CHAPTER 2

FUZZY LOGIC BASED DATA CLASSIFICATION

2.1 INTRODUCTION

The rapidly growing computing power enables faster processing of huge data sets and facilitates the use of elaborate and diverse methods for data analysis and classification. Most of the data classification algorithms proposed in the literature is based on statistical and machine learning approach in which individual items are placed into groups based on quantitative information on one or more characteristics inherent in the items.

The concept of Fuzzy Logic (FL) was conceived by Lotfi Zadeh, a Professor at the University of California at Berkley in the year 1965. Zadeh, defined the FL as a way of processing data by allowing partial set membership rather than crisp set membership or non-membership. FL incorporates a simple, rule-based approach for solving a data classification problem rather than attempting to model a system mathematically.

This chapter discusses the basic concepts of Fuzzy Logic and also presents the architecture and components of a Fuzzy Logic Based Classification System (FLBCS). Further it provides a detailed summary of the literature review undertaken for the design of FLBCS.
2.2 BASIC CONCEPTS OF FUZZY LOGIC SYSTEM

Fuzzy Logic is a problem-solving methodology that provides a simple way to arrive at a definite conclusion based upon vague, ambiguous, imprecise, noisy, or missing input information. FL’s approach to data classification problems mimics how a person makes decisions in complex situation. The following subsections present the various concepts underlying Fuzzy Logic (Celikyilmaz 2009, Jang et al. 1997):

2.2.1 Fuzzy Set

Fuzzy set theory generalizes classical set theory to allow partial membership with a smooth boundary. The degree of membership in a set is expressed as a number between 0 and 1. 0 means entirely not in the set, 1 means completely in the set, and a number in between 0 and 1 means partially in the set. The membership in a fuzzy set need not be complete, i.e., member of one fuzzy set can also be the member of other fuzzy sets in the same universe.

Fuzzy sets can be analogous to the thinking of people. If a person is to be classified as friend or enemy, people will not resort to absolute classification as friend or enemy. Rather, they will classify the person somewhere between two extremes of friendship and enmity. Thus vagueness is introduced in fuzzy set by eliminating the sharp boundaries that divide members from non-members in the group and introducing a gradual transition between full membership and non-membership.

A fuzzy set $A$ in the universe of discourse $U$ can be defined as a set of ordered pairs and is given by
\[ A = \left\{ (x, \mu_A(x)) \mid x \in U \right\} \]

(2.1)

where \( \mu_A(x) \) is the degree of membership of \( x \) in \( A \) and it indicates the degree that \( x \) belongs to \( A \). The degree of membership \( \mu_A(x) \) assumes values in the range from 0 to 1, i.e., \( \mu_A(x) \in [0, 1] \).

Fuzzy sets are the tools that convert the concept of fuzzy logic into algorithms. Since fuzzy sets allow partial membership, they provide computer with such algorithms that extend binary logic and enable it to take human-like decisions. In other words, fuzzy sets can be thought of as a media through which the human thinking is transferred to a computer.

### 2.2.2 Membership Function

Membership function defines the fuzziness in a fuzzy set irrespective of the elements in the set and can be thought of as a technique to solve empirical problems on the basis of experience rather than knowledge. In general, it is represented in graphical form but there exist certain limitations in the shape of the membership function used for representation.

Membership functions can be defined in a number of different ways, based on different shapes and different number of parameters, i.e., different shapes is suitable for different applications. Certain functions are often piecewise linear functions, such as triangular or trapezoidal functions because they are simple and fast to compute.

Figure 2.1 shows a trapezoidal membership function defined by four parameters a, b, c, d.
Mathematically, a Trapezoidal membership function is defined as given below:

\[
\text{Trapezoid (x: a, b, c, d)} = \begin{cases} 
0 & x < a \\
(x - a)(b - a) & a \leq x < b \\
1 & b \leq x < c \\
(d - x)(d - c) & c \leq x \leq d \\
0 & x \geq d 
\end{cases}
\]

\[(2.2)\]

Figure 2.2 shows a Triangular membership function defined by three parameters a, b, c.
Mathematically, Triangular membership function is defined as given below:

\[
\text{Triangle } (x: a, b, c) = \begin{cases} 
0 & x < a \\
(x-a)(b-a) & a \leq x \leq b \\
(e-x)(c-b) & b \leq x \leq c \\
0 & x > c 
\end{cases}
\] (2.3)

The shape of the membership function should be the representative of the variable. Figures 2.1 and 2.2 show the definition of a single membership function, corresponding to a single linguistic value. In practice a variable is fuzzified into two or more linguistic values and it is desirable to make sure that the corresponding membership functions interact in a meaningful way.

When designing the membership functions for an input variable, labels must be determined first. The range of each input variable is divided into several regions and each region is assigned a linguistic label for describing its behavior. Membership function can overlap with one another as shown in the Figure 2.3.

![Figure 2.3 Overlap of Membership Function](image-url)
A scope must be assigned to each membership function that numerically identifies the range of input values that correspond to a label.

2.2.3 Fuzzy If-Then Rules

A fuzzy if-then rule (also known as fuzzy rule, fuzzy implications or fuzzy conditional statement) assumes the form:

If \( x \) is \( A \) then \( y \) is \( B \) \hspace{1cm} (2.4)

where \( A \) and \( B \) are linguistic values defined by fuzzy sets on the universe of \( X \) and \( Y \), respectively. The linguistic value serves to summarize the information as it is expressed in terms of fuzzy sets instead of crisp numbers. Often “\( x \) is \( A \)” is called the antecedent or premise, while “\( y \) is \( B \)” is called the consequence or conclusion.

To understand the basic difference between crisp and fuzzy rules, consider the following two simple prediction rule antecedents, one of them is crisp and the other one is a simplified version of the former,

Crisp Rule: \( \text{IF (Age} \leq \text{25) AND (sex=male) THEN …} \hspace{1cm} (2.5) \)

Fuzzy Rule: \( \text{IF (Age is young) AND (sex=male) THEN …} \hspace{1cm} (2.6) \)

Consider a data instance with the value ‘23’ for the variable Age and the value ‘male’ for the variable Sex. In the case of the crisp rule, it is understood that the data instance satisfies both rule conditions – it matches the rule to a degree of 100%.

In the case of the fuzzy rule, first the matching between the ‘Age’ value 23 and the linguistic value ‘young’ is to be computed. As shown in Figure 2.4, if ‘Age’ is fuzzified by the three membership functions, then it is
understood that the data instance satisfies the fuzzy condition (Age is young) to a degree of 0.7, and it satisfies the crisp condition (Sex=male) to a degree of 1. With the application of fuzzy AND operator, the degree to which the data instance satisfies the rule antecedent as a whole would be given by min(0.7,1)=0.7.

![Three Linguistic Values of Age](image)

**Figure 2.4 Three Linguistic Values of Age**

In general only the variables in the antecedent part of the rules are fuzzified whereas the rule consequent is crisp. This is because crisp decisions are normally required in most of the problems. In this regard, crisp rule consequents tend to be more useful and more natural in practice, than fuzzy rule consequents.

### 2.2.4 Fuzzy Inference Engine

Fuzzy Inference Engine (FIE) is a popular computing framework based on the concept of fuzzy reasoning that derives conclusion from a set of fuzzy if-then rules and known facts. An FIE can take either fuzzy inputs or crisp inputs and produces fuzzy output by a simple nonlinear mapping from its input space to output space. Defuzzification is applied (if required) to convert the fuzzy output into crisp output.
In general, any one of the five methods namely centroid of area, bisector of area, mean of maximum, smallest of maximum and largest of maximum are used for defuzzification. There are three types of FIEs namely Mamdani, Sugeno and Tsukamoto are available and are widely employed in various applications. The major difference among these three lie in the consequents of their fuzzy rules, and thus their aggregation and defuzzification procedures differ accordingly.

2.3 ARCHITECTURE OF FUZZY LOGIC BASED CLASSIFICATION SYSTEM

Fuzzy Logic provides a framework to model uncertainty, the human way of thinking, reasoning, and the perception process. A data classification system designed using FL is simply an expert system that uses a collection of fuzzy membership functions and rules to reason about data. The architecture of Fuzzy Logic Based Classification System (FLBCS) (Kuncheva, 2000, Cordon et al. 2001) is shown in Figure 2.5. A typical FLBCS consists of three conceptual components:

- Rule Base, which contains a collection of fuzzy if-then rules
- Data Base, which defines the membership function used in the rules.
- Fuzzy Inference Engine, which derives the output class label using qualitative reasoning.

Rule Base consists of interpretable if-then rules with fuzzy antecedents and class labels in the consequent part as given below,

\[ R: \text{If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } A_2 \text{ and } \ldots x_n \text{ is } A_n \text{ then } y \text{ is class } C_1 \]  \hspace{1cm} (2.7)
The antecedents (if-parts) of the rules partition the input space into a number of fuzzy regions by fuzzy sets, while the consequents (then-parts) describes the class label in these regions. A collection of such rules (Rule Base) and membership function for each input variable (Data Base) are used as a Knowledge Base by the fuzzy classifier.

**Figure 2.5 Architecture of a Fuzzy Logic Based Classification System**

Among the three inference methods mentioned in section 2.2.4, in this research work, Mamdani Inference System (MIS) with product t-norm and max t-conorm is used since they build the classifier model with good interpretability using linguistics. (Ishibuchi *et al.* 2004).

In MIS, the set of input variable is matched against the if part of each if-then rule and the response of each rule is obtained through fuzzy implication operation. The response of each rule is weighted according to the extent to which each rule fires. The responses of all the fuzzy rules for a particular output class are combined to obtain the confidence with which the input is classified to the corresponding output class.
Figure 2.6 illustrates how a two rule MIS derives the overall output $z$ when subjected to two crisp inputs $x$ and $y$.

The operators used by the MIS do not use the strict definition of the compositional rule of inference and a suitable function is assigned for each of the following operators to completely specify the operation of MIS.

- **AND operator** (usually T-norm) for calculating the firing strength of a rule with AND’ed antecedents.
- **OR operator** (usually T-conorm) for calculating the firing strength of a rule with OR’ed antecedents.
- **Implication operator** (usually T-norm) for calculating qualified consequent MFs based on given firing strength.
- **Aggregate operator** (usually T-conorm) for aggregating qualified MFs to generate an overall output MF.
Since the class label associated with the fired rule is matched against the class label of the data instance for classification, the defuzzification operator is not required.

## 2.4 EXISTING APPROACHES FOR DESIGNING THE FLBCS

In the growing scenario, Data Mining industry successfully employs FLBCS for discovering comprehensible and interesting knowledge. The key to the success of the FLBCS is its ability to incorporate human expert knowledge and helps the user to make intelligent decision. An FLBCS consists of if-then rules whose antecedent defines a local fuzzy region using membership function and consequent describes the behavior within the region. Thus, formation of if-then rules and membership functions are the two important major issues for designing an FLBCS.

In general the rules and membership functions are formed from the experience of the human experts. With an increasing number of variables, the possible number of rules increases exponentially, which makes it difficult for experts to define a complete rule set for good system performance. Data-driven approaches (Wang et al. 1992) have been proposed for developing the fuzzy rule based system from numerical data without domain experts.

Abe et al. (1995) proposed a rule generation method in which each fuzzy if-then rule was represented by a hyper box in multidimensional pattern spaces. But they are very weak in self learning and determining the required number of fuzzy if-then rules. (Ishibuchi et al. 1995) proposed a heuristic method for generating fuzzy if-then rules using grid-type fuzzy partition in which a priori knowledge on linguistic values is required for specifying the membership function which fails to handle high dimensional problems with many input variables due to the dimensionality factors.
During the late 1990s attempts have been made to incorporate optimization techniques, which make the FLBCS as a self-learning system. This optimization based learning approach is responsible for generating rules and tuning membership function that are incorporated into an FLBCS. Genetic Fuzzy System (GFS) (Russo, 2000) is one approach in which a fuzzy system is augmented by a learning process based on Genetic Algorithm (GA).

Specifically, there exist two kinds of Genetic Fuzzy learning strategies: the Pittsburg and the Michigan approaches (Setnes et al. 2000). The former obtains a population of rule bases where the best suited one is selected as the final rule base and the later encodes and evolves each rule as an individual, and a single rule base is evolved.

Cordon et al. (2004) gives a brief review of most of the approaches found in the literature on GFS. Accordingly, the GFS proposed in the literature falls into four categories:

- Learning rules with fixed fuzzy membership functions (Casillas et al. 2005),
- Learning membership functions with fixed rules (Alcala et al. 2007),
- Learning rules and membership functions in stages (Zhou et al. 2007)
- Learning rules and membership functions simultaneously (GaneshKumar et al. 2009).

Majority of GA based approaches reported in the literature improve the accuracy of the FLBCS at the cost of their interpretability. Recently the existence of interpretability-accuracy tradeoff (Luengo et al. 2010, Gacto et al. 2011) has been realized in the design of GFS. A decision-tree-based
initialization of FLBCS proposed by Abonyi et al. (2003) iteratively use the similarity driven rule reduction and GA for obtaining compact and interpretable classifier model.

Evolutionary scatter partition of feature space (Ho et al. 2004), Variable input spread inference training algorithm (Chang et al. 2004), Modified Gath Geva and C4.5 (Pulkkinen et al. 2007) and multi-objective approaches (Ishibuchi et al. 2007, Alcala et al. 2007) are proposed to design the compact and interpretable FLBCS using GA with comparable accuracy. Herrera (2008) provided a detailed taxonomy, current research trends and prospects of genetic fuzzy system. Even though a number of works have been reported in the literature in designing a fuzzy system for data classification, more effort is still required to increase the efficiency of the learning of FLBCS and their interpretability issues.

2.5 SUMMARY

In this chapter, a brief review on the concepts of Fuzzy Logic is discussed. Further, the architecture of a data classification system designed using Fuzzy Logic is presented. Also, a detailed summary of the literature review undertaken for this research is discussed.