have a certain amount of random and fuzzy. How to make use of fuzzy analysis to such statistics these uncertain factors, combining with regional runoff model and complementing classification of real-time flood forecasting is the focus of this study. We have chosen Chao ER river of Lu Zhou station as a forecasting station and have tried to use fuzzy math and hydrological forecasting knowledge of history to extract the impact feature from the historical occurred floods and select the most appropriate impact factor to be fuzzy clustering. Fuzzy clustering shows the impact feature of same type of history flooding has same degree of influence to flooding. In real-time forecasting, the impact characteristics of real-time flooding and the characteristics of a sample historical flooding are analyzed by fuzzy clustering to get the membership of a real-time occurrence flooding relative to different categories of flood and do operations in real-time flood forecasting.

4.3.1 Fuzzy clustering analysis

Hydrological forecasting factor is very vague, as a result that the fuzzy cluster analysis is applied to classification of flood forecasting flood categories in order to make classification more realistic, the gotten result is more comprehensive, detailed, reasonable than the traditional classification but also provide the classification base for finial real-time forecasting.

The generally welcomed practical application is the fuzzy clustering method based on the objective function that is, clustering is boiled down into a constrained non-linear programming problem by optimizing solution to obtain the fuzzy partition and cluster of data sets. The method is designed simple a wide range of problemsolving and easy to achieve on the computer. In fuzzy clustering algorithm based on the objective function the algorithm theory of Fuzzy C-MEANS is better and applied more widely.

However, C-MEANS algorithm does not consider the problem of extent impact of the hydrological forecasting impact factor on forecasting results but in the actual application due to the calculation and theoretical reasons of prediction model the contribution of impact factor to the prediction result is non-uniform and in accordance with Principle calculation of the forecasting mode select impact factor one part is measured hydrologic data, the other part is the hydrological properties data. In order to reflect the extent of impact of numerical data on the forecasting results weighted optimization is complemented in sample feature with feature selection techniques Relief.

4.3.2 The fuzzy c-means clustering based on weighted feature

In order to consider each dimensional feature of vector sample of historical flood make contribution to pattern classification, to make use of feature selection technique, to put forward a C-MEAN fuzzy clustering algorithm based on the characteristic weighting, which makes use of feature selection technology features RELIEF algorithm to weighted select feature not only made the classification more effective than an ordinary C-MEAN fuzzy clustering algorithm and can also analyze the contributions and degree of impact of each dimensional characteristics to the classification results and provide a reference for the feature extraction of real-time forecasting.

When the characteristics weight will be directly applied to fuzzy C-MEAN clustering, the objective function fuzzy clustering can be expressed as:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{i,j}^{m} \left\| x_i - c_j \right\|^2, 1 \le m \le \infty$$

J_m (where m=1,……n) denotes the weight of each-dimensional characteristics of the sample. When the objective function J_m achieves the minimum, it gets the optimal result of fuzzy clustering, it is needed to point out that the samples of historical flood contain part of flooding property data needing artificial evaluation, a part of property data is not involved in the calculation of flood forecasting model which mainly works in fuzzy classification, therefore the weighted value of this part of attribute characteristics is defined as a relatively small number in order to reduce the impact made by artificial evaluation error of attribute data to forecasting result in real-time forecasting. At last relevance between reference characteristics factor and prediction objection make the posted-weight to all of the preferred characteristics.

Due to the limitations and particularity existing in the acquisition of historical flood samples in general, the big of basin flood is medium-sized floods, low incidence rate to the large flood. In the sample collection process, the difference between the quantities of various types of samples is very large, from the point of view of the calculation method of characteristics weight value. This will be the difference between the quantity will make a large effect on selection of sample characteristics. In order to minimize this impact as much as possible. In accordance with the distribution of all kinds of different samples, set number which is most close to points to be dynamic in RELIEF algorithm.

4.4 Example analysis calculation

In order to test the validity of fuzzy C-means algorithm -weighted based on the characteristics of weight, the proposed accuracy of historical flooding of "CAO ER river" basin is comparably analyzed around the proposed fuzzy by this method.

4.4.2 Flooding Feature Selection

The key of FCM algorithm lies the prior given cluster number of C, therefore, this review accords to the iterative self-organizing analysis technology, uses meaning of optimal classification to build function F of inter-class and extra-class distance of Fuzzy Cluster Analysis to achieve the C-MEAN fuzzy optimization algorithm together, to obtain the best cluster sample in this example for Category 5.

Since the prediction model applied the runoff model used in sub-humid areas, therefore firstly select rainfall intensity in greatest time, initial water content of soil, the average rainfall intensity and the initial flow to take part in model calculation which is characterized by basic numerical characteristics of fuzzy clustering. However, the hydrological element influencing formation of the impact of flooding has a strong randomness and fuzziness, and uncertainty predictor factor is very powerful, in order to fully reflect basin rainfall characteristics and distribution of rainfall. In accordance with distribution formation of pre-rainfall and its extend contribution to the flood peak, select three property impact characters such as a rainstorm center, rain trend and property on shape of rain type. In the calculation of fuzzy clustering analysis, it is necessary that the property characteristic numerical value of rainfall changes into a numerical value characteristics. Location of rainstorm center accords the five levels: the upper reaches, middle and upper reaches, middle reaches, the middle and lower reaches, evaluates from 1to 5 according to their degree of influence of flooding peak. Trend of rainfall accords from bottom to top, upstream, midstream, from top to bottom, lower reaches of the five-levels, gives evaluation from 1to 5 according to their degree of influence on the flooding speak. Shape of rain type interpolates evaluation in accordance with the figure 4-1.

The rain characteristics during the largest time based on forecasting calculating period as unity select the greatest periods of rainfall making impact on flood peak. Initial water content of soil refers to the water content of soil during the early period of prediction calculation. The average rainfall intensity is the peak-hour rainfall of rain peak. The initial flow refers to the flow value of early period of forecasting calculation in the forecast section. These characteristics compose series of characteristics samples of the basin historical flood, shown in Table 4- 14-1.

Property	1	2	3

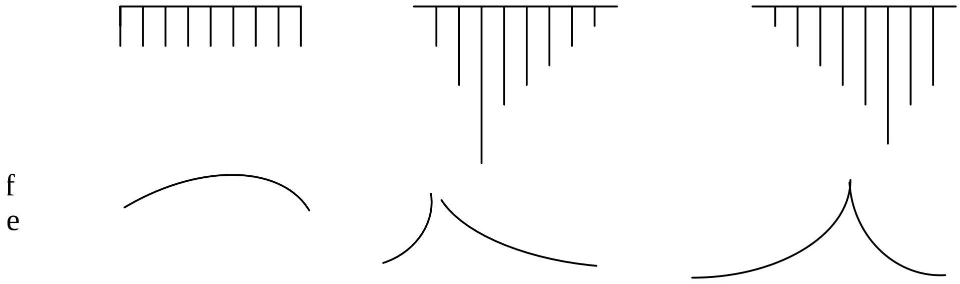


Figure 4-1: Sketch for evaluate property on the shape of rain type

Table 4- 1: Feature and flood peaks runoff of historical flood sample Table 4-2: Sample data for rainfall intensity according to optimal classification

Flood No.	No.	initial water content /mm	average rainfall intensity/(mm*h^{-1})	initial flow/(m^3*s^{-1})	flow of flooding peak/(m^3*s^{-1})
860410	1	89.1	12.6	74.51	691
860708	2	85	15.7	60.98	532
860905	3	69.8	19.7	38.6	937
860919	4	68.6	7.2	33.37	630
870620	5	68.7	7.2	33.55	1220
870722	6	43.2	7	93.13	592
870727	7	90	16.6	194	2093
870908	8	59.1	13.2	89.4	1490
880617	9	74.3	14.9	48.23	1080
880807	10	28.5	32	71.2	1389

890411	11	51.8	5.3	12.7	660
890522	12	81.6	8.7	241	825

Chapter 5:

Conclusion

The C-means algorithm is treated as a new search operator so that proposed algorithm obtains the characters of strong ability of part searching and small operation in order to faster the part search optimum speed. Therefore, the textual algorithm can converge to the optimum faster and has higher accuracy. The algorithm can be extended to other clustering model whose objective function can be represented in terms of optimization of clustering centres.

From the features, concepts and properties discussed about pattern recognition it is evident that Fuzzy logic can indeed be part of this vast universe of data thereby finding itself in various applications and measures in relation with pattern recognition.

5.1 Future of technology

The future regarding pattern recognition still looks bright. As long as man's determination to solve every complex problem precisely and perfectly persists, one can never see the horizon beyond i.e. the philosophy that to solve complex problems approximate reasoning and imperfections are good enough and this is where fuzzy logic and its variants, in this case fuzzy data and different patterns, will remain prevalent.

It has been a wonderful experience to learn new vast universe in the form of Fuzzy Logic and having the knowledge of Fuzzy pattern recognition has indeed given a different dimension. The thesis presented has done quite a bit of justice in knowing about this topic right from introduction to fuzzy logic and pattern recognition, the reasons and its uses, to C means algorithms enhancement, segmentation using Fuzzy and finally its applications.

REFERENCES

[1] K means clustering-algorithms A tutorial on clustering http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/kmeans.html

[2] Sheng W , TuckerA, Liu X. Clustering with Nichinggenetic K-means algorithm Proceedings of Genetic and Evolutionary Computation Conference
Berlin: Springer-Verlag, 2004.162-173.

[3] Maulik U, Bandyopadhyay S. Genetic algorithm-based Clustering technique, PatternRecognition. 1997.

[4] De Castro L N, von Zuben F J. Learning and optimization using the clonal selection principle, IEEE Transactionon Evolutionary Computation, 2002.

[5] ZHANG Honggang , GUO Shenglian . Recent advancement and prospect of hydrological forecasting uncertainty study. J of Shihezl University: Natural
Science Edition 2006, 24(I):15-21.

[6] San jose state university, Fuzzy Logic, available at: http://www.sjsu.edu/faculty/watkins/fuzzysets.htm

[7] Fuzzy logic and introduction by Steven D.Kaehler, available:http://www.seattlerobotics.org/Encoder/mar98/fuz/fl_part4.html#MEMBERSHIP FUNCTIONS

[8] Fuzzifying Clustering algorithms,The case study of major clust, Eugene levner,David pinto,availableat:
http://users.dsic.upv.es/~prosso/resources/LevnerEtAl_MICAI07.pdf

[9] A tutorial on clustering algorithms, Fuzzy C-means Clustering, available at:
http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/cmeans.html

[10] J. C. Bezdek (1981): "Pattern Recognition with Fuzzy Objective Function Algoritms", Plenum Press, New York

[11] Pattern Recognition with Fuzzy Objective Function Algorithms J.C Bezdek 1981 Kluwer Academic Publishers, Norwell, MA, 1981

[12] Data Mining. A Knowledge Discovery Approach Cios, K.J., Pedrycz, W., Swiniarski, R.W., Kurgan, L.A.. 2007

[13] Pattern Recognition S. Theodoridis, K. Koutroumbas

[14] Algorithms for Fuzzy Clustering methods in c-means clustering with applications Miyamoto S, Ichihashi H, Honda K.

[15] MATLAB Fuzzy Logic tool box, J.-S. Roger Jang Ned Gulley

[16] K-Means Clustering Tutorial By Kardi Teknomo,PhD

[17] Fuzziness and Probability Based on texts from Jim Bezdek, Didier Dubois, Bart Kosko, and Henri Prade, available at: http://www.neuronet.pitt.edu/~bogdan/research/fuzzy/fvsp/fvsp.html

[18] flexible completion of workflow activities, Georg Peters-Munich University of Applied Sciences, Department of Computer Science, Roger Tagg- University of South Australia, School of Computer and Information Science