



APPLICATION OF FUZZY LOGIC IN WEB MINING DOMAIN: A SURVEY

Mamta Kathuria*

Neelam Duhan*

C.K.Nagpal**

Abstract: *This paper summarizes the different characteristics of web data, the basic components of web mining and its different types, and importance of fuzzy logic its representation and the use of soft computing (comprising fuzzy logic (FL)) is discussed with the present web mining techniques. Fuzzy logic is used with one of the web mining technique i.e Clustering. So clustering is also discussed. Clustering is a widely used technique in data mining application for discovering patterns in underlying data. Most traditional clustering algorithms are limited in handling datasets that contain categorical attributes. Conventional clustering means classifying the given data objects as exclusive subsets (clusters). That means we can discriminate clearly whether an object belongs to a cluster or not. However such a partition is insufficient to represent many real situations. However, in many real situations, there not exists an exact boundary between different clusters. Therefore a fuzzy clustering object belongs to overlapping clusters with some membership degree. In other words, the essence of fuzzy clustering is to consider not only the belonging status to the clusters, but also to consider to what degree do the object belong to the cluster.*

Keywords: *Web mining, Classical Logic, Fuzzy logic (FL), Clustering, Fuzzy Clustering.*

*Department of computer Engineering, YMCA University of science & Technology, Faridabad

**Principal, Echleon Institute of Technology, Faridabad



1. INTRODUCTION

The World Wide Web can be seen as one of the largest databases in the world. This huge, and ever-growing, amount of data is a fertile area for data mining research. Data mining [6] as the process of extracting previously unknown information from (usually large quantities of) data, which can, in the right context, lead to knowledge. When data mining techniques are applied to web data, it becomes web mining. Web mining [1] as the whole of data mining and related techniques that are used to automatically discover and extract information from web documents and services. Knowledge discovered after mining the web, could pose a threat to people, when for instance personal data is misused. However, it is this same knowledge factor that can imply lots of different advantages, as it is of high value to all sorts of applications concerning planning and control. Kosala, Blockeel and Neven have already described some specific benefits of web mining, like improving the intelligence of search engines. Web mining can also contribute to marketing intelligence by analyzing the web user's on-line behavior and turning this information into marketing knowledge. Web mining is a technique with a large number of good qualities and potential. It shall be made clear that there are different ways to mine the web. By describing some scenarios this paper illustrate those categories.

Web data are those that can be collected and used in the context of Web personalization. These data are classified in four categories:

- **Content data** are presented to the end-user appropriately structured. They can be simple text, images, or structured data, such as information retrieved from databases.
- **Structure data** represent the way content is organized. They can be either data entities used within a Web page, such as HTML or XML tags, or data entities used to put a Web site together, such as hyperlinks connecting one page to another.
- **Usage data** represent a Web site's usage, such as a visitor's IP address, time and date of access, complete path (files or directories) accessed, referrers address, and other attributes that can be included in a Web access log.
- **User profile data** provide information about the users of a Web site. A user profile contains demographic information for each user of a Web site, as well as information about users' interests and preferences. Such information is acquired through registration forms or questionnaires, or can be inferred by analyzing Web usage logs.



The data components for web mining can involve user-provided information, server log files, cookies, form-generated datasets, email, as well as commercial demographics, life-style information, previous browsing activity, prior sales, transactional data, search keywords, redirects or referrals, and other consumer related behavior. In web mining responses and offers are driven by real time events and interactions. Web mining unlike data mining which existed prior the Internet explosion involves a new paradigm of data collection, integration and analysis. Web mining involves pattern recognition via a seamless stream of activity taking place over a decision network and not a static warehouse. Web mining works by performing data analysis via networks, using software agents to mine, collaborate and discover conditions and features which can lead to increases in sales, cross-selling opportunities and the targeting of specific products or services.

Web mining provides an enterprise the integrated tools for analyzing all type of data sources from multiple departments, from different locations from different formats for an assortment of deliverables such as propensity to purchase scores, risk scores for fraud, prediction of customer behavior or the creation of customer categories or groupings. Web mining enables an enterprise to leverage their everyday communications with actionable knowledge discovery, real-time business intelligence and targeted customer responses. Web mining enables an enterprise to make the right offer, to the right customer, as events take place in real time.

These data types makes the three axes of Web Mining, That are also known as web mining categories. The Web Mining categories are discussed below.

2. CATEGORIES OF WEB MINING

Web mining can involve all of the traditional data mining processes of classification, segmentation, clustering, association, prediction, and modeling the only difference is that the analyzes result in immediate action. Unlike data mining, web mining is dependent on the use of software agent to trigger targeted offers as events take place in real time. An agent is a program that takes action on behalf of a process, which in this case can be cross and up selling, customer retention, risk assessment, fraud detection and counterterrorism.

There are different ways to mine the web. The different ways to mine the web are closely related to the different types of web data. Actual data on web pages can be distinguished as, web structure data regarding the hyperlink structure within and across web documents,



and web log data regarding the users who browsed the web pages. Therefore, in accordance with Madria et al (1999), we shall divide web mining into three categories.

2.1 Web Content Mining (WCM)

Web content mining[8] examines the content of Web pages as well as results of Web searching. The content includes text as well as graphics data. Web content mining is further divided into Web page content mining and search results mining. Web content mining has to do with the retrieval of information (content) available on the Web into more structured forms as well as its indexing for easy tracking information locations. Web content may be unstructured (plain text), semistructured (HTML documents), or structured (extracted from databases into dynamic Web pages). Such dynamic data cannot be indexed and consist what is called “the hidden Web”. A research area closely related to content mining is text mining. Web content mining uses data mining techniques for efficiency, effectiveness and scalability.

Web content mining can be divided into:

- a) Agent-based approach
- b) Database-based approach

a) Agent-based approach

They have software systems (agents) that perform the content mining. For example, search engines. Intelligent search agents use techniques like user profiles or knowledge concerning specific domains. Information filtering utilizes IR techniques, knowledge of the link structures to retrieve and categorize documents. Personalize Web agents use information about user preferences to direct their search.

b) Database-based Approach

It views the Web data as belonging to a database. There have been approaches that view the Web as a multilevel database, and there have been many query languages that target the Web.

2.2 Web Structure Mining (WSM)

Web structure mining[8], which focuses on link information. It aims to analyze the way in which different web documents are linked together. With Web structure mining, information is obtained from the actual organization of pages on the Web. Discovering the model underlying link structures (topology) on the Web. E.g. discovering authorities and



hubs. Web structure mining is the process of using graph theory to analyze the node and connection structure of a web site. According to the type of web structural data, web structure mining can be divided into two kinds:

- 1) Extracting patterns from hyperlinks in the web: a hyperlink is a structural component that connects the web page to a different location.
- 2) It can be viewed as creating a model of the Web organization or a portion of it.

2.3 Web Usage Mining (WUM)

Usage mining analyses the transaction data that is logged when users interact with the web. Usage mining is sometimes referred to as 'log mining', because it involves mining the web server logs. Web usage mining is the process of extracting useful information from server logs i.e. users history. Web usage mining is the process of finding out what users are looking for on the Internet. Some users might be looking at only textual data, whereas some others might be interested in multimedia data. Goal of WUM[8] to analyze the behavioral patterns and profiles of users interacting with a Web site. The discovered patterns are usually represented as collections of pages, objects, or resources that are frequently accessed by groups of users with common interests. Data that are to be used in Web Usage Mining includes web server logs, site contents, data about the visitors, gathered from external channels, further application data. As there is large amount of information available on the web, finding the relevant information is not an easy task. Moreover, there are various ambiguities in the web data. Web data is high dimensional, overlapped in nature, heterogeneous, uncertain, imprecise, unreliable and can have several meanings. Soft Computing techniques can help handle such type of situations. In the next section, a discussion about soft computing and its relevance on the web has been carried out.

3 SOFT COMPUTING

A collection of methodologies which provide information processing capabilities for handling real life ambiguous situations.

Some domains of Soft Computing [7]

- a) Fuzzy Logic
- b) Neural Networks
- c) Genetic Algorithms
- d) Rough Sets



3.1 Soft Computing and Its Relevance

Its aim is to exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve tractability, robustness, low-cost solutions, and close resemblance to human-like decision making. Fuzzy Logic (FL) [2,3] provides algorithms for dealing with imprecision and uncertainty arising from vagueness rather than randomness, Rough Sets(RS) for handling uncertainty arising from limited discernibility of objects, Artificial Neural Network(ANN) the machinery for learning and adaptation, and Genetic Algorithm(GA) for optimization and searching. In the further section we will discuss about one area of soft computing i.e. Fuzzy logic. In the next section classical set theory is being discussed in brief before the detail description of fuzzy logic has been carried out.

4 CLASSICAL SET THEORY

Classical set theory allows the membership of the elements in the set in binary terms, a bivalent condition i.e. - an element either belongs or does not belong to the set.

Fuzzy set theory permits the gradual assessment of the membership of elements in a set, that is described with the aid of a membership function valued in the real unit interval [0, 1].

Example: Words like young, tall, good, or high are fuzzy.

- There is no single quantitative value which defines the term young.
- For some people, age 25 is young, and for others, age 35 is young.
- The concept young has no clean boundary.
- Age 20 is definitely young and age 100 is definitely not young;
- Age 35 has some possibility of being young and usually depends on the context in which it is being considered.
- Fuzzy Set theory is an extension of classical set theory where elements have degrees of membership.

A Set is any well defined collection of objects. An object in a set is called an element or member of that set. Sets are defined by a simple statement describing whether a particular element having a certain property belongs to that particular set. Classical set theory enumerates all its elements using

$$A = \{a_1, a_2, a_3, a_4 \dots a_n\}$$



If the elements a_i ($i = 1, 2, 3 \dots n$) of a set A are subset of universal set X, then set A can be represented for all elements $x \in X$ by its characteristic function

$$A = \begin{cases} 1 & \text{if } x \in X \\ 0 & \text{otherwise} \end{cases}$$

–A set A is well described by a function called characteristic function.

This function, defined on the universal space X, assumes:

- a value of 1 for those elements x that belong to set A, and
- a value of 0 for those elements x that do not belong to set A.

The notations used to express these mathematically are

$$\left. \begin{array}{l} A : X \rightarrow [0,1] \\ A(x) = 1 \quad \quad \quad x \text{ is a member of A} \\ A(x) = 0 \quad \quad \quad x \text{ is not a member of} \end{array} \right\} \text{Eq.(1)}$$

Alternatively, the set A can be represented for all elements $x \in X$ by its characteristic function $A(x)$ defined as

$$A = \begin{cases} 1 & \text{if } x \in X \\ 0 & \text{otherwise} \end{cases}$$

Thus in classical set theory A (x) has only the values 0 ('false') and 1 ('true'). Such sets are called crisp sets. In classical theory there is no intermediate term that can be used between 0 and 1, so fuzzy set theory has been proposed that is to be discussed in the next section.

5 FUZZY SET THEORY

Fuzzy set theory[2,3] is an extension of classical set theory where elements have varying degrees of membership. A logic based on the two truth values, True and False, is sometimes inadequate when describing human reasoning. Fuzzy logic uses the whole interval between 0 (false) and 1 (true) to describe human reasoning.

- A Fuzzy Set is any set that allows its members to have different degree of membership, called membership function, in the interval [0, 1].
- Fuzzy logic is derived from fuzzy set theory dealing with reasoning that is approximate rather than precisely deduced from classical predicate logic.
- Fuzzy Logic is used to manage continuous change.
- Fuzzy logic is used where boundary of indecision is high.



- Fuzzy logic is capable of handling inherently imprecise concepts.
- Fuzzy logic allows in linguistic form the set membership values to imprecise concepts like "slightly", "quite" and "very".
- Fuzzy set theory defines Fuzzy Operators on Fuzzy Sets.

5.1 Fuzzy Logic

The modeling of imprecise and qualitative knowledge, as well as handling of uncertainty at various stages is possible through the use of fuzzy sets. Fuzzy logic[2,3] is capable of supporting, to a reasonable extent, human type reasoning in natural form by allowing partial membership for data items in fuzzy subsets. Integration of fuzzy logic with data mining techniques has become one of the key constituents of soft computing in handling the challenges posed by the massive collection of natural data .Fuzzy logic is logic of fuzzy sets. A Fuzzy set has, potentially, an infinite range of truth values between one and zero. Propositions in fuzzy logic have a degree of truth, and membership in fuzzy sets can be fully inclusive, fully exclusive, or some degree in Between. The fuzzy set is distinct from a crisp set is that it allows the elements to have a degree of membership. The core of a fuzzy set is its membership function: a function which defines the relationship between a value in the sets domain and its degree of membership in the fuzzy set . The relationship is functional because it returns a single degree of membership for any value in the domain [2].

$$\mu=f(s,x)..... (2)$$

Here, μ : is he fuzzy membership value for the element

s: is the fuzzy set

x : is the value from the underlying domain.

Fuzzy sets provide a means of defining a series of overlapping concepts for a model variable since it represent degrees of membership. The values from the complete universe of discourse for a variable can have memberships in more than one fuzzy set.

Fuzzy logic [3] is a form of many-valued logic or probabilistic logic; it deals with reasoning that is approximate rather than fixed and exact. In contrast with traditional logic theory, where binary sets have two-valued logic: true or false, fuzzy logic variables may have a truth value that ranges in degree between 0 and 1. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true and completely false. The reasoning in fuzzy logic is similar to human reasoning. It allows for

approximate values and inferences as well as incomplete or ambiguous data (fuzzy data) as opposed to only relying on crisp data (binary yes/no choices). Fuzzy logic is able to process incomplete data and provide approximate solutions to problems other methods find difficult to solve. For example, let a 100 ml glass contain 30 ml of water. Then we may consider two concepts: Empty and Full. The meaning of each of them can be represented by a certain fuzzy set. Then one might define the glass as being 0.7 empty and 0.3 full. Note that the concept of emptiness would be subjective and thus would depend on the observer or designer. Another designer might equally well design a set membership function where the glass would be considered full for all values down to 50 ml. It is essential to realize that fuzzy logic uses truth degrees as a mathematical model of the vagueness phenomenon while probability is a mathematical model of ignorance.

Example of fuzzy logic

Fuzzy logic temperature

In this image, the meanings of the expressions cold, warm, and hot are represented by functions mapping a temperature scale. A point on that scale has three "truth values"—one for each of the three functions. The vertical line in the image represents a particular temperature that the three arrows (truth values) gauge. Since the red arrow points to zero, this temperature may be interpreted as "not hot". The orange arrow (pointing at 0.2) may describe it as "slightly warm" and the blue arrow (pointing at 0.8) "fairly cold".

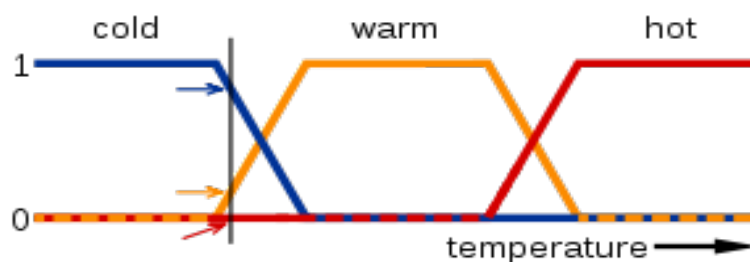


Fig 5.1: Representation of fuzzy set temperature

A fuzzy set A is written as a set of pairs $\{x, A(x)\}$ as

$A = \{\{x, A(x)\}, x \text{ in the set } X, \text{ where } x \text{ is an element of the universal space } X, \text{ and } A(x) \text{ is the value of the function } A \text{ for this element.}$

The value $A(x)$ is the membership grade of the element x in a fuzzy set A.

Representation of Crisp and Non-Crisp Set

Example: Classify students for a basketball team

This example explains the grade of truth value.

- tall students qualify and not tall students do not qualify
- if students 1.8 m tall are to be qualified, then should we exclude a student who is 1/10" less? or should we exclude a student who is 1" shorter?

Representation of crisp and non-crisp set of a tall person.

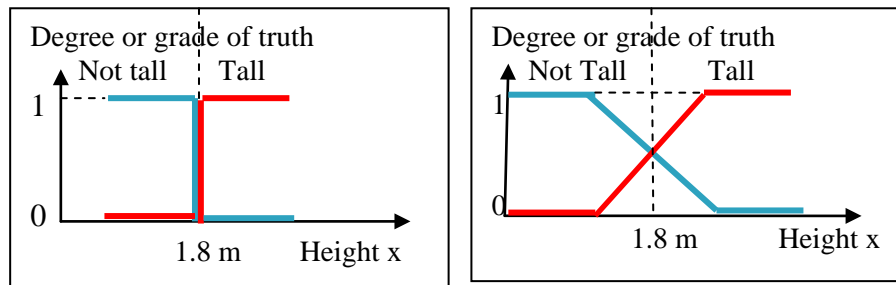


Fig. 5.2 Set Representation - Degree or grade of truth

A student of height 1.79m would belong to both tall and not tall sets with a particular degree of membership. As the height increases the membership grade within the tall set would increase whilst the membership grade within the not-tall set would decrease.

Examples of Crisp and Non-Crisp Set

Example 1: Set of prime numbers (a crisp set)

If we consider space X consisting of natural numbers ≤ 12

i.e. $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

Then, the set of prime numbers could be described as follows.

PRIME = $\{x \text{ contained in } X \mid x \text{ is a prime number}\} = \{2, 3, 5, 6, 7, 11\}$

Graphic Interpretation of Fuzzy Sets PRIME Numbers

The fuzzy set PRIME numbers, defined in the universal space

$X = \{x_i\} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ is presented as

SetOption [FuzzySet, UniversalSpace $\rightarrow \{1, 12, 1\}$]

The Set PRIME in set X is :

PRIME = FuzzySet $\{\{1, 0\}, \{2, 1\}, \{3, 1\}, \{4, 0\}, \{5, 1\}, \{6, 0\}, \{7, 1\}, \{8, 0\},$
 $\{9, 0\}, \{10, 0\}, \{11, 1\}, \{12, 0\}\}$

Therefore SetPrime is represented as

SetPrime = FuzzySet $\{\{\{1,0\},\{2,1\}, \{3,1\}, \{4,0\}, \{5,1\},\{6,0\}, \{7,1\}, \{8, 0\}, \{9, 0\}, \{10, 0\}, \{11, 1\},$
 $\{12, 0\}\}, \text{UniversalSpace} \rightarrow \{1, 12, 1\}\}$

FuzzyPlot [PRIME, AxesLabel→ {"X", "PRIME"}]

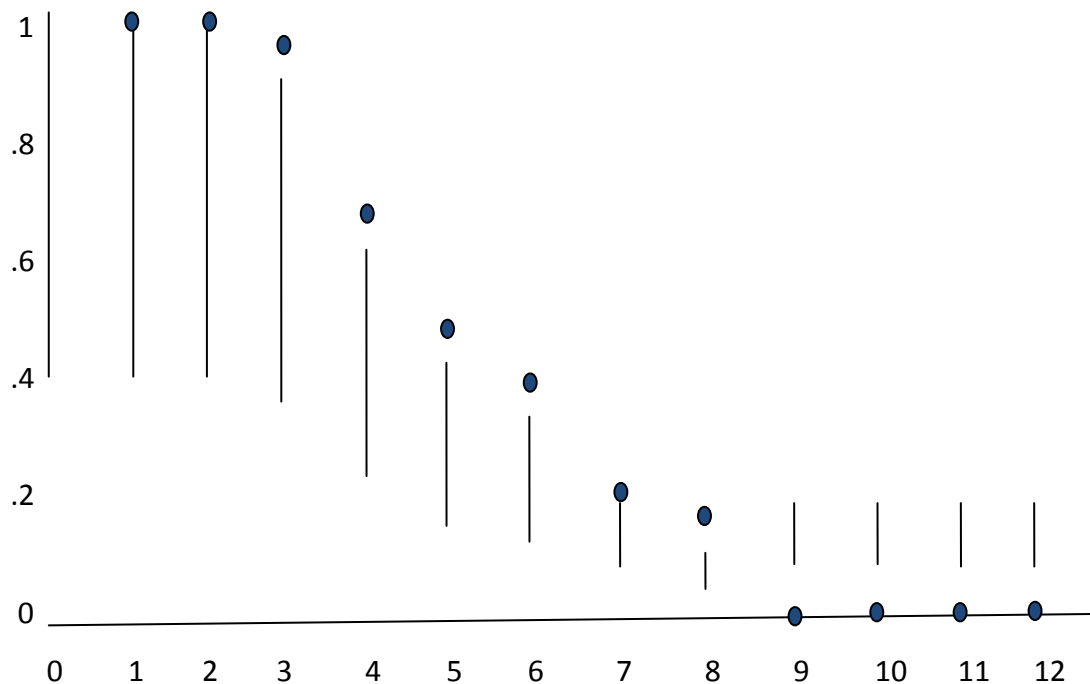


Fig. 5.3 Graphic Interpretation of Fuzzy Sets PRIME

Because web mining has fuzzy characteristics so fuzzy clustering is used as a technique to cluster data instead of conventional clustering. In the next section the use of clustering and fuzzy clustering has been discussed.

6 CLUSTERING

Clustering [8] can be considered the most important unsupervised learning problem; it deals with finding a structure in a collection of unlabeled data. A loose definition of clustering could be the process of organizing objects into groups whose members are similar in some way. A cluster is therefore a collection of objects which are similar between them and are dissimilar to the objects belonging to other clusters. Clustering algorithms can have different properties:

Hierarchical or flat: Hierarchical algorithms induce a hierarchy of clusters of decreasing generality, for flat algorithms, all clusters are the same.

Iterative: The algorithm starts with initial set of clusters and improves them by reassigning instances to clusters.

Hard and soft: Hard clustering assigns each instance to exactly one cluster. Soft clustering assigns each instance a probability of belonging to a cluster.

Disjunctive: Instances can be part of more than one cluster.

Clustering involves two forms of object cluster which are overlap and non-overlap cluster.

The cluster is said as overlap when the object can belong to other cluster and non-overlap is the object belongs to one and only cluster.

Clustering is a process which partitions a given data set into homogeneous groups based on given features such that similar objects are kept in a group whereas dissimilar objects are in different groups. It is the most important unsupervised learning problem. It deals with finding structure in a collection of unlabeled data. For better understanding please refer to Fig 1.

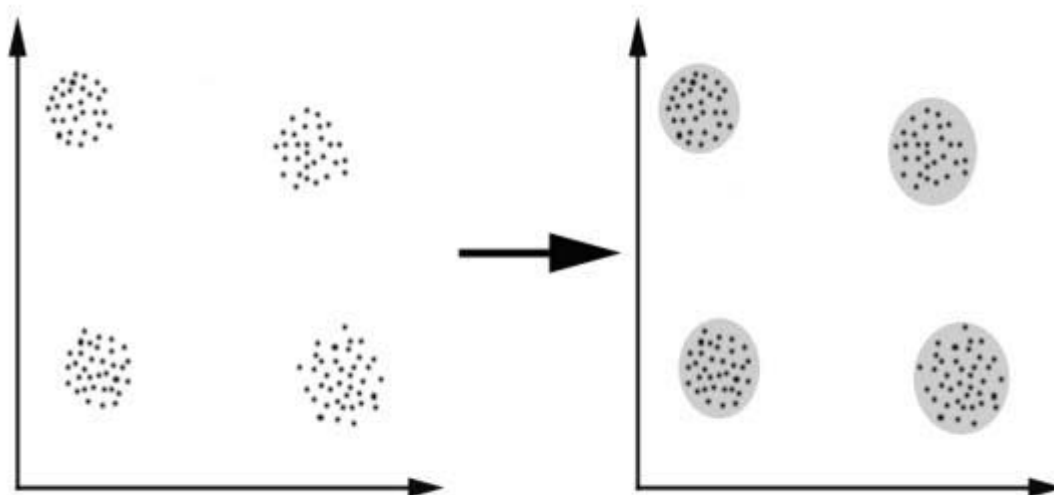


Fig. 6.1: showing four clusters formed from the set of unlabeled data

For clustering algorithm to be advantageous and beneficial some of the conditions need to be satisfied.

1) Scalability - Data must be scalable otherwise we may get the wrong result.

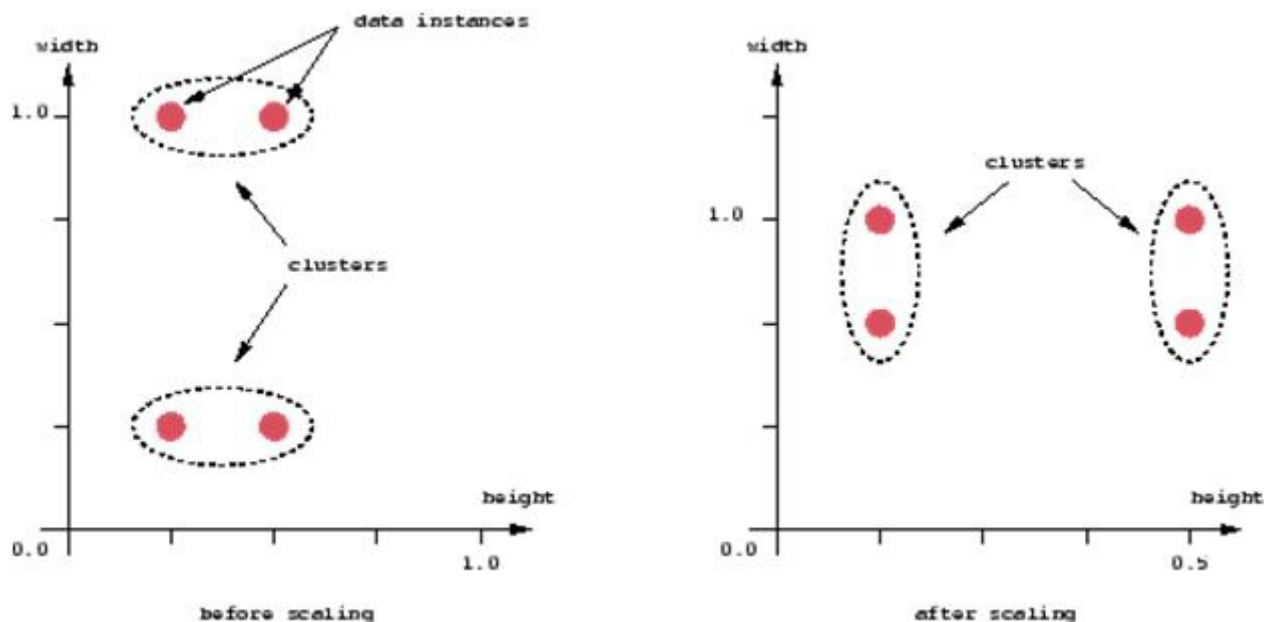


Fig 6.2: showing example where scalability may leads to wrong result

- 2) Clustering algorithm must be able to deal with different types of attributes.
- 3) Clustering algorithm must be able to find clustered data with the arbitrary shape.
- 4) Clustering algorithm must be insensitive to noise and outliers.
- 5) Interpret-ability and Usability - Result obtained must be interpretable and usable so that maximum knowledge about the input parameters can be obtained.
- 6) Clustering algorithm must be able to deal with data set of high dimensionality.

Clustering algorithms can be broadly classified into two categories:

- 1) Unsupervised linear clustering algorithms and
- 2) Unsupervised non-linear clustering algorithms

I. Unsupervised linear clustering algorithm

k-means clustering algorithm, Fuzzy c-means clustering algorithm, Hierarchical clustering algorithm, Gaussian(EM) clustering algorithm, Quality threshold clustering algorithm

II. Unsupervised non-linear clustering algorithm

MST based clustering algorithm, kernel k-means clustering algorithm, Density based clustering algorithm.

6.1 Fuzzy Clustering

In fuzzy clustering[10,18], data elements can belong to more than one cluster, and associated with each element is a set of membership levels. These indicate the strength of the association between that data element and a particular cluster. Fuzzy clustering is a



process of assigning these membership levels, and then using them to assign data elements to one or more clusters. There are few types of fuzzy clustering, such as Fuzzy c-varieties (FCV) algorithm, adaptive fuzzy clustering (AFC) algorithm, Fuzzy C-mean (FCM) algorithm, Gustafson-Kessel (GK) algorithm and Gath-Geva (GG) algorithm. Borgelt (2003) has used FCM in the high number of dimensions and the special distribution characteristics of the data. The selection of Fuzzy methods is to find fuzzy clusters of ellipsoidal shape and differing size since the data is too complex. As the studies from Borgelt, it is said that FCM succeeds in cluster the document and yields the best result in classification accuracy. In IR field, cluster analysis has been used to create groups of documents with the goal of improving the efficiency and effectiveness of retrieval, or to determine the structure of the literature of a field. The terms in a document collection can also be clustered to show their relationships. The two main types of cluster analysis methods are the non-hierarchical, which divide a data set of N items into M clusters and hierarchical clustering produces a nested data set in which pairs of items or clusters are successively linked. The non-hierarchical methods such as the single pass and reallocation methods are heuristic in nature and require less computation than the hierarchical methods. However, the hierarchical methods have usually been favored for cluster-based document retrieval. The commonly used hierarchical methods, such as single link, complete link, group average link, and Ward's method have high space and time requirements. While variables in mathematics usually take numerical values, in fuzzy logic[2,3] applications, the non-numeric linguistic variables are often used to facilitate the expression of rules and facts.[3]. A linguistic variable such as age may have a value such as young or its antonym old. However, the great utility of linguistic variables is that they can be modified via linguistic hedges applied to primary terms. The linguistic hedges can be associated with certain functions Logical analysis.

Objectives and challenges in fuzzy clustering. Create an algorithm for fuzzy clustering [10] that partitions the data set into an optimal number of clusters. This algorithm should account for variability in cluster shapes, cluster densities, and the number of data points in each of the subsets. Cluster prototypes would be generated through a process of unsupervised learning.



CONCLUSION

Web data has fuzzy characteristics, so fuzzy clustering is sometimes better suitable for Web Mining in comparison with conventional clustering. Researchers have identified the use of soft computing methods (Fuzzy logic, Neural networks, Genetic algorithms) in web mining as a tool to optimize search results. This paper discussed the part played by fuzzy logic. Finally the concept of Fuzzy Clustering was introduced as a solution for optimizing results of generalized search engines. This concept contains many aspects on which further research can be carried out.

REFERENCES

- [1] R. Kosala and H. Blockeel, "Web Mining Research: A Survey", SIGKDD Explorations ACM SIGKDD, July 2000.
- [2] M. Ganesh Introduction to FUZZY SETS & FUZZY LOGIC, PHI learning private limited
- [3] George J. Klir/Bo Yuan, Fuzzy sets and Fuzzy logic, Prentice Hall of India Private limited.
- [4] Jiawei Han, Micheline Kamber, Data mining, Morgan Kaufmann publishers
- [5] Improved FCM algorithm for Clustering on Web Usage Mining, K. Suresh
- [6] Guandong Xu, Yanchun Zhang, Web mining & social networking
- [7] Web Mining in Soft Computing Framework: Relevance, State of the Art and Future Directions, Sankar K. Pal, Fellow, IEEE, Varun Talwar, Student Member, IEEE, and Pabitra Mitra, Student Member, IEEE
- [8] Web Document Clustering, Anjali B. Raut, G. R. Bamnote
- [9] Web mining: a survey in the fuzzy framework. Dragos Arotariteia, Sushmita, Aalborg University Esbjerg, Niels Bohrs Vej 8, 6700 Esbjerg, Denmark
- [10] A new approach for Fuzzy Clustering of Web Documents, Menahem Friedman, Moti Schneider.
- [11] Real-World Fuzzy Logic Applications in Data Mining and Information Retrieval
Bernadette Bouchon-Meunier, Marcin Detyniecki, Marie-Jeanne Lesot, Christoph Marsala, and Maria Rifqi.
- [12] Web Document Clustering Using Fuzzy Equivalence Relations Anjali B. Raut, G. R. Bamnote.
- [13] The application of fuzzy logic to the construction of the ranking function of information retrieval systems n.O. Rubens.



- [14] Improved FCM algorithm for Clustering on Web Usage Mining ,K.Suresh R.Madana Mohana A.RamaMohan Reddy.
- [15] Soft Information retrieval based on Fuzzy set & Neural network, Fabio crastani, Gabriella Pasi.
- [16] Real-World Fuzzy Logic Applications in Data Mining and Information Retrieval
Bernadette Bouchon-Meunier, Marcin Detyniecki, Marie-Jeanne Lesot, Christoph Marsala, and Maria Rifqi.
- [17] Web Document Classification Based on Fuzzy Association , Choochart Haruechaiyasak, Mei-Ling Shyu , FL 33124, USA Shu-Ching Chen.
- [18] Improved FCM algorithm for Clustering on Web Usage Mining ,K.Suresh R.Madana Mohana A.RamaMohan Reddy.
- [19] Soft Information retrieval based on Fuzzy set & Neural network, Fabio crastani, Gabriella Pasi.