DOKUZ EYLUL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

# DEVELOPMENT AND IMPLEMENTATIONS OF FUZZY DECISION TREE ALGORITHMS

**Suzan KANTARCI SAVAŞ**

**February, 2017**
**İZMİR**

# DEVELOPMENT AND IMPLEMENTATIONS OF FUZZY DECISION TREE ALGORITHMS

**A Thesis Submitted to the**
**Graduate School of Natural and Applied Sciences of Dokuz Eylül University**
**In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Department of Statistics, Statistics Program**

**by**
**Suzan KANTARCI SAVAŞ**

**February, 2017**
**İZMİR**

## Ph.D. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"DEVELOPMENT AND IMPLEMENTATION OF FUZZY DECISION TREE ALGORITHMS"** completed by **SUZAN KANTARCI SAVAŞ,** under supervision of **PROF. DR. EFENDİ NASİBOĞLU** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. Efendi NASİBOĞLU

Supervisor

Prof. Dr. Esin FİRUZAN

Thesis Committe Member

Asst. Prof. Dr. Derya BİRANT

Thesis Committe Member

Prof. Dr. Şahin EMRAH

Examining Committe Member

Assoc. Prof. Dr. Ali MERT

Examining Committe Member

Prof. Dr. Emine İlknur GÖCEN

Director

Graduate School of Natural and Applied Sciences

## ACKNOWLEDGEMENTS

I would like to say my deepest gratitude to my supervisor, Prof. Dr. Efendi NASİBOĞLU for his patience, motivation and continuous support during my Ph.D study. His encouragement and supervision lead me not just to start my academic career in a good way, also lead me to find my own fuzzy solution whenever I come across with a complex problem in my life. It has been a great chance for me to work with him. In other respects, I would like to thank the rest of my thesis committee: Prof. Dr. Esin FIRUZAN, and Assit. Prof. Dr. Derya BIRANT for her helpful comments, but also for their confidence in me and for providing during my thesis process. Also, I would like to thank Prof. Dr. Ronald R.YAGER for your valuable inspiration, to Prof. Dr. Fazıl GÜLER and Assist. Prof. Dr. Esin Cumhur YALÇIN for their support and belief in me. This work has been supported by Dokuz Eylul University under Scientific Research Project No. 2011.KB.FEN.035. I would like to thank Mrs. Ummuhan Tibet and Dr. Aytac Gumuskesen for allowing me to use their data set, Dr. Gözde ULUTAGAY, FuDeS Group Members Dr. Alper VAHAPLAR and Dr. Ayşe Övgü KINAY for sharing all their experience and time with me generously.

I would like to thank my dear parents, Ayfer & Halil Cevdet KANTARCI, and my brother, Erden KANTARCI who have always been with me whenever I am in need, for their encouragement and trust in me. I would like to thank my all family members, to thank my cousin Elif Deniz YÜNLÜ for her valuable comments and to thank dear mother and father in law, Sibel & Hüseyin SAVAŞ for their encouragement and trust in me. I also owe thanks to my grandmothers, Emine ERDEN and Rükiye Suzan KANTARCI, for their love, care, and prayers through the years and for their confidence in me.

I would like to express my great appreciation to my invaluable and immortal beloved husband, Kerem SAVAŞ, for his infinite patience, support, and understanding. His love has always given me unlimited strength.

Suzan KANTARCI SAVAŞ

**DEVELOPMENT AND IMPLEMENTATIONS OF FUZZY DECISION TREE ALGORITHMS**

## ABSTRACT

In this work, fundamentally fuzzy ID3 algorithm and the effects of T-operators on its reasoning procedure on numeric data is investigated and two novel approaches working on linguistic data have been proposed. Firstly, the information about fuzzy ID3 approach is presented. The usage of T-operators is given in details. The implementation is applied on geographic classification of virgin olive oil. Then, the first proposed approach on linguistic data is the FID3-L-WABL (Fuzzy ID3 Algorithm Based on Linguistic Data by Using WABL Defuzzification Method) which is a novel version of the known Fuzzy Interactive Dichotomizer 3 (Fuzzy ID3) classification algorithm working on linguistic data. This approach is performed on L-R (Left-Right) Fuzzy Number. Each fuzzy number is defuzzified by using WABL (Weighted Averaging Based on Levels). Then, Fuzzy c-means algorithm is adapted to obtain membership values of each fuzzy term for each fuzzy variable. Consequently, Fuzzy ID3 algorithm is applied. The second proposed approach is the FID3-LR (Fuzzy ID3 Algorithm for L-R Fuzzy Data) which is a mixture of FkM-F (Fuzzy k-means Clustering Model for Fuzzy data) clustering algorithm working on L-R fuzzy data and Fuzzy Interactive Dichotomizer 3 (Fuzzy ID3) classification algorithms.

Fuzzy c-means algorithm was performed in MATLAB 2014a. The codes for the experiments, FuzzyID3 by using T-operators, FID3-L-WABL, and FID3-LR, have been developed in the MS Visual Studio C# IDE for the experimental study (Intel i7, 2.4 GHz, 4 Gb RAM). They have been designed as an integrated software system called as Fuzzy Artemis. In addition, OliveDeSoft is designed for current and future studies in order to analyze the olive oil quality and geographic characterization.

**Keywords:** Fuzzy logic, fuzzy decision tree, classification, defuzzification, linguistic data, geographic identification, olive oil.

**BULANIK KARAR AĞACI ALGORİTMALARININ GELİŞİMİ VE UYGULAMALARI**

**ÖZ**

Bu çalışmada, esasen fuzzy ID3 algoritması ve nümerik veri üzerinde T-operatörlerinin çıkarsama prosedürüne etkisi araştırılır ve sözel veri üzerinde çalışan iki yeni yaklaşım ortaya konulmaktadır. Öncelikle, fuzzy ID3 yaklaşımı sunulmaktadır. T-operatörlerinin kullanımı detaylıca incelenmektedir. Uygulaması, natürel zeytinyağının coğrafik sınıflandırılması üzerinde gerçekleştilmektedir. Sonra, sözel verilerde çalışan ilk sunulan yaklaşım Fuzzy Interactive Dichotomizer 3 (Bulanık ID3) sınıflandırma algoritmasının yeni bir versiyonu olan FID3-L-WABL (WABL durulaştırma methodu kullanılarak sözel veriye dayalı Bulanık ID3 Algoritması)'dır. Bu yaklaşım, L-R (Sol-Sağ) Bulanık Sayısı üzerinde çalıştılmaktadır. Her bulanık sayı WABL (Weighted Averaging Based on Levels) ile durulaştırılmaktadır. Sonra, Bulanık c-Ortalamalar algoritması, her bulanık değişkenin bulanık terimimin üyelik derecesini elde etmek için adapte edilmektedir. Son olarak, Bulanık ID3 uygulanır. İkinci sunulan yaklaşım, L-R bulanık veri üzerinde çalışan FkM-F (Bulanık Veri için Bulanık k-ortalamalar Kümeleme Modeli) kümeleme algoritmasının, Fuzzy Interactive Dichotomizer 3 (Bulanık ID3) algortimasıyla karışımından oluşan FID3-LR (Bulanık Veri için Bulanık ID3) algoritmasıdır.

Bulanık c-ortalamalar algoritması MATLAB 2014a'da yazılmıştır. T-operatörleri kullanılarak çalışan FuzzyID3, FID3-L-WABL, and FID3-LR, deneysel çalışmaları MS Visual Studio C# ortamı (Intel i7, 2.4 GHz, 4 Gb RAM) kullanılarak geliştirilmiştir. Deneyler, Fuzzy Artemis isimli entegre edilmiş bir yazılım geliştirilerek tasarlanmıştır. Ayrıca, OliveDeSoft, zeytinyağının kalitesi ve coğrafik karakterizasyon analizi yapmak için tasarlanmıştır.

**Anahtar Kelimeler:** Bulanık mantık, bulanık karar ağacı, sınıflandırma, durulaştırma, sözel veri, cografi karakterizasyon, zeytinyağı.

**CONTENTS**

## LIST OF FIGURES

**Page**

**LIST OF TABLES**

**Page**

# CHAPTER ONE
# INTRODUCTION

Classification has a crucial manner in machine learning and data mining. Humans use words in order to communicate. Words include more information than numbers in order to make decision and classification. Nowadays, fuzzy logic is preferred to handle the imprecise information for computing with words because of its flexibility.

The amount of data in the world goes on increasing each day in data repositories. As a result, databases reserve a huge amount of information that has not been discovered yet. In recent years, an effective tool which is called data mining has emerged to discover patterns and trends in these data repositories. According to Larose (2005) data mining is a widely used process in order to discover meaningful relations, patterns and trends by using statistical and mathematical techniques in large data repositories.

The technical basis of data mining is provided by machine learning. Machine learning aims to identify patterns and pick out some knowledgeable information based on the data inside data repositories. The algorithms of the machine learning can be used to illustrate relations among observed variables and have deeper information about data.

Decision trees are widely used machine learning tools in data mining among the most popular classification structures (Janikow, 1996, 1998; Lee et al., 1999; Kantarcı, Turanoğlu, & Ulutagay, 2013). Originally, they have been studied in the fields of decision theory and statistics. In recent years, it has also been implemented in pattern recognition. Quinlan (1990) argues that decision trees provide a powerful formalism for representing comprehensible and accurate classifiers.

In the literature, there are distinctive methods proposed to generate the decision trees. ID3 (Quinlan, 1986), CART (Breiman et al., 1984) and C 4.5 (Quinlan, 1993) are the most important decision tree learning algorithms. Trees produced by these algorithms are very sensitive to small changes in data.

However, Abu-halaweh & Harrison (2010) argue that ID3 algorithm is not efficient in handling data uncertainties due to measurement error and/or noise. It was also argued that these algorithms work well in symbolic domains. They designate symbolic decisions to new samples.

In recent years, neural networks have become popular for classification (Sarle, 1994; Aroca-Santos, 2016; Binetti et al., 2017). But when the users want to understand or justify the rules, neural networks do not fit for the purpose. Fuzzy representation deals with problems of uncertainty, noise, and inexact data (Zadeh, 1965). Some scholars have attempted to combine fuzzy sets and decision trees. They advise fuzzy decision trees. Fuzzy decision trees are also useful to evaluate the rules used for the classification. The usage of linguistic labels as fuzzy terms gives power to the knowledge systems.

First, Chang & Pavlidis developed a kind of fuzzy decision tree algorithm in 1977. This fuzzy decision tree approach was based on binary search trees. In the literature, several versions of fuzzy decision tree algorithms are proposed by Umano et al. (1994), Janikow (1998), Chiang & Hsu (2002), Liu & Pedrycz (2007), Sanz et al. (2012) as fuzzy ID3 algorithm. Jang et al. (1997) also applied a fuzzy CART approach to estimate the structure of a fuzzy inference system, and Tokumaru & Muranaka (2009) used Fuzzy C 4.5 algorithm for product impression.

In our study, firstly, it is aimed to observe the behaviour of various types of T-operators on the reasoning process of fuzzy ID3 structure. Secondly, two novel fuzzy ID3 algorithms called Fuzzy ID3 algorithm based on linguistic data by using WABL defuzzification method (Fuzzy ID3-L-WABL) and Fuzzy ID3 algorithm for L-R fuzzy data (Fuzzy ID3-LR) are proposed for the classification on linguistic data set. In both two approaches, linguistic variables are defined by using triangular fuzzy numbers given as L-R fuzzy numbers. Fuzzy ID3-L-WABL uses weighted level-based averaging method to make defuzzificaiton. Then, fuzzy c-means algorithm is performed in order to handle the membership degrees for each variable given in the data sets. Then, Fuzzy ID3-LR performs directly on L-R fuzzy data. It makes fuzzification via fuzzy k-means clustering model for fuzzy data (FkM-F) algorithm which works on L-R fuzzy data. After fuzzification process, both two approaches use Fuzzy ID3 algorithm for the induction of fuzzy decision tree. And, finally the fuzzy reasoning performs with different T-operators.

The rest of the thesis is organized as follows. In Chapter II, we briefly introduce fuzzification, tree induction and fuzzy reasoning procedure on fuzzy decision trees. In Chapter III, the information about Fuzzy ID3 algorithm based on Fuzzy c-means for numeric data is given. In Chapter IV, a brief information is given about linguistic data, Weighted Averaging Based on Levels (WABL), and then Fuzzy ID3 algorithm based on linguistic data by using WABL defuzzification method (Fuzzy ID3-L-WABL) is proposed. In Chapter V, weighted dissimilarity measure, fuzzy k means clustering model for fuzzy data (FkM-F) is informed, and Fuzzy ID3 algorithm for L-R fuzzy data (Fuzzy ID3-LR) is presented. In Chapter VI, the experimental framework is reported. In Chapter VII, Conclusion is stated as a final chapter.

# CHAPTER TWO
# FUZZY DECISION TREES AND ITS REASONING PROCEDURE

## 2.1 Introduction

Decision trees which are combined with fuzzy approach are called fuzzy decision trees. While fuzzy decision tree is being constructed, some scholars prefer to initially fuzzify variables and splitting criteria, thus obtain fuzzy rules and implement these rules in inference procedure (pre-fuzzification). On the other hand, other scholars prefer using classical variables and splitting criteria, and then fuzzifying the obtained classical rules (post-fuzzification). First fuzzy decision trees were introduced by Chang & Pavlidis in 1977. The binary search method and fuzzy approach was combined in their paper. In 1994, Umano et al. proposed a novel approach to set up a fuzzy decision tree from data defined using fuzzy sets by way of experts. Their algorithm was called fuzzy ID3 algorithm. Fuzzy decision tree handled with fuzzy ID3 algorithm consists of nodes for testing variables, thresholds used for branching via test values of fuzzy sets given via an expert and leaves for determining the class with certainties (pre-fuzzification).

Yuan & Shaw (1995) introduced an induction-learning algorithm for fuzzy decision trees. They were interested in incorporating subjective uncertainties into knowledge induction procedure for classification. In their study, the aim is to reduce the classification ambiguity in order to generate fuzzy decision tree (pre-fuzzification).

Hsu et al. (1995) used Classical ID3 algorithm to generate the rules for mobile robot control. A post-fuzzification is applied to the generated rules. Then, a neural network architecture represented the fuzzy rules. While performing an on-line training, the membership function of each linguistic variable is assigned via the gradient descent approach.

In 1998, Janikow presented the fuzzy decision tree approach with various inference procedures, which were depended on conflict resolution in rule-based systems and powerful approximate reasoning methods. Proposed fuzzy decision tree uses a version of Fuzzy ID3 algorithm (pre-fuzzification). This paper has become an important reference for future studies.

Chiang & Hsu (2002) developed a new fuzzy classification tree for data analysis. This algorithm integrates the fuzzy classifiers with decision trees that can work well in classifying data with noise. When the accuracy of this algorithm is compared with the accuracy of three algorithms, which are proposed by Yuan & Shaw (1995), Hsu et al. (1995), and Janikow (1998), it is seen that the fuzzy classification algorithm has a better accuracy rate.

Liu & Pedrycz (2007) proposed a new algorithmic framework for building fuzzy sets (membership functions), their logic operators and forming the design process of fuzzy decision trees. They compared their findings with the outcomes produced by the fuzzy decision trees presented by Janikow (1998).

Sap & Khokhar (2004) constructed a new fuzzy decision tree by using weighted fuzzy production rules. In this approach, each proposition is assigned a weight parameter in the antecedent of a fuzzy production rule and a certainty factor is assigned to each rule. In this paper, the implementation is applied to stock market databases. Certainty factors have been calculated by using important variables such as effect of other companies, effect of stock exchanges etc. in dynamic stock market. And, this approach was used to predict stock share indices.

Chang et al. (2011) established a case based fuzzy decision tree model to predict the behaviour of stock prices movements in financial time series for trend discovery. In this paper, the fuzzy decision tree is generated from the stock database and then converted to fuzzy rules; and these rules are used in decision making of stock price's movement.

Sanz et al. (2012) aimed to improve the performance of fuzzy decision trees (FDT) by using IVFS and Genetic Algorithms. They presented a novel methodology called "Ignorance functions based Interval-Valued Fuzzy Decision Tree with Genetic Tuning (IIVFDT)". The induction of the base FDT using the fuzzy ID3 algorithm proposed in Yuan & Shaw (1995). A new modeling of the linguistic labels of the classifier is proposed by means of IVFSs. With this aim, they defined a novel construction method of IVFSs starting from the fuzzy sets by learning algorithm and using weak ignorance functions to measure the degree of ignorance when assigning punctual values as membership degrees (Sanz et al., 2011). The extension of the Fuzzy Reasoning Method (FRM) accomplished with the full power of IVFSs in the inference process. In each step of the FRM, the computation is made by using intervals and the ignorance degree is taken into account from the beginning to the end of the process.

In this chapter, before explaining proposed fuzzy decision tree approaches, what fuzzification is; what fuzzy logic is; how the fuzzification is done by using fuzzy numbers and fuzzy c-means (FCM) is defined, how the tree induction is performed, and how the fuzzy reasoning procedure on fuzzy decision tree is done will be covered.

## 2.2 Fuzzification

Fuzzification process can be defined as making a crisp quantity fuzzy. Membership functions can be used for transforming a crisp variable into a fuzzy variable. Also, if fuzzy c-means algorithm is used, membership values are achieved. These values show the membership degrees for each cluster defined as a fuzzy variable (Ross, 2010).

### *2.2.1 Fuzzy Logic and Fuzzy Sets*

In real life, there are many more unrealistic situations contained vagueness and ambiguity. In order to deal with the problems appeared because of these contexts, Zadeh (1965) suggested fuzzy set theory. In consequence of this set theory, fuzzy logic emerged. While the classical logic uses binary sets, the fuzzy logic uses fuzzy sets. In classical set theory, an element in the universe either matches with a set or does not.

Let U be a space of objects and x is generic element of U. A classical set A, $A \subseteq U$, is defined as a collection of elements or objects $x \in U$, such as each x can either belong or not belong to the set A. By using a characteristic function for each element x in U, a classical set A can be represented by a set of ordered pairs $(x, 0)$ or $(x, 1)$. It means $x \notin A$ or $x \in A$, respectively.

**Definition 2.1. Fuzzy sets and membership functions** If x is a collection of objects denoted generically by x, then a fuzzy set A in U is defined as a set of ordered pairs (Zadeh, 1965):

$$A = \{x, \mu_A(x) | x \epsilon U\} \tag{2.1}$$

where $\mu_A(x)$ is called membership function for the fuzzy set A. The membership function maps each element of U to a membership grade (or membership value) between 0 and 1. X referred to as the universe of discourse, and this universe can be defined with discrete (ordered or non-ordered) or continuous space.

**Definition 2.2. Normality** A fuzzy set A is normal if its core is nonempty. There is always a point $x \in U$ such that $\mu_A(x) = 1$ (Jang et al., 1997).

**Definition 2.3. Convexity** A fuzzy set A is convex if and only if for any $x_1, x_2 \in U$ and any $\lambda \in [0,1]$ (Jang et al, 1997):

$$\mu_A(\lambda x_1 + (1 - \lambda) x_2) \geq min\{\mu_A(x_1), \mu_B(x_2)\} \tag{2.2}$$

**Definition 2.4. Fuzzy numbers** A fuzzy number A is a fuzzy set in the real line (R) that satisfies the conditions for normality and convexity.

A fuzzy set is defined by using a membership function. In literature, there are different kinds of membership functions such as triangular, trapezoidal, etc. A mathematical formula is described in order to express each membership function (Jang et al., 1997).

**Definition 2.5. A parametric triangular fuzzy number** A fuzzy number with membership function in the form

$$\mu_A(x) = \begin{cases} \left(\frac{x-a}{b-a}\right)^{1/s}, x \in [a,b), \\ \left(\frac{c-x}{c-b}\right)^{1/s}, x \in [b,c), \\ 0 \qquad\quad , \quad otherwise \end{cases} \tag{2.3}$$

where $s > 0$ is a parameter (Figure 2.1), will be termed a parametric triangular fuzzy number $A = (a, b, c)$ (Nasibov and Mert, 2007).

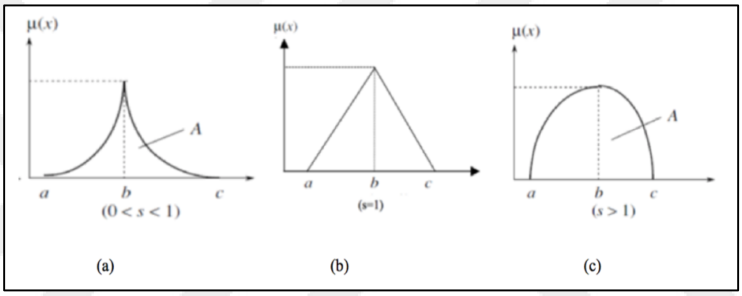



Figure 2.1 Forms of parametric triangular fuzzy numbers (a) 0<s<1 (b) s=1 (c) s>1.

**Definition 2.5. Triangular membership function** A triangular membership function is presented by three parameters $\{a, b, c\}$ as follows (Jang et al.(1997)):

$$triangle(x; a, b, c) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x < b \\ \frac{c-x}{c-b}, & b \leq x < c \\ 0, & c \leq x \end{cases} \tag{2.4}$$

These parameters $\{a, b, c\}$ $(with\ a < b < c)$ determine the x coordinates of the three corners of the underlying triangular membership function (Figure 2.2).

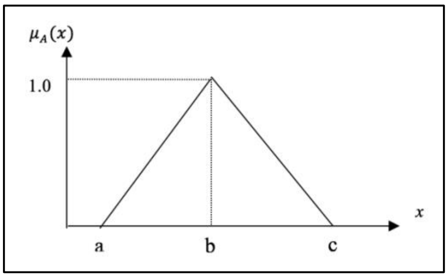


Figure 2.2 Triangular Membership function characterized by three parameters $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

### *2.2.2 Fuzzy c-Means(FCM)*

Fuzzy c-means (FCM) is a path for assigning the membership degrees for fuzzy terms in each fuzzy variable. This algorithm proposed in Dunn (1973) and it became better in Bezdek (1981). The main aim of it is to reach a fuzzy $c$ partition matrix $U$. An objective function $J_m$ is defined in order to minimize. It is given as follows for fuzzy partition (Eq. 2.5):

$$J_m(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_{ik})^m \, (d_{ik})^2 \tag{2.5}$$

where $\mu_{ik}$ is defined as the membership degree of the $k^{th}$ data point in the $i^{th}$ class. p is the dimensionality of the data space. The parameter $m > 1$ reflects sharpness of the fuzzification process. In Eq.2.6, $d_{ik}$ shows any distance measure (usually the Euclidean distance) between $k^{th}$data point and $i^{th}$ cluster center in p dimensional space. The distance could be defined as:

$$d_{ik} = d(x_k, v_i) = \left[\sum_{j=1}^{p} \left(x_{kj} - v_{ij}\right)^2\right]^{1/2}, \quad k = 1, \dots, n \quad i = 1, \dots c \tag{2.6}$$

Here, $v_i$ shows the $i^{th}$ cluster center. The following, Eq.2.7. is used for the calculation of each clusters' center:

$$v_{ij} = \frac{\sum_{k=1}^{n} \mu_{ik}^{m} x_{kj}}{\sum_{k=1}^{n} \mu_{ik}^{m}}, i = 1, \dots, c, j = 1, \dots, p \ . \tag{2.7}$$

Membership degrees are computed in accordance with the Eq.2.8.

$$\mu_{ik} = \frac{1}{\sum_{z=1}^{c} \left( \frac{\|x_k - v_i\|}{\|x_k - v_z\|} \right)^{\frac{2}{m-1}}}, i = 1, \dots, c; k = 1, \dots, n. \tag{2.8}$$

The cluster number of each variable identifies the number of fuzzy linguistic terms for each fuzzy variable.

### *2.2.3 Linguistic Variable*

Linguistic variables contain words or sentences instead of numbers as used in natural or artificial languages (Zadeh; 1975). These variables are more useful for the individuals to tell their knowledge based on various statements. Generally, humans think and need to explain their thoughts by using words and sentences in daily life. Yet, it is difficult to give the information directly. Regardless, using linguistic variable is an impressive way in order to model the human thought.

Linguistic variables also use fuzzy variables as its values. Fuzzy variable is identified by a triple (X, U, R(X)) in which X is the variables' name, U is a universe of discourse, and R(X) is a fuzzy subset of U which represents a fuzzy restriction imposed by X (Nasibov & Mert, 2007; Ross, 2010).

#### *2.2.3.1 Selecting the Linguistic Term Set*

The selection of the linguistic term set is an important issue. The aim is to introduce the minimum number of words to the user. The user needs enough argument in order to explain his/her statement efficiently. Hence, the number of the linguistic terms should be as few as achievable. It should also be as many as possible in order to evaluate the statements in different levels. The differentiation of levels in fuzzy variable can be called as “Granularity of Fuzziness”.

Miller (1955) proposed that humans generally organize the information designated with odd numbers in order to keep in their memory. It is observed that the terms are symmetrically distributed around the mean term which is approximately 0.5. Hence, generally, 7, 9, 11, or not more than 13 are preferred as element numbers in linguistic term sets (Nasibov & Kantarcı, 2010).

In addition, the numeric variables can be transformed into linguistic variables defined as fuzzy variable. A fuzzy clustering algorithm such as fuzzy c-means is used in order to assign a membership degree for each numeric value. Each fuzzy term of a fuzzy variable is defined by using linguistic term. The cluster number is identified by using experts’ opinion, by evaluating the variables’ distributions or by using a partition coefficient index in order to reach the optimal cluster number (Bezdek, 1974a; Bezdek, 1974b; Dunn, 1974).

#### *2.2.3.2 Generating Linguistic Descriptors.*

After the determination of element number of the linguistic term set, it is necessary to generate the linguistic descriptors. There are two kinds of approaches, called as Context-Free Grammar Approach and Ordered Structure of Linguistic Terms Approach, respectively. These approaches define the linguistic descriptors.

*2.2.3.3.1 Context-Free Grammar Approach.* Context-free grammar is used in order to define the linguistic term set in this approach. $G$ is defined as the grammar that accomplishes the sentences. Grammar consists of a four-order notation presented as $(V_N, V_T, I, P)$. $V_N$ proposes non-terminal symbols set, $V_T$ proposes terminal symbols set, I propose the initial symbol, and $P$, the generation rules. The expanded Backus Naur Form (Bordogna & Passi, 1993) may be performed for $P$ generation rule.

The basic terms are identified as {many, medium, few…}, constraints as {none, a lot, quite, ...}, relations are given as {fewer, lower, …} and links defined as {and, but, or, …}. Firstly, an initial term I is selected. Then, a linguistic terms set can be generated as S= {low, lower, not low, lower or medium, …} using P. Miller (1955) observed that the language used should be clearly understandable. Therefore, the selection of grammar terms composes the shape of the linguistic term set (Kantarcı, 2010).

*2.2.3.3.2 Approach Depending on the Ordered Structure of Linguistic Terms.*

Ordered structure of linguistic terms range on an indicator chart. It is defined by using an example; assume $a < b$ then $s_a < s_b$ (Kantarcı, 2010).

$S = \{s_0{:}\, none, s_1{:}\, very\ few, s_2{:}\, few, s_3{:}\, average, s_4{:}\, high, s_5{:}\, very\ high, s_6{:}\, perfect\}$

When this approach will be used, it is necessary that the term set provides the following characteristics:

There is a negation operator, i.e.,$Neg(s_i) = s_j, j = T - i$.

($\mathrm{T} + 1 =$ The number of the elements of the term set).

Maximization Operator $= \mathrm{Max}(s_i, s_j) = s_i,\ s_i \geq s_j$.

Minimization Operator$=\mathrm{Min}(s_i,s_j)=s_i,\ s_i \leq s_j$.

#### *2.2.3.3 The Semantic Depending on Membership Functions and a Semantic Rule as Linguistic Term Set*

The number of elements of the linguistic term set and its descriptors are identified by using the methods defined in former subchapters. Then, the meanings of the linguistic term set should be assigned. In the literature, there are three main issues that assign the definition of the linguistic term set. These are defined as “semantic depending on membership functions and a semantic rule”, “semantic depending on the ordered structure of the linguistic term set”, and “mixed semantic”, respectively. In this study, “semantic depending on membership functions and a semantic rule” is used. The definition is given as follows:

Propositions as “Suzan is tall”, “The olive tree is extremely small” are named as fuzzy propositions. These statements are not reflected certain situations. Each fuzzy proposition generates fuzzy terms. And each of the fuzzy term is modelled by using a “fuzzy set”. This set is characterized by mathematically designating a value from the real numbers in the range of [0,1]. This value shows that each individual’s membership degree belonging to the fuzzy set defined as fuzzy term.

The semantic issue follows the two steps given below:

- The primary fuzzy sets combined to the primary linguistic terms.
- The semantic rule M, to construct the fuzzy sets of the non-primary linguistic terms from primary fuzzy sets.

The representation of the primary fuzzy terms counts on parameters expressed by humans. Yet, it is so hard to explain their behaviour and preferences within the similar parameters. Same primary terms may not have same representations. Each researcher can prefer different membership function for linguistic evaluations (Delgado et al., 1992; Bordogna & Passi, 1993; Kantarcı, 2010).

## 2.3 Tree Induction

Decision tree method is widely used in data mining, machine learning, expert systems, and multivariate analysis. Decision tree algorithms aim to part the input space of data set into mutually exclusive regions while each input has its class label. The structure of a decision tree consists of internal and external nodes connected via branches. There is a top root node. The other internal nodes follow the root node.

A function is used in order to decide which internal node will come to next. There are external nodes at the end of the tree. These nodes are also known as leaf or terminal nodes. Each terminal node is combined with a class label or value. A decision tree has paths from the root node to the terminal nodes. These paths are used as rules. And these constructed rules are used for classification.

Generally, a decision tree is constructed as follows. Firstly, a decision function is chosen in order to decide the root node for the starting. Then, the data set is participated according to the values of this root node. The tree begins to branch. The decision function is performed iteratively. The child nodes, named as internal nodes, are defined via the result of this decision function. This is repeated until a terminal node is reached. If the terminal node is reached, a class label or value is designated to the terminal node. Decision trees are often used for the classification problems.

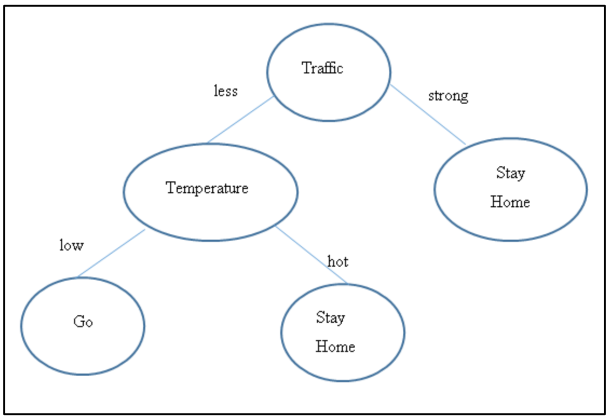


Figure 2.3 Example of classical decision tree.

Fuzzy decision trees are the structures, which are the adaptation of decision trees on fuzzy sets. These structures are also used for the classification problems. Fuzzy decision tree algorithms perform on fuzzy variables. Unlike a classical decision tree, each leaf node includes each class label with a normalized weight. All rules generated by fuzzy decision tree is used in order to make classification. Then, the classification is completed according to the result of fuzzy rule-based reasoning structure based on these generated rules (Harrington, 1991; 1993; 2017; Harrington et al., 2009).

Fuzzy decision trees are the structures, which are the adaptation of decision trees on fuzzy sets. These structures are also used for the classification problems. Fuzzy decision tree algorithms perform on fuzzy variables. As distinct from classical decision tree, each leaf node includes each class label with a normalized weight. All rules generated by fuzzy decision tree is used in order to make classification. Then, the classification is completed according to the result of fuzzy rule-based reasoning structure based on these generated rules (Harrington, 1991; 1993; 2017; Harrington et al., 2009).

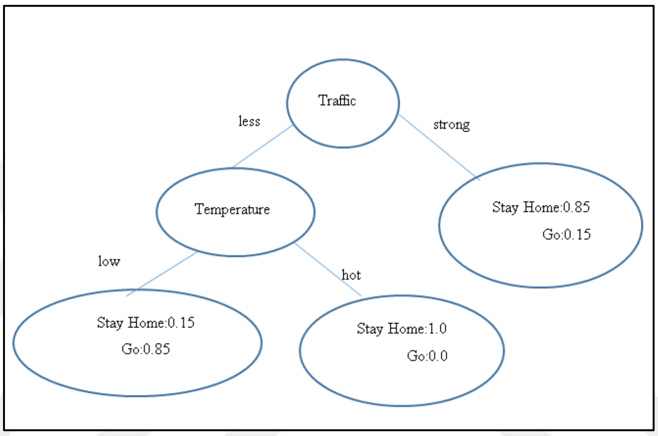



Figure 2.4 Example of fuzzy decision tree.

## 2.4 Fuzzy Reasoning Procedure on Fuzzy Decision Trees

Fuzzy reasoning procedeure is an inference process that reproduces consequences by using a set of fuzzy if-then-rules and known facts. The inference procedure can be formalized upon these rules. In general, this inference procedure is named approximate reasoning or fuzzy reasoning.

**Definition 2.7. Approximate Reasoning** Let A, A', and B be fuzzy sets of X, X' and Y, respectively. Assume that the fuzzy implication $A \to B$ is expressed as a fuzzy relation R on $X \times Y$. Then, the fuzzy set B is induced by "$x\ is\ A'$" and the fuzzy rule "$if\ x\ is\ A\ then\ y\ is\ B$" is defined by

$$\mu_{B'}(y) = max_x min[\mu_{A'}(x), \mu_R(x,y)] \tag{2.9}$$

### *2.4.1 Multiple Rules with Multiple Antecedents*

The multiple rules are interpreted by using the union of the fuzzy relations corresponding to the fuzzy rules. In summary, the process of fuzzy reasoning or approximate reasoning can be divided into four steps given in Figure 2.5. Also, the steps are defined as the following (Jang et al., 1997):

**Step 1. Degrees of compatibility:** The antecedents of fuzzy rules are compared with respect to each antecedent membership function to find the degrees of compatibility.

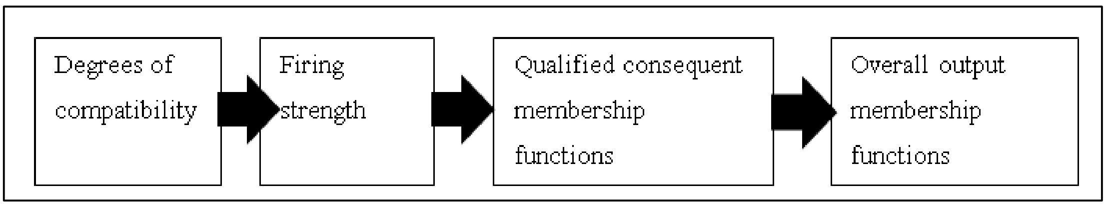



Figure 2.5 The steps of fuzzy reasoning of multiple rules with multiple antecedents.

**Step 2. Firing strength:** Degrees of compatibility are compared with respect antecedent membership functions in a rule using fuzzy AND or OR operators. The aim is to form a firing strength. It shows the degree at which the antecedent part of the rule is satisfied.

**Step 3. Qualified (induced) consequent Membership Functions (MF):** The firing strength is applied to the consequent membership functions of a rule to generate a qualified consequent membership functions. It shows how the firing strength performed in a fuzzy implication.

**Step 4. Overall output membership functions:** All the qualified consequent membership functions are applied in order to reach an overall output membership function.

Assume that a generalized modus potent problem given as

**Premise 1 (fact):** $x\ is\ A'$ and $x\ is\ B'$

**Premise 2 (rule 1):** $if\ x\ is\ A_1\ and\ y\ \ is\ B_1\ then\ z\ is\ C_1$

**Premise 3 (rule 2):** $if\ x\ is\ A_2 and\ y\ \ is\ B_2\ then\ z\ is\ C_2$

**Consequence (conclusion)** $z\ is\ C''$.

Let $R_1 = A_1 \times B_1 \to C_1$ and $R_2 = A_2 \times B_2 \to C_2$ . If the the max-min composition operator $\circ$ is assumed the distributive over the U operator, it can be given as follows:

$$
\begin{aligned}
C' &= (A' \times B') \circ (R_1 \cup R_2) \\
&= [(A' \times B') \circ R_1\ ] \cup [(A' \times B') \circ R_2] \\
&= ({C_1}' \times {C'}_2)
\end{aligned}
$$

where ${C'}_1$ and ${C'}_2$ are the inferred fuzzy sets for rules 1 and rules 2, respectively.

The fuzzy reasoning for multiple rules with multiple antecedents are given schematically in Figure 2.6 which shows graphically the operation of the fuzzy reasoning for multiple rules with multiple antecedents in graphics.

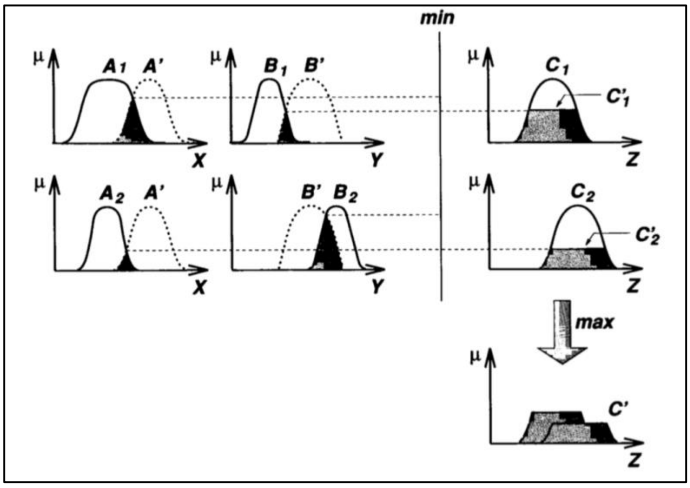



Figure 2.6 Fuzzy reasoning for multiple rules with multiple antecedents.

### 2.4.2 *Fuzzy If-Then Rules*

Let a fuzzy if-then rule (also known as fuzzy rule, fuzzy implication, or conditional statement) is given as the following form

$$if\ x\ is\ A\ then\ y\ is\ B$$

where A and B are linguistic values defined by fuzzy sets on universes of discourse X and Y, respectively. Generally, "x is A" is named as the antecedent or premise, and "y is B" is named as the consequence or conclusion. In daily life, humans generally use examples of fuzzy if-then rules, such as the following: “If an olive is black, then it is ripe”. It is necessary to formalize the rule $A \rightarrow B$ as a binary fuzzy relation R on the product space $X \times Y$. In literature, a fuzzy rule $A \rightarrow B$ can be represented as,

$$R = A \rightarrow B = A \times B = \int_{X \times Y} \mu_A(x) \tilde{*} \mu_B(x) / (x, y)$$

where $\tilde{*}$ is a T-norm operator and $A \rightarrow B$ is used again to represent the fuzzy relation R.

### 2.4.3 *Fuzzy Rule Based Classification System (FRBCS)*

Fuzzy rule based classification system (FRBCS) is an important issue in the field of pattern recognition in order to solve classification problems. This uses linguistic labels in the antecedents of their rules. This behaviour of the system provides computational flexibility. In real life, it has been used for the solution of various kind of problems, such as image processing (Nakashima, Schaefer, Yokota, & Ishibuchi, 2007), medical problems (Sanz et al., 2014), chemometrics (Nasibov et al., 2016), etc.

It is necessary to define a classification problem for FRBCS. Classification problem is a supervised learning problem. A set of training samples, named as training set, is used in order to solve this kind of problems. Each training samples has its class label. A mapping function called as classifier is used as the construction of the classification model. The model is used in order to assign class label to a new sample.

Assume that a training set consists of p samples. $x_p = (x_{p1}, \dots, x_{pn})$ is the p$th$ sample of the training set where $x_{pi}$ is the value of the i$th$ attribute ($i = 1,2 \dots, n$) of the p$th$ training sample. Target variable includes class labels as given $y_p \in C = \{C_1, C_2, \dots, C_m\}$, where m is the number of classes of the problem (Ishibuchi, Nakashima, & Nii, 2004; Elkano et al., 2015).

FRBCSs can be summarized with two main components as given below (Elkano et al. 2015; Nasibov et al., 2016):

**Knowledge Base:** The rule base (RB) and the database is included in knowledge base. The rules and the membership functions are stored in it.

**Fuzzy Reasoning Method:** The classification model is performed on the samples via the information stored in the knowledge base.

Fuzzy decision tree approach is also a kind of fuzzy classifier. The algorithm constructs a fuzzy decision tree. Each path handled from the root node to the terminal node is assumed as a fuzzy rule. It means that the constructed fuzzy decision tree model includes multivariate rules with multivariate antecedents. This model can be thought as fuzzy rule based system while it is being used to make the classification.

The structure of FRBCS uses a preferred fuzzy decision tree algorithm in learning process in order to generate fuzzy rules based on linguistic labels. Each path of the tree shows the fuzzy rules. Rule Weights (RW) for each $l$ class are stored at each leaf node. $RW_{jl}$ shows j*th* rule weight handled from fuzzy confidence value $CF_{jl}$ which equals to $RW_{jl}$. A classification problem with fuzzy decision tree model combined with Fuzzy Rule Based Classification System is summarized in Figure 2.7.

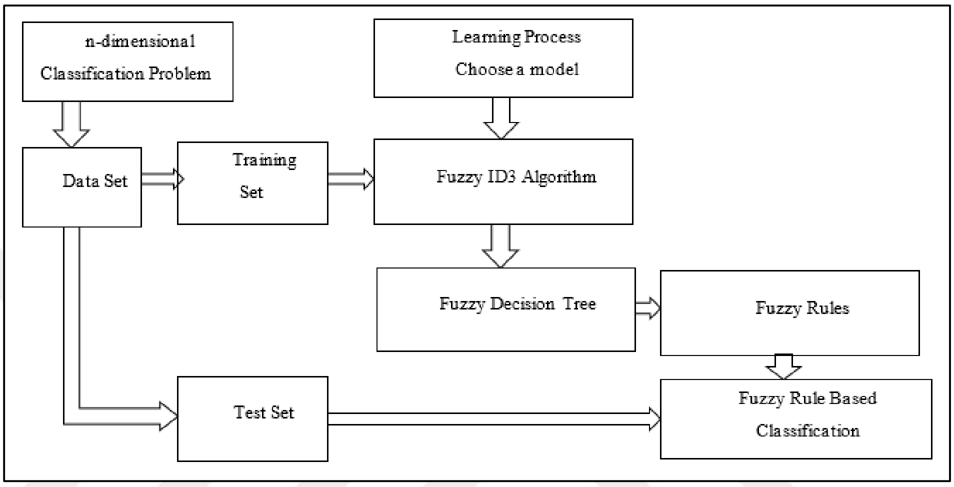



Figure 2.7 A classification problem with Fuzzy ID3 algorithm combined with FRBC

# CHAPTER THREE
# FUZZY ID3 ALGORITHM BASED ON FCM FOR NUMERIC DATA

## 3.1 Introduction

In this chapter, information about Fuzzy ID3 Interactive Dichotomizer 3 (Fuzzy ID3) algorithm and the adaptation of this algorithm via fuzzy c-means (FCM) which can work on numeric data is given. The phases of this approach will be discussed in the following subchapters.

## 3.2 Fuzzy ID3 Interactive Dichotomizer3 (Fuzzy ID3)

Umano et al. (1994) proposed Fuzzy Interactive Dichotomizer 3 (Fuzzy ID3) which is a fuzzy decision tree builder algorithm. This algorithm is a kind of fuzzified version of ID3 algorithm proposed in Quinlan (1986). It uses crisp and fuzzy variables. This algorithm divides the training set in accordance with a variable. This variable is chosen via a measure called information gain which is based on fuzzy entropy. This measure aims to seek that the variable includes the highest qualified information.

Let *N* labelled fuzzified patterns and *n* attributes are given as $A = \{A_1, A_2, \dots, A_n\}$. For each $k$ assume that $(1 \leq k \leq n)$. The attribute $A_k$ takes $m_k$values of fuzzy subsets $\left(A_{k1}, A_{k2}, \dots, A_{km_k}\right)$. $C$ denotes the classification target attribute, taking *m* values $C_1, C_2, \dots, C_m$. The cardinality of a given fuzzy set is denoted by $M(.)$, that is, the sum of the membership values of the fuzzy set (Umano et al., 1994; Nasibov et al., 2016). The induction process of fuzzy ID3 is given step by step as the following:

**Step 1:** A root node which has a set of all data is generated. This data set should be fuzzified data set which is fuzzified with fuzzy c-means (FCM) algorithm (class number (c) of each fuzzy variable is set to 3), and it is initialized with the membership values equal to 1 for all data.

**Step 2:** The expanded attribute is selected by using the following steps:

**Step 2a:** For each linguistic label $A_{ki}(i = 1,2, \ldots, m_k)$, compute its relative frequencies with respect to class $C_j$ $(j = 1,2, \ldots, m)$.

$$p_{ki}(j) = \frac{M(A_{ki} \cap C_j)}{M(A_{ki})} \tag{3.1}$$

**Step 2b:** For each linguistic label $A_{ki}$, $(i = 1,2, \ldots, m_k)$. Compute its fuzzy classification entropy.

$$Entr_{ki} = -\sum_{j=1}^{m} p_{ki}(j) \log(p_{ki}(j)) \tag{3.2}$$

**Step 2c:** Compute the average fuzzy classification entropy of each attribute.

$$E_k = \sum_{i=1}^{m_k} \frac{M(A_{ki})}{\sum_{j=1}^{m_k} M(A_{kj})} Entr_{ki} \tag{3.3}$$

**Step 2d:** Select the attribute that maximizes the information gain $(G_k)$.

$$Attr = \underset{1 \le k \le n}{\operatorname{argmax}}(G_k), \text{ where } G_k = E - E_k \tag{3.4}$$

For class label $C_i$, $i = 1,2 \ldots, m$, compute its relative frequencies depending on class $C_i$:

$$p_i(j) = \frac{M(C_i)}{N} \tag{3.5}$$

In Eq. 3.6, E is called as the total entropy, E is calculated as below:

$$E = -\sum_{j=1}^{m} p_i(j)\log(p_i(j)) \tag{3.6}$$

**Step 2e:** Assign the selected attribute as the root node and the linguistic labels as candidate branches of the tree.

**Step 3:** Select one branch to analyze. If it is empty, the selected branch is deleted. If the selected branch is non-empty, the relative frequencies are computed by using (Eq.3.1) all objects within the branch into each class. If the relative frequency of each class is above the given threshold $\theta_r$ or all the attributes are used for the induction, the branch is terminated as a leaf node. Otherwise, the attribute is selected with the smallest average fuzzy classification entropy (Eq.3.4) among those, which has not been used as a new decision node yet and its linguistic labels are added as candidates. At each leaf, each class will have its relative frequency.

**Step 4:** Repeat Step 3 as long as there are branches to analyze. If there are no candidate branches the decision tree is completed.

### *3.2.1 The Rule Structure Generated from Each Branch of The Fuzzy Decision Tree*

After the fuzzy decision tree induction, the rules are induced from each branch. Each branch behaves as a kind of path. The rule $R_j$ is given as follows:

Rule $R_j$: If $x_1$is $A_{j1}$ and … and $x_n$ is $A_{jn}$ then Class $= C_j$ with $RW_{jl}$, where $R_j$ is the label of the j*th* rule with class *l*. $x = (x_1, \dots, x_n)$ is an n-dimensional pattern vector that represents the samples. $A_{ji}$ is a fuzzy set. $C_j \in C$ is the class label, and $RW_j$ is the rule weight. In fuzzy decision tree, each leaf node has rule weights which are computed as the relative frequency for each class (as in Step 3).

### *3.2.2 Reasoning (Classification)*

Let $x_p = (x_{p1}, \dots, x_{pn})$ be the p*th* example of the training set, which is composed of P samples, where $x_{pi}$ is the value of the i*th* attribute $i = (1,2, \dots, n)$ of the p*th* sample. Each sample belongs to class $y_p \in C = \{C_1, C_2, \dots, C_m\}$, where *m* is the number of classes of the problem. Assume that $x_p$ is a novel sample to be classified with FID3 reasoning procedure suggested in Umano et al. (1994). It is adapted into fuzzy reasoning method as given in FARC-HD (Elkano et al., 2015).

Four steps are given below which are adapted with Fuzzy ID3 reasoning structures in (Sanz et al., 2012).

**Step 1. Matching degree:** In this step, the strength of activation of the if-part for all rules handled from each path of the fuzzy decision tree in the rule base with the pattern $x_p$ is computed as;

$$\mu_{A_j}(x_p) = T\left(\mu_{A_{j1}}(x_{p1}), \dots, \mu_{A_{jn_j}}(x_{pj})\right) \tag{3.7}$$

where $\mu_{A_j}(x_{pj})$ is the matching degree of the example with i*th* antecedent of the rule $R_j$. T is a T-norm (Algebraic Product/Sum T-norm operator given in Table 3.1), and $n_j$ is the number of antecedents of the rule.

**Step 2.Association degree:** The association degree of the pattern $x_p$ with each rule in the rule base and for the class *l* is computed as follows, where $RW_{jl}$, is an associated degree of each leaf node which is at the end of each path, $R_j$, with the class *l*.

$$b_{jl}(x_p) = \mu_{A_j}(x_p).\mathrm{RW}_{jl} \tag{3.8}$$

**Step 3. Confidence degree:** In this stage, the confidence degree for each class $l$ is computed to obtain the confidence degree of a class, and the association degrees of the rules of that class are summed as given in (Umano et al., 1994).

$$\mathrm{conf}_l(x_p) = \sum_{R_j \in RB; C_j = l} b_{jl}(x_p) \qquad l = 1,2 \dots, m \tag{3.9}$$

and in which $b_{jl}(x_p)$, $j = 1,2, \dots, R$ is the association degree of the pattern $x_p$ to the class *l* according to the $jth$. rule.

**Step 4. Classification:** The class with the highest confidence degree is assigned as the predicted one (Umano et al., 1994; Elkano et al. 2015).

$$Class = \underset{l=1,\dots,m}{\operatorname{argmax}}(conf_l(x_p)) \tag{3.10}$$

Table 3.1 T-Operators used in fuzzy reasoning method.

| **Non-Parametric Operators** | | | |
|---|---|---|---|
| **Ref** | **T-norm operators** | **T-conorm operators** | |
| Zadeh(1965) | $T_1(x,y) = \min(x,y)$ | $T_1^*(x,y)) = \max(x,y)$ | |
| Algebraic Product/Sum (Weber, 1983; Bandler & Kohout, 1980) | $T_2(x,y) = \mathrm{x.y}$ | $T_2^*(x,y)) = x + y - x.y$ | |
| Bounded Product/Sum (Giles, 1976) | $T_3(x,y) = \max(0, x + y - 1)$ | $T_3^*(x,y) = \min(1, x + y)$ | |
| Nonparametric Hamacher (Oussallah, 2003) ($\lambda = 0$) | $T_4(x,y)) = \dfrac{x.y}{(x + y - x.y)}$ | $T_4^*(x,y) = \dfrac{x + y - 2.x.y}{1 - x.y}$ | |
| **Parametric Operators** | | | |
| **Ref** | **T-norm operators** | **T-conorm operators** | **Parametric Range** |
| Hamacher (Oussallah, 2003) | $T_5(x,y)) = \dfrac{x.y}{\lambda + (1-\lambda)(x + y - x.y)}$ | $T_5^*(x,y)) = \dfrac{x + y - (2-\lambda).x.y}{\lambda + (1-\lambda)(1 - x.y)}$ | $\lambda \geq 0$ |
| Yager (1980) | $T_6(x,y) = \max(1 - ((1-x)^p + (1-y)^p)^{1/p}, 0)$ | $T_6^*(x,y) = \min((x^p + y^p)^{1/p}, 1)$ | $p = (0,1)$ |
| Dombi (1982) | $T_7(x,y)) = \dfrac{1}{1 + ((\frac{1}{x} - 1)^{\lambda} + (\frac{1}{y} - 1)^{\lambda})^{1/\lambda}}$ | $T_7^*(x,y)) = \dfrac{1}{1 + ((\frac{1}{x} - 1)^{-\lambda} + (\frac{1}{y} - 1)^{-\lambda})^{-1/\lambda}}$ | $\lambda = (0,1)$ |
| Dubois&Prade (1986) | $T_8(x,y)) = \dfrac{x.y}{\max(x,y,\lambda)}$ | $T_8^*(x,y) = 1 - \dfrac{(1-x).(1-y)}{\max(1-x,\ 1-y,\lambda)}$ | $\lambda = (0,1)$ |
| Weber(1983) | $T_9(x,y)) = \max(\dfrac{x + y - 1 + \lambda.x.y}{1 + \lambda}, 0)$ | $T_9^*(x,y)) = \min(x + y + \lambda.x.y, 1)$ | $\lambda = (0,1)$ |

## 3.3 Proposed Reasoning Approach for Fuzzy ID3 Based on T-Operators

In this study, it is aimed to search the effects of different T-Operators on the reasoning process in fuzzy decision tree based on fuzzyID3 algorithm. Therefore, T-operators are adapted into the reasoning process (Farahbod & Eftekhari, 2012).

It is assumed that $x_p$ is a novel sample to be classified. In this subchapter, the steps of novel proposed reasoning approach in order to make reasoning will be studied by using the rules generated from a fuzzy decision tree.

### *3.3.1 Overview of T-Operators*

T-norm and T-conorm operators developed from the triangular inequalities, are also named as T-Operators. These operators were generated from the studies of probabilistic metric spaces (Menger, 1942; Schweizer & Sklar, 1973).

Their aim is to calculate the intersection and union of two fuzzy sets. In literature, there are various types of T-operators which work better in some decision-making situations (Dubois & Prade, 1986).

While determining a set of T-operators for a decision-making problem, their properties, the accuracy model, their simplicity, computer and hardware implementations, etc. gain importance are taken account.

*Union (Disjunction):* The union of two fuzzy sets A and B is a fuzzy set C written as $C = A \; or \; B$, whose membership function (MF) is related to those of A and B by

$$\mu_C(x) = \left(\mu_A(x) \vee \mu_B(x)\right) \tag{3.11}$$

*Intersection (Conjunction):* The intersection of two fuzzy sets A and B is a fuzzy set C, written as $C = A \; or \; B$, whose MF is related to those of A and B by

$$\mu_C(x) = \left(\mu_A(x) \wedge (x)\right) \tag{3.12}$$

T-norms and T-conorms are two placed functions from $[0,1] \times [0,1]$ to $[0,1]$ that are monotonic, commutative and associative.

**Definition 3.1.** Let $T: [0,1] \times [0,1] \rightarrow [0,1]$. T is a T-norm if and only if (iff) for all $x, y, z \in [0,1]$ (Gupta & Qi, 1991):

$T(x, y) = T(y, x)$ commutativity
$T(x, y) \leq T(x, z)$ if $y \leq z$ (monotonicity)
$T\big(x, T(y, z)\big) = T(T(x, y), z)$ (associativity)
$T(x, 1) = x$

**Definition 3.2.** Let $T^*: [0,1] \times [0,1] \rightarrow [0,1]$. T* is a T*-conorm if and only if (iff) for all $x, y, z \in [0,1]$ Gupta & Qi, 1991):

$T^*(x, y) = T^*(y, x)$ commutativity
$T^*(x, y) \leq T^*(x, z)$ if $y \leq z$ (monotonicity)
$T^*\big(x, T^*(y, z)\big) = T^*(T^*(x, y), z)$ (associativity)
$T^*(x, 0) = x$

### *3.3.2 Proposed Reasoning Approach*

It is assumed that $x_p$ is a novel sample to be classified with proposed reasoning approach by using different T operators as defined in Gupta & Qi (1991). Then, the steps of this novel approach are given as following:

**Step 1. Matching degree:** In this step, the strength of activation of the if-part for all rules handled from each path of the fuzzy decision tree in the rule base on the pattern $x_p$ is computed (Eq. 3.13.):

$$\mu_{A_j}\big(x_{pi}\big) = T(\mu_{A_{j1}}\big(x_{p1}\big), \dots, \mu_{A_{jn_j}}(x_{pn_j})) \tag{3.13}$$

where $\mu_{A_j}(x_{pi})$ is the matching degree of the example with $ith$ antecedent of the rule $R_j$. T is a T-norm (listed in Table 3.1) and $n_j$ is the number of antecedents of the rule.

**Step 2: Association degree:** The association degree of the pattern $x_p$with each rule in the rule base and for the class $l$ is computed as follows where $RW_{jl}$ is an associated degree of each leaf node which is at the end of each path, $R_j$, with the class $l$ (Eq. 3.14). T is a T-norm listed in Table 3.1.

$$b_{jl}(x_p) = T(\mu_{A_j}(x_p), RW_{jl}) \tag{3.14}$$

**Step 3: Confidence degree:** In this stage, the confidence degree for each class is computed. To obtain the confidence degree of a class, the association degrees of the rules of that class are aggregated by using conjunction operators (Eq. 3.15) where T* is a T-conorm listed in Table 3.1.

$$conf_l(x_p) = T^*(b_{1l}(x_p), b_{2l}(x_p), \ldots, b_{Rl}(x_p)) \tag{3.15}$$

and in which $b_{jl}(x_p)$, $j = 1,2, \ldots, R$, is the association degree of the pattern $x_p$ to the class $l$ according to the $jth$ rule.

**Step 4: Classification:** The class with the highest confidence degree is assigned as the predicted one (Eq. 3.16) (Umano et al., 1994).

$$Class = \underset{l=1,\ldots,m}{\operatorname{argmax}}(conf_l(x_p)) \tag{3.16}$$

# CHAPTER FOUR
# FUZZY ID3 ALGORITHM BASED ON LINGUISTIC DATA BY USING WABL DEFUZZIFICATION METHOD

## 4.1 Introduction

In this chapter, a novel fuzzy ID3 algorithm for linguistic data is proposed. The information is given briefly about linguistic data defined by using triangular fuzzy numbers given as L-R fuzzy numbers. Weighted Averaging Based on Levels (WABL), the adaptation of fuzzy c-means (FCM) and Fuzzy ID3 Classification Model for High Dimensional Problems are summarized, respectively.

## 4.2 Linguistic Data and Its Representation

Humans use words or sentences in daily life. These words or sentences are used in order to make the decisions. Each word or sentence is defined as the value of linguistic variables. The information is given about the concept of linguistic variable in Subchapter 2.2.3. In this approach, L-R representation of fuzzy number is used as given in Definition 4.1 (Nasibov, Baskan, & Mert, 2005).

**Definition 4.1.** The set of real numbers is denoted by E. Let $\mathbf{L} = \{L | L: [0,1] \rightarrow E\}$ be a class of monotone non-decresasing functions and $\mathbf{R} = \{R | R: [0,1] \rightarrow E\}$ be a class of monotone non-increasing functions. Both L and R are left-continuous functions, and $\forall \alpha \in [0,1], L(\alpha) > -\infty, \forall L \in \boldsymbol{L}$ and $R(\alpha) < \infty, \forall R \in \boldsymbol{R}$**.** Any fuzzy subset A of the number of axis E or fuzzy number A can be defined by the following L-R representation:

$$A = \cup_{\alpha \in (0,1]} (\alpha, A^{\alpha}) \quad (4.1)$$

where

$$A^{\alpha} = [L_A(\alpha), R_{\alpha}(\alpha)] = \{t \in E | L_A(\alpha) \leq t \leq R_{\alpha}(\alpha)\} \quad (4.2)$$

For this representation, it is assumed that $A^1 \neq \emptyset$

## 4.3 Weighted Averaging Based on Levels (WABL)

Weighted Averaging Based on Levels (WABL) method is defined in (Nasibov, 2002; Nasibov & Mert, 2007). It is a kind of defuzzification method such as mean of maxima (MOM), centroid etc. Let A is a fuzzy number indicated by L-R representation. The average representative of this fuzzy number is defined by the formula given below:

$$I(A) = \int_0^1 \big(c_L L_A(\alpha) + c_R R(\alpha)\big) p(\alpha) d\alpha \tag{4.3}$$

where coefficients $c_L$ and $c_R$ are the weight coefficients of left and right sides respectively (pessimism/optimism parameters) , and $p(\alpha)$ is the distribution function of the importance of the level sets. The weights satisfy the following normality and non-negativity conditions:

$$c_L \geq 0, c_R \geq 0, c_L + c_R = 1 \tag{4.4}$$

$$p: [0,1] \rightarrow E_+, \int_0^1 p(\alpha) d_\alpha = 1 \tag{4.5}$$

$c_L$, $c_R$ and $p(\alpha)$ are called as WABL strategy parameters.

A way to determine the $p(\alpha)$ function is given by using the following formula (Eq.4.6) (Nasibov & Mert, 2007):

$$p(\alpha) = (k+1)\alpha^k \tag{4.6}$$

where $k > 0$ is a parameter.

It is valid for the theorem given below and proven in the study (Nasibov & Mert, 2007):

Theorem 1. Let $A = (a, b, c)$ will be termed a parametric triangular fuzzy number with parameter $s > 0$ and suppose that the distribution function of the importance of the degrees has the form (Eq. 4.7). Then, the following formula for WABL is valid:

$$I(A) = c_R\left(c - \frac{k+1}{k+s+1}(c-b)\right) + c_L\left(c - \frac{k+1}{k+s+1}(b-a)\right) \qquad (4.7)$$

## 4.4 Fuzzy ID3 Algorithm Based on Linguistic Data By Using WABL Defuzzification Method (Fuzzy ID3-L-WABL Approach)

A novel fuzzy ID3 approach is proposed for linguistic data which is defined as fuzzy. It aims to generate a fuzzy decision tree on L-R fuzzy data. In order to apply fuzzy ID3, it is necessary to apply WABL for the defuzzification on L-R fuzzy data. Then, Fuzzy c-means (FCM) algorithm is used in order to fuzzify each variable. At the end, fuzzy ID3 algorithm is performed. Then, the reasoning can be performed by using the rules generated from the fuzzy decision tree. This procedure is summarized in Figure 4.1, in graphic.

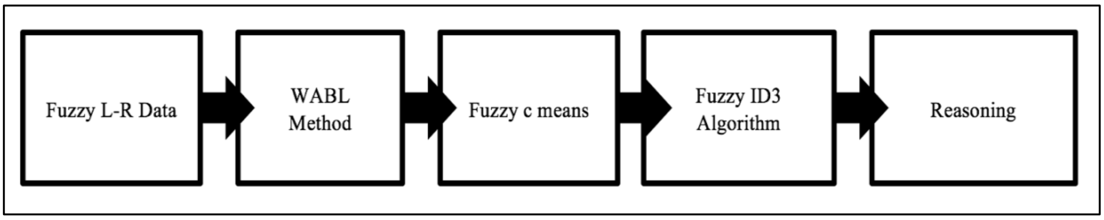



Figure 4.1 The process of fuzzy decision tree induction based on linguistic data by using WABL defuzzification method.

***Algorithm of the induction process Fuzzy ID3 Algorithm Based on Linguistic Data by Using WABL Defuzzification Method***

Assume that the data set is defined with fuzzy data as linguistic variable.

**Step 1:** The average representative of each fuzzy term is calculated by using Eq. 4.7. characterized as $I(A)$ for each fuzzy variable.

**Step 2:** Fuzzy c-means algorithm defined in Subchapter 2.2.4. is applied to the modified dataset to construct fuzzy sets of linguistic terms in Step 1.

**Step 3:** Generate a root node, which has a set of all fuzzified data set by using FCM algorithm.

**Step 4:** The expanded attribute is selected by using the following steps:

**Step 4a:** For each linguistic label$A_{ki}(i = 1,2, \dots, m_k)$, compute its relative frequencies depending on class $C_j(j = 1,2, \dots, m)$

$$.p_{ki}(j) = \frac{M(A_{ki} \cap C_j)}{M(A_{ki})} \tag{4.8}$$

**Step 4b:** For each linguistic label $A_{ki}(i = 1,2, \dots, m_k)$, compute its fuzzy classification entropy.

$$Entr_{ki} = -\sum_{j=1}^{m} p_{ki}(j) \log(p_{ki}(j)) \tag{4.9}$$

**Step 4c:** Compute the average fuzzy classification entropy of each attribute.

$$E_k = \sum_{i=1}^{m_k} \frac{M(A_{ki})}{\sum_{j=1}^{m_k} M(A_{kj})} Entr_{ki} \tag{4.10}$$

**Step 4d:** Select the attribute that maximizes the information gain ($G_k$).

$$Attr = \underset{1 \le k \le n}{\operatorname{argmax}}(G_k), \text{ where } G_k = E - E_k \tag{4.11}$$

For class label $C_i$, $i = 1,2 \dots, m$, compute its relative frequencies considering class $C_i$ .

$$p_i(j) = \frac{M(C_i)}{N} \tag{4.12}$$

In Eq.4.13, E is a total entropy, and is calculated as below:

$$E = -\sum_{j=1}^{m} p_i(j) \log(p_i(j)) \tag{4.13}$$

**Step 4e:** Assign the selected attribute as the root node and the linguistic labels as candidate branches of the tree.

**Step 5** Select one branch to analyze. If it is empty, the selected branch is deleted. If the selected branch is non-empty, the relative frequencies are computed by using (Eq.4.8) all objects within the branch into each class. If the relative frequency of each class is above the given threshold $\theta_r$ or all the attributes are used for the induction, the branch is terminated as a leaf node. Otherwise, the attribute is selected with the smallest average fuzzy classification entropy (Eq.4.11) among those, which has not been used as a new decision node yet and its linguistic labels are added as candidates. At each leaf, each class will have its relative frequency.

**Step 6:** Repeat Step 4-5 as long as there are branches to analyze. It is completed.

### *4.4.1 The Rule Structure Generated from Each Branch of The Fuzzy Decision Tree*

After the fuzzy decision tree induction, the rules are induced from each branch. Each branch behaves as a kind of path. The rule $R_j$ is given as follows:

Rule $R_j$: If $x_1$is $A_{j1}$ and … and $x_n$ is $A_{jn}$ then Class $= C_{jl}$ with $RW_{jl}$, where $R_j$ is the label of the j*th* rule with class *l*. $x = (x_1, \ldots, x_n)$ is an n-dimensional pattern vector that represents the example. $A_{ji}$ is a fuzzy set. $C_{jl} \in C$ is l*th* class label for j*th* rule, and $RW_{jl}$ is the rule weight.

In fuzzy decision tree, at each leaf node has rule weights which are computed as the relative frequency for each class (as in Step 4d).

### *4.4.2 Reasoning (Classification)*

Let $x_p = (x_{p1}, \ldots, x_{pn})$ is the p*th* sample of the training set, which is composed of P samples, where $x_{pi}$ is the value of the i*th* attribute $i = (1,2, \ldots, n)$ of the p*th* sample. Each sample belongs to class $y_p \in C = \{C_1, C_2, \ldots, C_m\}$, where *m* is the number of classes of the problem. Assume that $x_p$ is a new example to be classified with FID3 reasoning procedure suggested in Subchapter 3.2.2. The steps are given below:

**Step 1:** Matching degree: In this step, the strength of activation of the if-part for all rules obtained from each path of the fuzzy decision tree in the rule base with the pattern $x_p$ is computed as

$$\mu_{A_j}(x_p) = T\left(\mu_{A_{j1}}(x_{p1}), \ldots, \mu_{A_{jn_j}}(x_{pj})\right) \tag{4.16}$$

where $\mu_{A_j}(x_{pj})$ is the matching degree of the example with i*th* antecedent of the rule $R_j$. T is one of the non-parametric T-norm operators (listed in Table 3.1), and $n_j$ is the number of antecedents of the rule.

**Step 2:** Association degree: The association degree of the pattern $x_p$ with each rule in the rule base and for the class *l* is computed as follows where $RW_{jl}$ is an association degree of each leaf node which is at the end of each path, $R_j$, with the class *l*. T is one of the non-parametric T-norm operators (listed in Table 3.1),

$$b_{jl}(x_p) = T\left(\mu_{A_j}(x_p), RW_{jl}\right) \tag{4.17}$$

**Step 3:** Confidence degree: In this stage, the confidence degree for each class is computed. To obtain the confidence degree of a class, the association degrees of the rules of that class are aggregated by using conjunction operators where T* is one of the non-parametric T-conorm operators (listed in Table 3.1),

$$conf_l(x_p) = T^*\left(b_{1l}(x_p), b_{2l}(x_p), \ldots, b_{Rl}(x_p)\right) \tag{4.18}$$

and in which $b_{jl}(x_p)$, $j = 1,2,\ldots,R$, is the association degree of the pattern $x_p$ to the class *l* according to the $j$*th* rule.

**Step 4:** Classification: The class with the highest confidence degree is assigned as the predicted one (Umano et al., 1994; Elkano et al. 2005; Nasibov et al. 2016).

$$Class = \underset{l=1,\ldots,m}{\operatorname{argmax}}(conf_l(x_p)) \tag{4.19}$$

# CHAPTER FIVE
# FUZZY ID3 ALGORITHM FOR L-R FUZZY DATA

## 5.1 Introduction

In this chapter, a novel fuzzy ID3 algorithm is presented working on directly L-R fuzzy data for the classification problems. The information is given about weighted dissimilarity measure. Then, Fuzzy k-means clustering model for fuzzy data (FkM-F) is defined. At the end, the adaptation of FkM-F and Fuzzy ID3 Classification Model for High Dimensional Problems are summarized, respectively.

## 5.2 L-R (Left-Right) Fuzzy Data

In a matrix form, a general class of fuzzy data, called L-R fuzzy data, can be defined as follow (Coppi, D'Urso, & Giordani, 2012; Ulutagay & Kantarci, 2013; Ulutagay & Kantarci, 2014; Ulutagay & Kantarci, 2015):

$$\tilde{X} = \left\{\tilde{x}_{ij} = \left(c_{1ij}, c_{2ij}, l_{ij}, r_{ij}\right)_{LR} : i = 1, \dots, n; j = 1, \dots, p\right\} \text{ fuzzy data matrix} \quad (5.1)$$

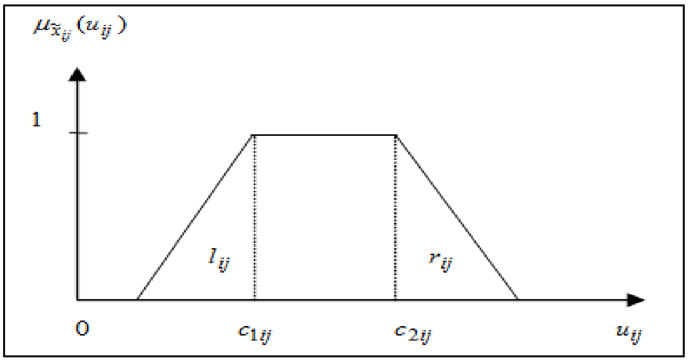



Figure 5.1 Trapezoidal membership function.

where $\tilde{x}_{ij} = \left(c_{1ij}, c_{2ij}, l_{ij}, r_{ij}\right)_{LR}$ represents the LR fuzzy variable j observed on the i-th object, $c_{1ij}$ and $C_{2ij}$, denote the left and right center, respectively, and $l_{ij}$ and $r_{ij}$ the left and the right spread, respectively, with the following membership function:

$$\mu_{\tilde{x}_{ij}}\left(u_{ij}\right) = \begin{cases} L\left(\frac{c_{1ij}-u_{ij}}{l_{ij}}\right) & , u_{ij} \leq c_{1ij}(l_{ij} > 0) \\ 1, & c_{1ij} \leq u_{ij} \leq c_{1ij} \\ R\left(\frac{u_{ij}-c_{2ij}}{l_{ij}}\right) & , u_{ij} \leq c_{2ij}(l_{rj} > 0) \end{cases} \quad (5.2)$$

where $L\left(z_{ij}\right)$ $(and\ R\left(z_{ij}\right))$ is a decreasing 'shape' function from $\Re^{+}$ to $[0,1]$ with $L(0) = 1$; $L\left(z_{ij}\right) < 1$ for all $z_{ij} > 0$, $\forall i, j; L(z_{ij}) > 0$ for all $z_{ij} < 1$, $\forall i, j; L(1) = 0$ (or $L(z_{ij}) > 0$ for all $z_{ij}$, $\forall i, j$, and $L(+\infty)$=0) (Coppi, D'Urso, & Giordani (2012)). One of the particular case of L-R fuzzy data (Figure 5.1.) is the trapeziodal one where $L\left(\frac{c_{1ij}-u_{ij}}{l_{ij}}\right) = 1 - \frac{c_{1ij}-u_{ij}}{l_{ij}}$ and $R\left(\frac{u_{ij}-c_{2ij}}{l_{ij}}\right) = 1 - \frac{u_{ij}-c_{2ij}}{l_{ij}}$ (Zimmerman, 2001).

### 5.3 Weighted Dissimilarity Measure

A weighted dissimilarity measure for fuzzy data observed on each object i.e. is suggested by considering, separately, the distances for the centers and spreads of the fuzzy data observed on each object, i.e. and using a suitable weighting system for such distance components (Coppi, D'Urso, & Giordani, 2012). Thus, by considering the $ith$ and $i'th$ objects, the results are:

$$d_F^2(\tilde{x}_i, \tilde{x}_{i'}) = w_c^2[d^2(c_{1i}, c_{1i'}) + d^2(c_{2i}, c_{2i'})] + w_s^2[d^2(l_i, l_{i'}) + d^2(r_i, r_{i'})] \quad (5.3)$$

$d^2(c_{1i}, c_{1i'}) = \|c_{1i} - c_{1i'}\| =$ Euclidean distance between the left centers $c_{1i}$ and $c_{1i'}$.

$d^2(c_{2i}, c_{2i'}) = \|c_{2i} - c_{2i'}\| =$ Euclidean distance between the right centers $c_{2i}$ and $c_{2i'}$.

$d^2(l_i, l_{i'}) = \|l_i - l_{i'}\| =$ Euclidean distance between the left spreads $l_{1i}$ and $l_{1i'}$.

$d^2(r_i, r_{i'}) = \ \|r_i - r_{i'}\| =$ Euclidean distance between the right spreads $r_{1i}$ and $r_{1i'}$.

$c_{1i} \equiv (c_{1i1}, \dots, c_{1ij}, \dots, c_{1ip})$ 'and $c_{1i'} \equiv (c_{1i'1}, \dots, c_{1i'j}, \dots, c_{1i'p}{}'$

$c_{2i} \equiv (c_{2i1}, \dots, c_{2ij}, \dots, c_{2ip})$ 'and $c_{2i'} \equiv (c_{2i'1}, \dots, c_{2i'j}, \dots, c_{2i'p}{}')$

$l_i \equiv (l_{i1}, \dots, l_{ij}, \dots, l_{ip})$ 'and $l_{i'} \equiv (l_{i'1}, \dots, l_{i'j}, \dots, l_{i'p}{}')$

$r_i \equiv (r_{i1}, \dots, r_{ij}, \dots, r_{ip})$ 'and $r_{i'} \equiv (r_{i'1}, \dots, r_{i'j}, \dots, r_{i'p}{}')$

$w_c, w_s \geq 0$ are suitable weights for the center component and the spread component of $d^2(\tilde{x}_i, \tilde{x}_{i'})$, where $\tilde{x}_i$ and $\tilde{x}_{i'}$ denote the fuzzy data vectors, respectively, for the $ith$ and $i'th$ objects., i.e. $\tilde{x}_i = \{\tilde{x}_{ij} = (c_{1ij}, c_{2ij}, l_{ij}, r_{ij})_{LR}: j = 1, \dots., p\}$ and $\tilde{x}_{i'} = \{\tilde{x}_{i'j} = (c_{1i'j}, c_{2i'j}, l_{i'j}, r_{i'j})_{LR}: j = 1, \dots., p\}$. The weights $w_c, w_s \geq 0$ can be stored in the two dimensional vector $w \equiv (w_c w_s)'$.

If the membership function value of the centers is maximum, it is advised to accept that the weight of (left and right) center distances is higher than or at least equals to the weight of (left and right) spread distances. Then, the following conditions are assumed:

$w_c + w_s = 1$ (normalization condition) and $w_c, w_s \geq 0$ (coherence condition).

It is assumed that the weights are equal for center distances and spreads, respectively.

### 5.4 Fuzzy k-Means Clustering Model for Fuzzy Data (FkM-F)

A fuzzy clustering method based on the generalized class of the so-called symmetric $LR_2$ fuzzy data given in 5.1 and the measure given in Eq.5.2 by means of Eq.5.3 is explained in (Coppi, D'Urso, & Giordani, 2012): This technique, called Fuzzy k-means clustering model for fuzzy data (FkM-F), can be formalized as follows:

$$\min_{u_{ig}, h_g, w} : J_{FkM-F} \equiv \sum_{i=1}^{n} \sum_{g=1}^{k} u_{ig}^m d_F^2(\tilde{x}_i, h_g)$$

$$= \sum_{i=1}^{n} \sum_{g=1}^{k} u_{ig}^m \left[ w_c^2 \left[ d^2\left(c_{1i}, h_g^{c_1}\right) + d^2(c_{2i}, h_g^{c_2}) \right] + w_s^2 \left[ d^2\left(l_i, h_g^L\right) + d^2(r_i, h_g^R) \right] \right] \quad (5.4)$$

$s.t.u_{ig} = 1,$

$w \equiv (w_c, w_s)' \geq 0_2, w_c \geq w_s, w_c + w_s = 1.$

where $m > 1$ is a weighting exponent that controls the fuzziness of the obtained partition:

$u_{ig}$ indicates the membership degree of the $ith$ object in the $gth$ cluster

$d_F^2(\tilde{x}_i, \tilde{h}_g)$ shows the suggested dissimilarity measure (Eq.5.3) between the $ith$ object and the prototype of the $gth$ cluster; analogously for its components $d^2\left(c_{1i}, h_g^{c_1}\right)$ $, d^2(c_{2i}, h_g^{c_2})$ $d^2\left(l_i, h_g^L\right)$, $d^2(r_i, h_g^R)$, where the fuzzy vector $\tilde{h}_g \equiv \left\{\tilde{h}_{gi} = (h_g^{c_1}, h_g^{c_2}, h_g^L, h_g^R)_{LR} : j = 1, \ldots, p\right\}$ represents the fuzzy prototype of the $gth$ cluster, $h_g^{c_1} \equiv (h_{g_1}^{c_1}, \ldots, h_{g_j}^{c_1}, \ldots, h_{g_p}^{c_1})', h_g^{c_2} \equiv (h_{g_1}^{c_2}, \ldots, h_{g_j}^{c_2}, \ldots, h_{g_p}^{c_2})',$ $h_g^L \equiv (h_{g_1}^L, \ldots, h_{g_j}^L, \ldots, h_{g_p}^L)'$, $h_g^R \equiv (h_{g_1}^R, \ldots, h_{g_j}^R, \ldots, h_{g_p}^R)'$ are $p$ vectors, whose $jth$ element refers to the $jth$ variable, that denote, respectively, the (left and right)
centers and the (left and right) spreads of the $gth$ fuzzy protoype.

By solving the constrained quadratic minimization problem (Eq.5.4) via the Lagrangian multiplier method with respect to $u_{ig}$ and by setting the first derivative of Eq. 5.4.The following iterative solution is obtained with rescpect to Eq.5.5, Eq.5.6 and Eq.5.7.

$$u_{ig} = \frac{[w_c^2\left[d^2\left(c_{1i}, h_g^{C_1}\right) + d^2\left(c_{2i}, h_g^{C_2}\right)\right] + w_s^2[d^2\left(c_{1i}, h_g^{C_1}\right) + d^2\left(c_{2i}, h_g^{C_2}\right)]]^{-\frac{1}{m-1}}}{\sum_{g'=1}^{k}[w_c^2\left[d^2\left(c_{1i}, h_g^{C_1}\right) + d^2\left(c_{2i}, h_g^{C_2}\right)\right] + w_s^2[d^2\left(l_i, h_g^{C_1}\right) + d^2(r_i, h_{g'}^R)]]^{-\frac{1}{m-1}}} \tag{5.5}$$

$$h_g^{c_1} = \frac{\sum_{i=1}^n u_{ig}^m c_{1i}}{\sum_{i=1}^n u_{ig}^m}, h_g^{c_2} = \frac{\sum_{i=1}^n u_{ig}^m c_{2i}}{\sum_{i=1}^n u_{ig}^m}, h_g^L = \frac{\sum_{i=1}^n u_{ig}^m l_i}{\sum_{i=1}^n u_{ig}^m}, h_g^R = \frac{\sum_{i=1}^n u_{ig}^m r_i}{\sum_{i=1}^n u_{ig}^m} \tag{5.6}$$

and

$$w_c = \frac{\sum_{i=1}^n \sum_{g=1}^k u_{ig}^m [d^2(l_i, h_g^L) + d^2(r_i, h_g^R)]}{\sum_{i=1}^n \sum_{g=1}^k u_{ig}^m \left[d^2\left(c_{1i}, h_g^{c_1}\right) + d^2\left(c_{2i}, h_g^{c_2}\right) + d^2(l_i, h_g^L) + d^2(r_i, h_g^R)\right]}, (w_s = 1 - w_c) \tag{5.7}$$

By recognizing (Eq.5.7), the normalization condition is satisfied. It can be demonstrated that (Eq.5.4) is a parabola with respect to $w_s$ to provide for the coherence condition. Actually (Eq.5.7), matches up with the abscissa of its vertex. When it is taken $w_s > 0.5.$ , the solution in (Eq.5.7) is impossible. In addition, among the feasible solutions, it is seen that the minimum value of (Eq.5.4) is obtained with respect to $w_s$ as long as $w_s = 0.5.$ and ($w_c = 0.5.$) (D'Urso & Giordani, 2006).

***Algorithm of Fuzzy k-means clustering Algorithm FkM-F($\widetilde{X}, m, k$)***

**Step0a.** Produce randomly the membership degree matrix $U^{(0)}$ subject to (Eq.5.4).

**Step0b.** Calculate the prototypes $\widetilde{H}^{0}$ according to (Eq.5.6) using $U^{(0)}$.

**Step 1.** Upgrade the weights $w_c^{(t)}$ and $w_s^{(t)}$ according to (Eq.5.7) keeping fixed $U^{(t-1)}$ and $\widetilde{H}^{(t-1)}$ where $t \geq 1$ shows the iteration number, and set $w_c^{(t)} = w_s^{(t)} = 0.5$ if $w_s^{(t)} \geq 0.5$.

**Step 2.** Update the prototypes $\widetilde{H}^{(t)}$ according to (Eq.5.6.) keeping fixed $U^{(t-1)}$.

**Step 3**. Update the membership degree matrix $U^{(t)}$ according to (Eq.5.5) keeping fixed $\widetilde{H}^{(t)}$ and $w_c^{(t)}$ and $w_s^{(t)}$.

**Step 4.** If $\left\| U^{(t)} - U^{(t-1)} \right\| < \varepsilon$, where $\varepsilon$ is non-negative small number fixed in advance, the algorithm has converged, otherwise go to Step 1.

### 5.5 Fuzzy ID3 for LR Fuzzy Data (FuzzyID3-LR)

Fuzzy ID3 for LR Fuzzy Data (FuzzyID3-LR) aims to generate a fuzzy decision tree on L-R fuzzy data by making the fuzzification directly. In contrast with Fuzzy ID3-L-WABL approach, it is not necessary to apply WABL method for the defuzzification on L-R fuzzy data and FCM algorithm. In this approach, FkM algorithm is applied directly to Fuzzy L-R data in order to handle membership degrees. At the end, fuzzy ID3 algorithm is performed. Then, reasoning can be performed with the rules generated from the fuzzy decision tree. This procedure is given in Figure 5.2, graphically.

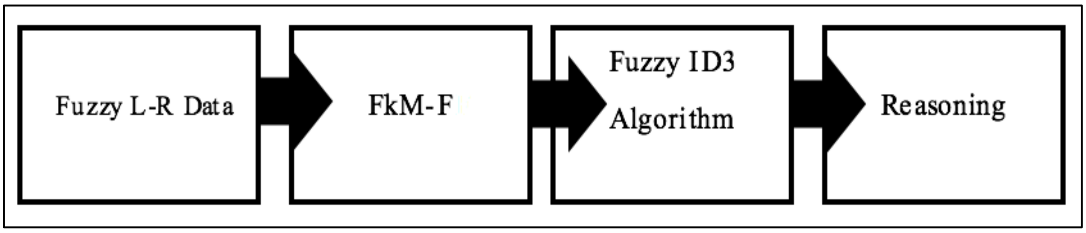



Figure 5.2 The process of fuzzy decision tree induction based on linguistic data by using WABL defuzzification method.

***Algorithm of the induction process Fuzzy ID3 Algorithm for L-R Fuzzy Data***

Assume that the data set is defined with fuzzy data as linguistic variable.

**Step 1:** Algorithm FkM-F is performed and each membership degree is calculated for each fuzzy term defined for each fuzzy variable.

**Step 2:** A root node which has a set of all fuzzified data set is generated by using the data set obtained from Step 1.

**Step 3:** The expanded attribute is selected by using the following steps:

**Step 3a:** For each linguistic label$A_{ki}(i = 1,2, \ldots, m_k)$, compute its relative frequencies considering class $C_j(j = 1,2, \ldots, m)$

.
$$p_{ki}(j) = \frac{M(A_{ki} \cap C_j)}{M(A_{ki})} \tag{5.8}$$

**Step 4:** The expanded attribute is selected with the following steps:

**Step 4a:** For each linguistic label$A_{ki}(i = 1,2, \ldots, m_k)$, compute its relative frequencies depending on class $C_j(j = 1,2, \ldots, m)$

$$p_{ki}(j) = \frac{M(A_{ki} \cap C_j)}{M(A_{ki})} \tag{5.9}$$

**Step 4b:** For each linguistic label $A_{ki}(i = 1,2, \dots, m_k)$, compute its fuzzy classification entropy.

$$\mathrm{Entr}_{ki} = -\sum_{j=1}^{m} p_{ki}(j)\log(p_{ki}(j)) \quad (5.10)$$

**Step 4c:** Compute the average fuzzy classification entropy of each attribute.

$$E_k = \sum_{i=1}^{m_k} \frac{M(A_{ki})}{\sum_{j=1}^{m_k} M(A_{kj})} \mathrm{Entr}_{ki} \quad (5.11)$$

**Step 4d:** Select the attribute that maximizes the information gain ($G_k$).

$$\mathrm{Attr} = \underset{1 \le k \le n}{\mathrm{argmax}}(G_k), \text{ where } G_k = E - E_k \quad (5.12)$$

For class label $C_i$, $i = 1,2 \dots, m$, compute its relative frequencies depending on class $C_i$ .

$$p_i(j) = \frac{M(C_i)}{N} \quad (5.13)$$

In Eq.5.14, E is a total entropy, and is calculated as below:

$$E = -\sum_{j=1}^{m} p_i(j)\log(p_i(j)) \quad (5.14)$$

**Step 4e:** Assign the selected attribute as the root node and the linguistic labels as candidate branches of the tree.

**Step 5** Select one branch to analyze. If it is empty, the selected branch is deleted. If the selected branch is non-empty, the relative frequencies are computed by using (Eq.5.8) all the objects within the branch into each class. If the relative frequency of each class is above the given threshold $\theta_r$ or all the attributes are used for the induction, the branch is terminated as a leaf node. Otherwise, the attribute is selected with the smallest average fuzzy classification entropy (Eq.5.12) among those, which has not been used as a new decision node yet and its linguistic labels are added as candidates. At each leaf, each class will have its relative frequency.

**Step 6:** Repeat Step 4 as long as there are branches to analyze. It is completed.

### *5.5.1 The Rule Structure Generated from Each Branch of The Fuzzy Decision Tree*

After the fuzzy decision tree induction, the rules are induced from each branch. Each branch behaves as a kind of path. The rule $R_j$ is given as follows:

Rule $R_j$: If $x_1$ is $A_{j1}$ and … and $x_n$ is $A_{jn}$ then Class = $C_{jl}$ with $RW_{jl}$, where $R_j$ is the label of the j*th* rule with the class *l*. $x = (x_1, \dots, x_n)$ is an n-dimensional pattern vector that represents the example. $A_{ji}$ is a fuzzy set. $C_{jl} \in C$ is the class label for j*th* rule, and $RW_{jl}$ is the rule weight.

At each leaf node has rule weights which are computed as the relative frequency for each class (as in Step 4d).

### *5.5.2 Reasoning (Classification)*

Let $x_p = (x_{p1}, \dots, x_{pn})$ be the p*th* example of the training set, which is composed of P samples, and where $x_{pi}$ is the value of the i*th* attribute $i = (1,2, \dots, n)$ of the p*th* sample. Each sample belongs to class $y_p \in C = \{C_1, C_2, \dots, C_m\}$, where *m* is the number of classes of the problem. A novel sample $x_p$ is classified with FID3 reasoning procedure whose steps are performed as given in Subchapter 4.4.2.

# CHAPTER SIX
# EXPERIMENTAL FRAMEWORK

## 6.1 Introduction

In this chapter, the experimental framework has been described to evaluate the beneficialness of the proposals. In the following subschapters, firstly a survey on geographic classification of virgin olive oil is explained. This survey focuses on fuzzy decision trees to solve the geographic characterization problem. The behaviour of different T-operators on the fuzzy reasoning procedure is also examined on virgin olive oil dataset. The study is encouraged with statistical tests.

Secondly, the behaviour of Fuzzy ID3-L-WABL and Fuzzy ID3-LR are analyzed on six data sets chosen from the real-world databases. In this part of the experimental framework, the performances of classical Fuzzy ID3, Fuzzy ID3-L-WABL, and Fuzzy ID3-LR are evaluated. The statistical comparisons are performed based on non parametric T-operators given in Table 3.1.

## 6.2 A Survey on Geographic Classification of Virgin Olive Oil with using T-operators in Fuzzy Decision Tree Approach

In this subchapter, the information about a survey on geographic classification of Virgin Olive Oil is given. A geographic classification system is proposed based on fuzzy decision tree approach. Proposed reasoning approach for fuzzy ID3 algorithm is also analyzed. Firstly, the description of the olive oil samples and the methodology used in chemical analyses of olive oil samples are given. Secondly, PCA results are discussed. Thirdly, data normalization is clarified. Then, the results are presented. Finally, discussion and conclusion parts are covered (Vahaplar et al., 2013; Nasibov et al., 2013; Nasiboğlu et al., 2013; Kantarci et al.,2015a; 2015b).

### *6.2.1 Olive Oil Samples*

Olives were collected from certain trees of the cultivars, some of which stands as the subject matter of this work: Ayvalik, Memecik, Kilis Yaglik, Nizip Yaglik. The samples were collected in 2002-2003, 2004-2005 and 2005-2006 harvest seasons. 101 olive oils; collected from different regions (North Aegean (33), South Aegean (53), Mediterranean (4), and South East (11)) were chosen for the experimental study (Gumuskesen and Yemişçioğlu, 2007). The analyses of fatty acids were performed according to the official method (European Community Regulation, 1991).

The olive oil samples were esterified in a methanol solution of 2N KOH for 30 minutes at 50 ºC. The gas chromatographic analyses of fatty acid methyl esters were performed on a Perkin Elmer 8600 gas chromatograph, equipped with a flame ionization detector: The column was a fused silica capillary coating with CP-WAX 52CB (Varian) length 25 meters, inner diameter 0.32 m. film thickness 0.20 m. Helium was the carrier gas at a flow rate of 1.5 mL/min. The column temperature program was initially isotherm for 10 min at 140ºC, an initially programmed rate of 1o C/min up to 160ºC, then a second rate of 2ºC/min up to 220º C and a final isotherm for 15 min. The injector and flame ionization detector temperatures were 250ºC. Samples of 0.2 L were injected into the split mode with a split ratio of 1:10. The apparatus itself carried out recording and integration. The analyses were repeated in triplicate.

The gas-chromatographic peaks were identified as corresponding fatty acid methyl esters by checking the elution order on the column and compared the retention times with those of pure standards. Results were expressed as peak area ratio percentage. The analysis of triglycerides was performed according to the official chromatographic method of the EC no. 2472/97 (European Community Regulation, 1997).

The apparatus was a Hewlett Packard HPLC instrument model 1100 consisted by a degasser, quaternary pum, manual six-way injection valve, refractometer detector, and Chemstation Software package for instrument control, data acquisition, and data analysis. A Lichrosorb FP 18 (4.6 0.25 mm) analytical column was used. The analysis of sterols was performed according to the official method of the EC no. 2568/91 (European Community Regulation, 1991).

The apparatus was a Hewlett Packard instrument model 6890 gas chromatograph, equipped with a flame ionization detector (FID); a HP-5 (Crosslinked 5% PH ME Siloxane) capillary column (30 m 0.25 mm 0.25 lm) and a 6890 Agilent automatic injector.

The determination of content of acidity, index of peroxide was performed according to the official methods of the EC. While PCA was applied in SPSS 20.0, partition coefficients and Fuzzy *c*-means algorithm were performed in MATLAB 2015. A software called as OliveDeSoft is programmed in the Visual C# for the experimental study (intel i7, 2.4 GHz, 4 Gb RAM).

### *6.2.2 OliveDeSoft*

OliveDeSoft is a novel improved version of the software named as SAPOO (Sensory Analysis Package: Olive Oil) developed using Borland C++ Builder environment (Kantarci, 2010). SAPOO was designed in order to determine the type of virgin olive oil. It uses the sensory evaluation model based on linguistic decision analysis proposed in (Martínez, Espinilla, & Perez, 2008).

In addition to SAPOO's facilities, OliveDeSoft (Olive Decision Software) includes the characterization of olive oil according to various regions. It enables making the characterization of the olive oil by using fuzzy decision tree approach. In addition, it includes different reasoning T-operators in order to observe the behaviour of the reasoning procedure. OliveDeSoft in the Visual C# is for the experimental study (intel i7, 2.4 GHz, 4 Gb RAM). When the software is run, the opening screen welcomes as given in Figure 6.1. (Kantarci et al., 2013).

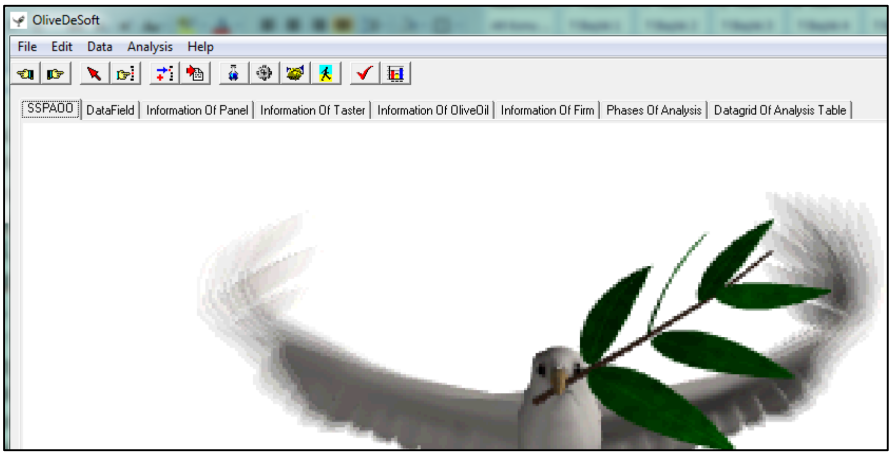



Figure 6.1 OliveDeSoft opening screen.

The aim is to get information from data and to create a useful tool for geographic characterization.

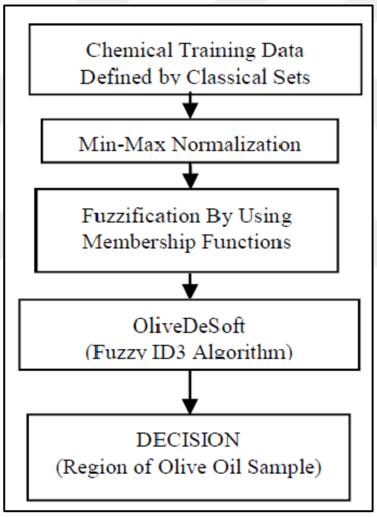



Figure 6.2 Classification Procedure by using fuzzy ID3.

The classical chemical data achieved from the olive oil samples is evaluated with Fuzzy ID3 algorithm. A fuzzy decision tree is constructed. Hence, this tree can be put into account for the classification of olive oil samples according to region. In OliveDeSoft, this can be performed fuzzy ID3 with the following menu item (given in Figure 6.3):

***Analysis→Fuzzy Decision Tree →Fuzzy ID3.***

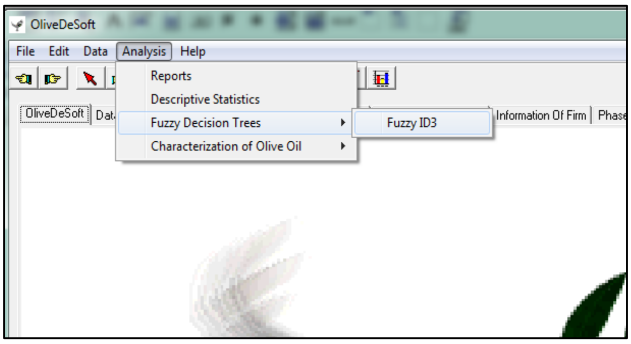



Figure 6.3 OliveDeSoft Screen "Analysis-Fuzzy Decision Tree"

Then, Screen "Fuzzy ID" is opening in order to observe the performance of different T-operators (Figure 6.4.).

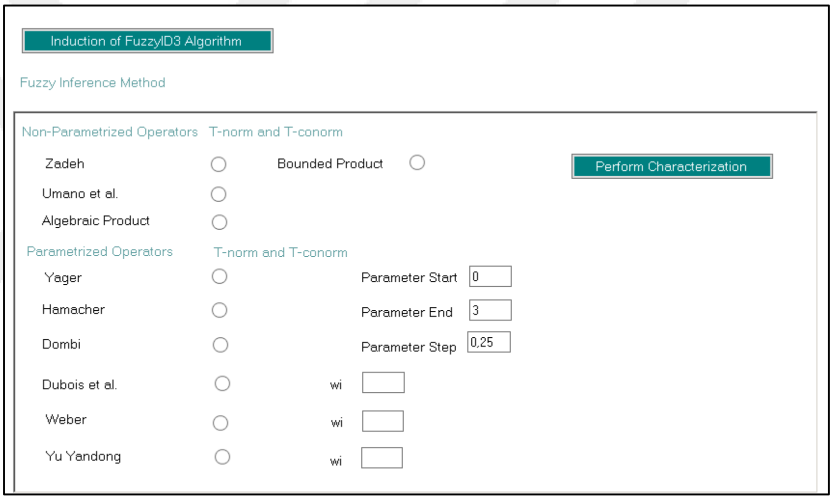



Figure 6.4 Screen "Fuzzy ID3"

With OliveDeSoft, it is possible to decide the region of olive oil sample with the following menu item (given in Figure 6.5.):

***Analysis →Characterization of Olive Oil →Satisfy the region.***

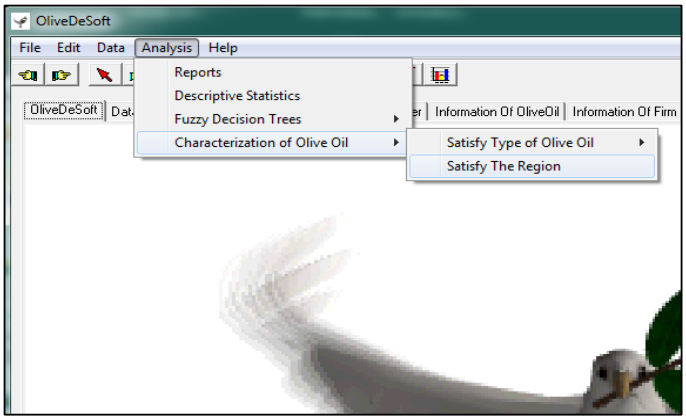



Figure 6.5 Screen “Analysis-Charcterization of olive oil”

### 6.2.3 *Implementation of PCA*

Principal component analysis was performed on this data set in order to explore the data structure. The principal components plot is given in Figure 6.6. It is clear that there is information related to the geographic origin of virgin olive oils on the results obtanied from the chemical analyses, but there is a region (Mediterranean) which has less data than the other regions so it cannot be viewed clearly. This region can be seen by collecting many more data from this region. The data implementation is performed in IBM SPSS 20.

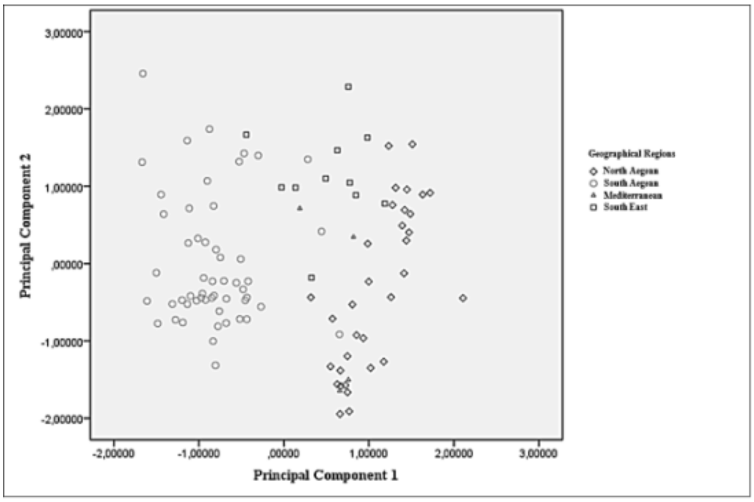



Figure 6.6 The Principal components plot on the virgin olive oil sample.

### 6.2.4 Min-Max Normalization and Fuzzy c-means Algorithm

The data set is normalized by min-max normalization. Normalization is performed to avoid domination between attributes of the data. It is a linear transformation. Let *B* is an attribute. Min *B* and Max *B* are the minimum and the maximum values of this attribute. In this case, min-max normalization maps a value $v$ of *B* into $v'$ in a new range between 0 and 1. The following formula is used for min-max normalization (Eq.6.1):

$$v' = \frac{v - min_A}{max_A - min_A} \tag{6.1}$$

In this experimental study, the data fuzzification process was performed with fuzzy *c*-means (FCM) algorithm.

### 6.2.5 Partition coefficient index

The determination of the correct number of clusters (c) for fuzzy c-means (FCM) algorithm has a crucial issue. In literature, there are some scalar measures of partitioning fuzziness, called *validity indicators* (Bezdek, 1974a; 1974b; Dunn, 1974). Partition coefficient is a scalar measure as formulized as below (Eq.6.2):

$$V_{pc} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \mu_{ij}^2 \tag{6.2}$$

whereas optimal cluster number is $\max(V_{PC}, U, c)$. Partition coefficient was used in order to determine the number of clusters (Bezdek, 1974b; Dunn, 1974). The calculated partition coefficient value for each cluster is given in Table 6.1.

### *6.2.6 Performance Measure and Statistical Tests*

The chemical measurements have imprecise information. In this study, fuzzy ID3 algorithm based on fuzzy logic is chosen. Normally, ID3 algorithm works with categorical variables. Yet, Fuzzy ID3 algorithm deals with numerical variables by using fuzzy variables. Each numeric variable transforms into fuzzy variable.

In this study, the chemical data was fuzzified by using Fuzzy c-means (FCM) algorithm. Each fuzzy variable has fuzzy terms inside of it as it is described before. The clusters are determined by using partition coefficient value. This approach uses nine different T-operators into the reasoning procedure. Classical FuzzyID3 (Umano et al., 1994) and C4.5 (Quinlan, 1993) algorithms are also performed to examine the performances.

Leave-one-out validation procedure was performed in order to measure the performances of the algorithms. Accuracy rate is a technic widely used in order to test different methods. This metric is defined as percentage of correctly classified samples (Elkano et al., 2015). Also, threshold value is set to $\theta_r = 0.75$ for the analysis. Parameters are set as Yager p=2, Hamacher p=0.25, Dombi=1, Dubois=0.25 and Weber=15 for parametric operators' experimental study. While $\theta_r = 0.75$, each operator reaches the maximum accuracy rates.

### *6.2.7 Studying fuzzy reasoning method with non-parametric operators*

C4.5 algorithm also uses entropy as splitting criteria, like ID3 algorithm. It was proposed in Quinlan (1993) to deal with the numerical data. The observed performance of this algorithm is 86.14%. Then, it is seen that the performance of classical Fuzzy ID3 algorithm with its' own reasoning method has the same performance with 86.14%.

Table 6.1The calculated partition coefficient value for each cluster number (c=2, c=3, c=4)

| **Attributes** | **c=2** | **c=3** | **c=4** | **Attributes** | **c=2** | **c=3** | **c=4** |
|---|---|---|---|---|---|---|---|
| Myristic Acid (C14:0) | 0.9202 | **0.9389** | 0.9189 | Campesterol | **0.8099** | 0.7607 | 0.7439 |
| Palmitic Acid (C16:0) | **0.8735** | 0.8207 | 0.7765 | Campestenol | **0.9998** | 0.9170 | 0.9116 |
| Palmitoleic Acid (C16:1) | **0.8313** | 0.7994 | 0.7960 | Stigmasterol | **0.9181** | 0.8006 | 0.7980 |
| Heptadecanoic Acid (C17:0) | **0.9066** | 0.8443 | 0.8035 | Delta 7 Campesterol | **0.8251** | 0.8161 | 0.8171 |
| Heptadecenoic Acid (C17:0) | **0.9153** | 0.8528 | 0.8240 | Delta 5-23 Stigmastadienol | **0.9899** | 0.8724 | 0.8852 |
| Stearic Acid (C18:0) | **0.8013** | 0.7930 | 0.7435 | Clerosterol | **0.8086** | 0.8037 | 0.7511 |
| Oleic Acid (C18:1) | **0.8797** | 0.8013 | 0.7436 | Beta-Sitosterol | **0.9027** | 0.8450 | 0.7576 |
| Linoleic Acid (C18:2) | **0.8368** | 0.7724 | 0.7441 | Sitostenol | **0.8982** | 0.8018 | 0.8076 |
| Linolenic Acid(C18:3) | **0.9998** | 0.8383 | 0.9239 | Delta 5 Avenasterol | **0.8901** | 0.8286 | 0.7609 |
| Arachidic Acid(C20:0) | **0.7967** | 0.7741 | 0.7567 | Delta 5-24 Avenasterol | **0.9143** | 0.8224 | 0.8254 |
| Gadoleic Acid (C20:1) | **0.8554** | 0.8291 | 0.7772 | Delta 7 Stigmastenol | **0.8356** | 0.7880 | 0.7352 |
| Behenic Acid(C22:0) | **0.8024** | 0.7900 | 0.7978 | Delta 7 Avenasterol | **0.8757** | 0.8368 | 0.7957 |
| Lignoceric Acid(C24:0) | 0.7754 | **0.7934** | 0.7670 | Total Beta Sitosterol | **0.8370** | 0.7756 | 0.8132 |
| Cholesterol | **0.8389** | 0.8002 | 0.7988 | Total Sterol | **0.8660** | 0.8109 | 0.7557 |
| Brassicasterol | **0.8640** | 0.7955 | 0.8046 | Erythrodiol_Uvaol | **0.8693** | 0.8178 | 0.7623 |
| 24-Methylene | **0.9011** | 0.7803 | 0.7989 | Trilinolein | **0.8719** | 0.8035 | 0.7649 |

The performance results of non-parametric approaches given in Table 6.2 shows that the result handled from four nonparametric versions have the same performance value with handled from C4.5 algorithm. Fuzzy ID3 algorithm reasoning with Bounded Product T-operators has the minimum performance value with 85.15%.

Table 6.2 The performance results of each algorithm for non-parametric operators.

| Algorithms | Accuracy Rate (%) |
|---|---|
| C4.5 | **86.14*** |
| FuzzyID3_reasoning with Classical | **86.14*** |
| FuzzyID3_ reasoning with Zadeh T-Opeators $T_1$ & $T_1^*$ | **86.14*** |
| FuzzyID3_ reasoning with Algebraic Product/Sum $T_2$ & $T_2^*$ | **86.14*** |
| FuzzyID3_ reasoning with Bounded Product/Sum $T_4$ & $T_4^*$ | 85.15 |
| FuzzyID3_ reasoning with Non Parametric Hamacher $(\lambda = 0)$ $T_3$ & $T_3^*$ | **86.14*** |

### *6.2.8 Studying Fuzzy Reasoning Method with Parametric Operators*

The performance of fuzzy reasoning method is controlled within different parameters. It is thought that a better classification accuracy rate is reached by changing the parameters value. The performance results for each Fuzzy ID3 reasoning with parametric T-operators (listed in Table 3.1.) are given in Table 6.3.

Fuzzy ID3 reasoning with Weber T-operators (lambda: (15-17)) has the highest performance value with 87.13%. It is observed that in different parameter values, the algorithm can reach the highest performance value. The other operators reach maximum 86.14%, same as non-parametric operators.

Table 6.3 The performance results of each algorithm for parametric operators.

| **Algorithms** | **Parameter Value** | **Accuracy Rate (%)** |
|---|---|---|
| FuzzyID3_ reasoning with Hamacher max. Values $T_5$ & $T_5^*$ | (0.25-6.50) | 86.14 |
| FuzzyID3_ reasoning with Yager max. Values $T_6$ & $T_6^*$ | (2-300) | 86.14 |
| FuzzyID3_ reasoning with Dombi max. values $T_7$ & $T_7^*$ | (1-155) | 86.14 |
| FuzzyID3_ reasoning with Dubois max. values $T_8$ & $T_8^*$ | (0.25-1) | 86.14 |
| FuzzyID3_ reasoning with Weber max. values $T_9$ & $T_9^*$ | (15-17) | **87.13*** |
| FuzzyID3_ reasoning with Yuyandong max. values $T_{10}$ & $T_{10}^*$ | (100-105) | 86.14 |

The graph of accuracy rates handled from different parametric operators with $\theta_r = 0.75$ in (0-20) are given in Figure 6.7. It is supported that Fuzzy ID3 reasoning with Weber has a good performance average of 85.90% within range (0-20) and Hamacher has a performance with an average of 84.82%.

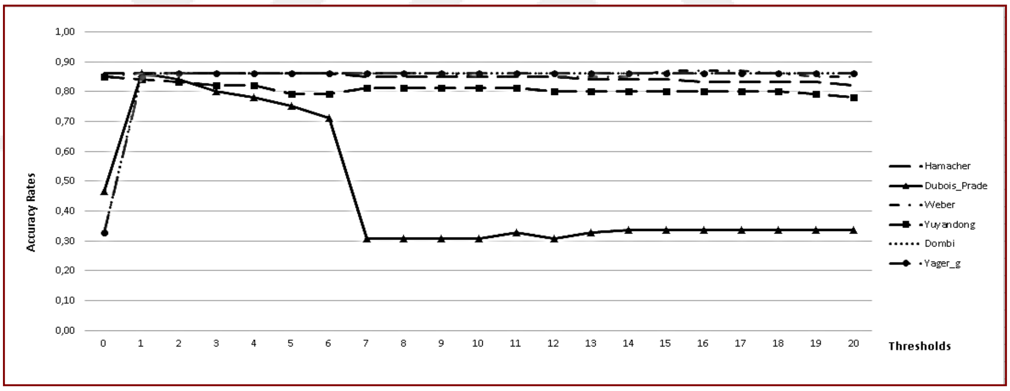



Figure 6.7 Accuracy rates handled from different parametric operators Range =(0-20) and $\theta_r$=0.75.

### 6.2.9 *Study of the Behaviour of Fuzzy ID3 Reasoning Method Based on Different T-Operators.*

The Friedman aligned ranks have been applied as a non-parametric statistical procedure in order to detect statistical differences among a group of results for 20 threshold ($\theta_r$) values in Table 6.4. This test obtains p-value as equal to zero, which shows that there are significant differences among the results.

Table 6.4 Friedman aligned ranks

| **Algorithm** | **Rank** | **Friedman aligned ranks** | |
|---|---|---|---|
| **Zadeh** | 7.68 | | |
| **Classical** | 4.72 | | |
| **Algebraic Product** | 4.72 | **Total N** | 20 |
| **Bounded Product** | 4.42 | | |
| **Non-Parametric Hamacher ($\lambda = 0$)** | 4.72 | **Test Statistic** | 95.605 |
| **Yager** | 7.65 | | |
| **Hamacher** | 4.72 | **Degrees of Freedom** | 9 |
| **Dombi** | 3.32 | | |
| **Dubois** | 4.18 | **Asymptotic Sig. (2 sided test)** | 0.000 |

The pairwise comparisons are performed. The adjusted *p*-values are taken into account in order to evaluate these pairwise comparisons among the non-parametric algorithms. The results are given in Table 6.5. Friedman aligned ranks test and pairwise comparisons were performed in IBM SPSS 20.

There are thirteen significance comparison as follows:

Dombi vs. Yager with adj. *p*-value=0.000.

Dombi vs. Zadeh with adj. *p*-value=0.000**.**

**Dombi vs. Weber** with adj. *p*-value=0.000.

Dubois vs. Yager with adj. *p*-value=0.013.

Dubois vs. Zadeh with adj. *p*-value=0.012.

**Dubois vs. Weber** with adj. *p*-value=0.000.

Bounded Product/Sum vs. Yager with adj. *p*-value=0.034.

Bounded Product/Sum vs. Zadeh with adj. *p*-value=0.031.

**Bounded Product/Sum vs. Weber** with adj. *p*-value=0.000.

**Classical vs. Weber** with adj. *p*-value=0.001.

**Algebraic Product/Sum vs. Weber** with adj. *p*-value=0.001.

**Hamacher** ($\lambda = 0$) **vs. Weber** with adj. *p*-value=0.001.

**Hamacher vs. Weber** with adj. *p*-value=0.001.

In Table 6.6, it is also seen that the highest average is handled from Weber. It is seen from pairwise comparisons that Weber has better results than Classical which is the standard version. Weber has also better results than Hamacher. Yager has better results than Dombi and Dubois. Yager, Zadeh, and Weber have better results than Bounded Product. As a result, Weber, which is a parametric operator given in bold as above, has better results than all non-parametric operators.

Also, the graph of the accuracy rates is handled for different thresholds within all approaches in Figure 6.8. Accuracy rates handled for different thresholds within different fuzzy reasoning methods are given in Table 6.7.

Table 6.5 The results of pairwise comparisons for FuzzyID3 reasoning operators with 20 different thresholds (range=0.71-0.90) via adjusted significance values

| | Weber | Zadeh | Yager | Hamacher | NP Hamacher ($\lambda = 0$) | Algebraic Product | Classical | Bounded Product | Dubois |
|---|---|---|---|---|---|---|---|---|---|
| Dombi | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Dubois | 0.000 | 0.012 | 0.013 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | |
| Bounded Product | 0.000 | 0.031 | 0.034 | 1.000 | 1.000 | 1.000 | 1.000 | | |
| Classical | 0.001 | 0.093 | 0.101 | 1.000 | 1.000 | 1.000 | | | |
| Algebraic Product | 0.001 | 0.093 | 0.101 | 1.000 | 1.000 | | | | |
| NPHamacher ($\lambda = 0$) | 0.001 | 0.093 | 0.101 | 1.000 | | | | | |
| Hamacher | 0.001 | 0.093 | 1.000 | | | | | | |
| Yager | 1.000 | 1.000 | | | | | | | |
| Zadeh | 1.000 | | | | | | | | |

It is seen that maximum value has Dombi T-operators handled for $\theta_r = 0.85$ with 88.11%. As a result, it is observed that we can also reach better results by using different threshold values. In future work, the behaviour of threshold values is planned to be researched.

### *6.2.10 Discussion and Conclusion*

In this study, it is aimed to make the geographic classification of olive oil. It is one of the basic agricultural products of Turkey, and is an important food product for the human health from past to present. So, the quality control of this product has a crucial importance and it is very difficult. In accordance with this study, chemical measurements were used in order to make on experimental study. Chemical measurements contain uncertainty. In order to deal with uncertain information, Fuzzy ID3 classifier was chosen to construct the classification of olive oil samples. Additionally, fuzzy ID3 reasoning method which is based on T-operators has been proposed. The study has targetted to see the performance of proposed fuzzy reasoning method to solve the geographic classification problem.

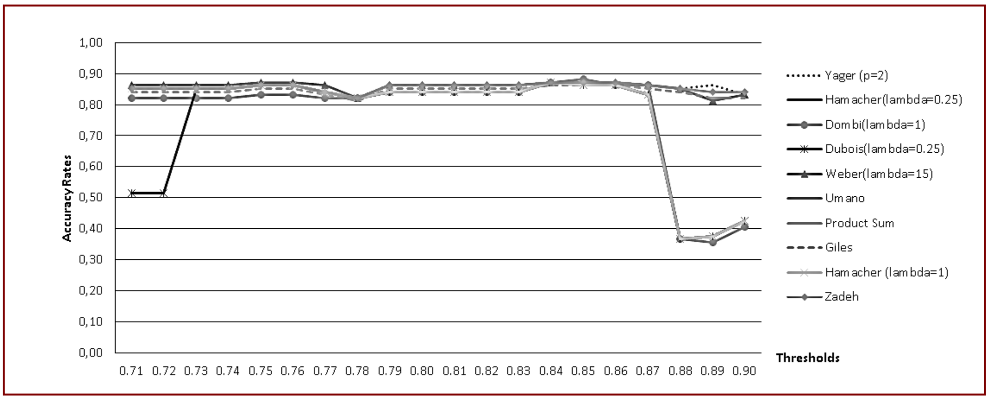



Figure 6.8 Accuracy rates handled for different thresholds.

It is observed that the results obtained from four non-parametric versions have the same performance value with the results obtained from C4.5 algorithm. Then, the performance of parametric operators are checked. As a result, it is seen that Fuzzy ID3 reasoning with Weber T-operators (lambda: (15-17)) has the highest performance value with 87.13%. Statistical procedure was performed in order to detect statistical differences among a group of results for 20 threshold ($\theta_r$) values.

It is observed that there are significant differences among the results. Also, the pairwise comparisons are performed for each approach. Weber has better results than Classical which is the standard version.

Table 6.6 The performance results (%) of each algorithm for parametric operators range =(0-20) and $\theta_r$=0.75.

| **Parameter** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** | **Average** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Hamacher** | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 85.12 | 85.12 | 85.15 | 85.15 | 85.15 | 85.15 | 84.16 | 84.16 | 84.16 | 83.17 | 83.17 | 83.17 | 83.17 | 82.18 | **84.82** |
| **Dubois_Prade** | 46.53 | 86.14 | 84.16 | 80.20 | 78.22 | 75.25 | 71.29 | 30.69 | 30.69 | 30.69 | 30.69 | 32.67 | 30.69 | 32.67 | 33.66 | 33.66 | 33.66 | 33.66 | 33.66 | 33.66 | 33.66 | 46.49 |
| **Weber** | 85.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 85.12 | 85.12 | 85.12 | 85.12 | 85.12 | 87.13 | 87.13 | 87.13 | 86.14 | 85.12 | 85.15 | **85.90** |
| **Dombi** | 32.67 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 83.59 |
| **Yager** | 32.67 | 85.12 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 86.14 | 83.54 |

Table 6.7 Accuracy rates handled for different thresholds (%).

| $\boldsymbol{\theta_r}$ | **Zadeh** | **Classical** | **Algebraic Product/Sum** | **Bounded Product/Sum** | **NP Hamacher ($\lambda=0$)** | **Yager (p=2)** | **Hamacher (p=0.25)** | **Dombi (1)** | **Dubois (0.25)** | **Weber (15)** |
|---|---|---|---|---|---|---|---|---|---|---|
| **0.71** | 85.15 | 85.15 | 85.15 | 84.16 | 85.15 | 85.15 | 85.15 | 82.18 | 51.48 | 86.14 |
| **0.72** | 85.15 | 85.15 | 85.15 | 84.16 | 85.15 | 85.15 | 85.15 | 82.18 | 51.48 | 86.14 |
| **0.73** | 85.15 | 85.15 | 85.15 | 84.16 | 85.15 | 85.15 | 85.15 | 82.18 | 85.15 | 86.14 |
| **0.74** | 85.15 | 85.15 | 85.15 | 84.16 | 85.15 | 85.15 | 85.15 | 82.18 | 85.15 | 86.14 |
| **0.75** | 86.14 | 86.14 | 86.14 | 85.15 | 86.14 | 86.14 | 86.14 | 83.16 | 86.14 | **87.13** |
| **0.76** | 86.14 | 86.14 | 86.14 | 85.15 | 86.14 | 86.14 | 86.14 | 83.16 | 86.14 | **87.13** |
| **0.77** | 84.16 | 84.16 | 84.16 | 83.17 | 84.16 | 84.16 | 84.16 | 82.18 | 84.16 | 86.14 |
| **0.78** | 82.18 | 82.18 | 82.18 | 81.19 | 82.18 | 82.18 | 82.18 | 82.18 | 82.18 | 82.18 |
| **0.79** | 86.14 | 84.16 | 84.16 | 85.15 | 84.16 | 86.14 | 84.16 | 84.16 | 84.16 | 86.14 |
| **0.80** | 86.14 | 84.16 | 84.16 | 85.15 | 84.16 | 86.14 | 84.16 | 84.16 | 84.16 | 86.14 |
| **0.81** | 86.14 | 84.16 | 84.16 | 85.15 | 84.16 | 86.14 | 84.16 | 84.16 | 84.16 | 86.14 |
| **0.82** | 86.14 | 84.16 | 84.16 | 85.15 | 84.16 | 86.14 | 84.16 | 84.16 | 84.16 | 86.14 |
| **0.83** | 86.14 | 84.16 | 84.16 | 85.15 | 84.16 | 86.14 | 84.16 | 84.16 | 84.16 | 86.14 |
| **0.84** | **87.13** | **87.13** | **87.13** | **86.14** | **87.13** | **87.13** | **87.13** | 87.13 | **87.13** | **87.13** |
| **0.85** | **87.13** | 86.14 | 86.14 | **86.14** | 86.14 | **87.13** | 86.14 | **88.11** | 86.14 | **87.13** |
| **0.86** | **87.13** | 86.14 | 86.14 | **86.14** | 86.14 | **87.13** | 86.14 | 86.14 | 86.14 | **87.13** |
| **0.87** | 86.14 | 83.17 | 83.17 | 85.15 | 83.17 | 86.14 | 83.17 | 86.14 | 83.17 | 86.14 |
| **0.88** | 85.15 | 36.63 | 36.63 | 84.16 | 36.63 | 85.15 | 36.63 | 36.63 | 36.63 | 85.15 |
| **0.89** | 84.16 | 37.62 | 37.62 | 83.17 | 37.62 | 86.14 | 37.62 | 35.64 | 37.62 | 81.19 |
| **0.90** | 84.16 | 42.57 | 42.57 | 83.17 | 42.57 | 83.17 | 42.57 | 40.59 | 42.57 | 83.17 |
| **Average** | 85.54 | 77.97 | 77.97 | 84.46 | 77.97 | 85.59 | 77.97 | 77.03 | 74.60 | **85.74** |

Hence, Weber has better results than all non-parametric operators' results. So, it is claimed that by using different parameters, better reasoning performance can be handled for the classification procedure with fuzzy ID3. In future research, there are several works to be addressed related with the adaptation of n-dimensional overlap functions (Elkano et al., 2015).

## 6.3 Study of the Behaviour of Fuzzy ID3-L-WABL and Fuzzy ID3- LR

In this Subchapter, Behaviour of Fuzzy ID3-L-WABL and Fuzzy ID3-LR are analyzed. First, the datasets selected for the experimental studies (see Subchapter 6.3.1) are explained. Second, the parameter set-up for each method (see Subchapter 6.3.2) is given. Third, Fuzzy Artemis is presented, which is programmed to make the experimental studies (see Subchapter 6.3.3). Then, the detailed information is given for the experimental study (see Subchapter 6.3.4). At last, a brief analysis is done about the results obtained from two proposed linguistic approaches.

### *6.3.1 Datasets*

In order to analyze the performance of our proposal, we have considered six datasets selected from the KEEL dataset repository (Alcala-Fdez et al., 2009; Datasets., n.d.) prepared by using 5-fold stratified cross-validation model. Table 6.8 summarizes the features of the selected datasets, showing for each dataset the number of examples, number of numerical attributes (Num Atts), and the number of classes(Class).

Table 6.8 Summary of the datasets' features used in experimental study

| ID | Dataset | Examples | Num Attr | Class |
|---|---|---|---|---|
| **1** | Iris | 150 | 4 | 3 |
| **2** | Phoneme | 5404 | 5 | 2 |
| **3** | Pima | 768 | 8 | 2 |
| **4** | Ring | 7400 | 20 | 2 |
| **5** | Sonar | 208 | 60 | 2 |
| **6** | Wdbc | 569 | 30 | 2 |

### *6.3.2 Performance Measure and Statistical Tests*

The accuracy rate is one of the most common metric measures to test performance of the different methods. It is explained with percentage of correctly classified examples related to the total number of examples. In our study, we use 5-fold cross validation. The statistical tests are done with the average of 5-fold experiments. Fuzzy c-means (FCM) algorithm and FkM-F algorithm are performed in MATLAB 2015 in order to handle the fuzzification phase of the datasets. A software called FuzzyArtemis is programmed in the Visual C# for the experimental study (intel i7, 2.4 GHz, 4 Gb RAM).

In order to give a statistical support to the analysis of the results, it is carried out Wilcoxon Signed Rank Test to perform pairwise comparisons and Aligned Friedman Test to check whether there are statistical differences among a group of methods or not. Statistical non-parametric tests are performed in IBM SPSS 23.

### *6.3.3 FuzzyArtemis*

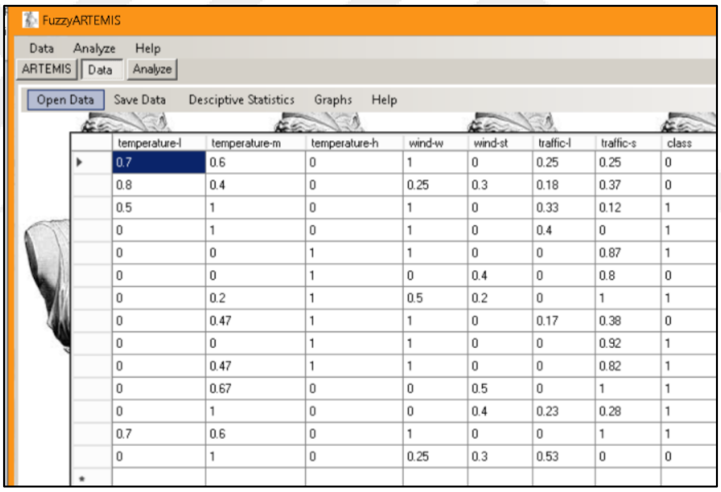
*FuzzyArtemis* aims to make experimental studies about fuzzy classification and clustering. The first menu is developed for fuzzy classification approach. It uses fuzzy ID3 algorithm in order to solve the classification problems. It is planned that fuzzy clustering approach will be made add-in for future studies.

You can handle fuzzy decision tree based Fuzzy ID3 algorithm. When FuzzyArtemis is run, the opening screen welcomes as given in Figure 6.9.


Figure 6.9 FuzzyArtemis opening screen.

Data menu help to see the data as seen in Figure 6.10. If user wants to construct fuzzy decision tree, user needs to follow the menus as given below:

***Data→Open Data →Fuzzy ID3.***




| temperature-l | temperature-m | temperature-h | wind-w | wind-st | traffic-l | traffic-s | class |
|---|---|---|---|---|---|---|---|
| 0.7 | 0.6 | 0 | 1 | 0 | 0.25 | 0.25 | 0 |
| 0.8 | 0.4 | 0 | 0.25 | 0.3 | 0.18 | 0.37 | 0 |
| 0.5 | 1 | 0 | 1 | 0 | 0.33 | 0.12 | 1 |
| 0 | 1 | 0 | 1 | 0 | 0.4 | 0 | 1 |
| 0 | 0 | 1 | 1 | 0 | 0 | 0.87 | 1 |
| 0 | 0 | 1 | 0 | 0.4 | 0 | 0.8 | 0 |
| 0 | 0.2 | 1 | 0.5 | 0.2 | 0 | 1 | 1 |
| 0 | 0.47 | 1 | 1 | 0 | 0.17 | 0.38 | 0 |
| 0 | 0 | 1 | 1 | 0 | 0 | 0.92 | 1 |
| 0 | 0.47 | 1 | 1 | 0 | 0 | 0.82 | 1 |
| 0 | 0.67 | 0 | 0 | 0.5 | 0 | 1 | 1 |
| 0 | 1 | 0 | 0 | 0.4 | 0.23 | 0.28 | 1 |
| 0.7 | 0.6 | 0 | 1 | 0 | 0 | 1 | 1 |
| 0 | 1 | 0 | 0.25 | 0.3 | 0.53 | 0 | 0 |

Figure 6.10 FuzzyArtemis "Data" screen (Open Data Command).

FuzzyArtemis "Analyze" screen is given in Figure 6.11. It is seen that the user can make the induction of fuzzy decision tree by using the button "Induction of Fuzzy ID3 Algoirthm". Then, user can choose one of the fuzzy reasoning methods among various parameteric and non-parametric approaches. In addition, the user can see the performance of the selected method by using the button "Perform The Selected Inference Method".

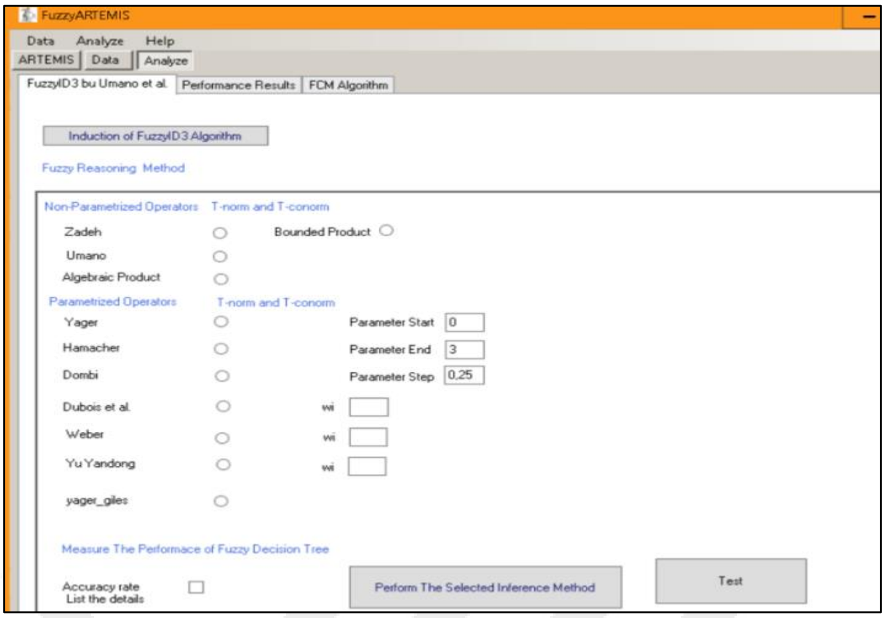



Figure 6.11 *FuzzyArtemis* “Analyze” screen.

### *6.3.4 Experimental Study*

The experimental study consists of two steps. It is aimed to show that novel fuzzy ID3 approaches working on linguistic data have good performances at least as classical fuzzy ID3 approach working on numeric data. In the first step, the performance behaviour of three approaches is analyzed by using a fixed threshold value ($\theta_r = 0.75$) and different T-operators. In the second step, the performance behaviour of of three approaches is examined by using various T-operators and different thresholds. The experimental study is supported by statistical analysis.

#### *6.3.4.1 Study of The Behaviour of Fuzzy ID3 Induction Process on Numerical Data, Fuzzy ID-L-WABL and Fuzzy ID3-LR*

The behaviour of classical Fuzzy ID3, Fuzzy ID-L-WABL, and Fuzzy ID3-LR approaches are analyzed by using T-operators on six well-known data sets. Then, the behaviour of different t-operators on fixed threshold results for three different approaches are obtained. It is aimed to show that the results handled from proposed approaches for linguistic data has good performance at least as the results handled from Fuzzy ID3 Induction process for the classical approach. Triangular fuzzy numbers (defined in Subchapter 2.2.2) are consisted of the following steps:

Step 1. Each attribute's value is assigned as the center (b).

Step 2. Each center value is multipled with a random number generated between (0-0.20). This random number represents $l_{ij}$ or $r_{ij}$ which is defined in Subchapter 5.2.

Step 3. Then, left (a) value is computed with the substraction of $l_{ij}$ from b and the right value is computed with the summation of $r_{ij}$ with right (c).

The set-up parameters are given in Table 6.9 for three methods.

First method is defined as classical Fuzzy ID3 Induction Process (FID3). In this approach, fuzzification is done by using fuzzy c-means algorithm. Second method indicates Fuzzy ID3 Algorithm Based on Linguistic Data by using WABL Defuzzification. Fuzzy c-means (FCM) algorithm is used again in order to get membership degrees.

Then, Fuzzy ID3-L-WABL is applied in order to achieve the fuzzy decision tree. Hence, third method works directly on fuzzy data. It uses FkM-F algorithm to perform the fuzzification phase. Then, the non-parametric reasoning methods (given in Table 3.1) are applied to the six data sets for both three approaches to examine the performances of the classification.

The accuracy rates are obtained for the three methods. The performance results are given in Table 6.10, Table 6.11 and Table 6.12, respectively.

It is seen that Wdbc has the highest performance result with 91.57%, which is obtained from tha adaptation of FID3-L-WABL and Bounded Product/Sum reasoning among the other approaches. Iris data set has best accuracy rate with 95.33% FID3-L-WABL and Zadeh reasoning among all reasoning methods.

Table 6.9 Set up of the methods parameters

| Methods | Algorithm | Parameters |
|---|---|---|
| **1** | Fuzzy ID3 Induction Process (FID3) | FCM:c=3 classes<br>$\theta_r = 0.75$<br>max depth=11<br>Non-parametric operatos given in Table 3.1. |
| **2** | Fuzzy ID3 Algorithm Based on Linguistic Data By Using WABL Defuzzification Method (FID3-L-WABL) | FCM: c=3 classes<br>WABL:<br>$k = 0; s = 1.0;$<br>$c_L = c_K = 0.50$<br>FID3: $\theta_r = 0.75$<br>max depth=11<br>Reasoning:Non-parametric operatos given in Table 3.1. |
| **3** | Fuzzy ID3 Algorithm for L-R Fuzzy Data (Fuzzy ID3-LR) | FkM: c=3 classes<br>L-R fuzzy data used for Method 2.<br>FID3: $\theta_r = 0.75$<br>max depth=11<br>Non-parametric operatos given in Table 3.1. |

Moreover, while Ring data set has the best performance with 74.88% FID3-L-WABL and Algebraic Product/Sum reasoning, it has the least performance with 49.64% FID3-LR and Bounded Product/Sum reasoning.

Table 6.10 Accuracy rates (%) obtained from FID3

| Datasets | Zadeh | Algebraic Product/Sum | Bounded Product/Sum | Nonparametric Hamacher ($\lambda = 0$) |
|---|---|---|---|---|
| Wdbc | 90.51 | 91.04 | **91.39** | 84.18 |
| Iris | **94.00** | **94.00** | **94.00** | **94.00** |
| Pima | 73.56 | **74.86** | 70.70 | 72.78 |
| Ring | 73.00 | **74.55** | 63.49 | 68.99 |
| Sonar | **73.08** | 63.11 | 71.64 | 58.72 |
| Phoneme | **73.67** | 73.59 | 73.33 | 73.17 |

Also, *Sonar* has 77.89% accuracy rate with FID3-LR and Algebraic Product/Sum reasoning. *Phoneme* has the highest accuracy rate with 74.91% for FID3-L-WABL and Algebraic Product/Sum reasoning.

Table 6.11 Accuracy rates (%) obtained from FID3-L-WABL

| Datasets | Zadeh | Algebraic Product/Sum | Bounded Product/Sum | Nonparametric Hamacher $(\lambda = 0)$ |
|---|---|---|---|---|
| Wdbc | 90.87 | 79.82 | **91.57** | 73.83 |
| Iris | **95.33** | 94.67 | 94.67 | 94.67 |
| Pima | 72.91 | **73.82** | 69.66 | 73.56 |
| Ring | 74.14 | **74.88** | 64.61 | 68.88 |
| Sonar | 73.10 | **74.08** | 73.09 | 64.41 |
| Phoneme | 74.70 | **74.91** | 73.58 | 74.48 |

The experimental study is encouraged by the statistical tests. The Aligned Friedman is performed to check whether there are statistical differences among the performances of reasoning methods. Also, Wilcoxon signed rank test is applied to make the comparison among the methods. It is aimed to show that the performances of the proposed approaches, FID3-L-WABL and FID3-LR, have as good performance as classical Fuzzy ID3 in general.

Table 6.12 Accuracy rates (%) obtained from FID3-LR

| Datasets | Zadeh | Algebraic Product/Sum | Bounded Product/Sum | Nonparametric Hamacher $(\lambda = 0)$ |
|---|---|---|---|---|
| Wdbc | 90.16 | 90.86 | 90.86 | **91.04** |
| Iris | **55.33** | 33.99 | 33.33 | 50.67 |
| Pima | 65.75 | 65.75 | 65.10 | **66.01** |
| Ring | **51.08** | 49.76 | 49.64 | 49.76 |
| Sonar | 74.03 | **77.89** | 57.75 | 74.99 |
| Phoneme | **70.65** | **70.65** | **70.65** | **70.65** |

According to the result of the Aligned Friedman test, it is seen that *p*-value equals to 0.008 as given in Table 6.13. It shows that there are significant differences among the results ($\alpha$=0.05). Then, Wilcoxon signed rank test is applied to the approaches in order to test the comparison among the reasoning methods on three induction approaches as FID3-L-WABL, FID3-LR, and classical FID3. The results of this test for six well-known selected data sets are given in Table 6.14. (L_ implies FID3-L-WABL, L2 implies FID3_LR).

Table 6.13 Friedman aligned ranks results to test the performance of reasoning methods

| Algorithm | Rank | Algorithm | Rank | Algorithm | Rank | Friedman aligned ranks | |
|---|---|---|---|---|---|---|---|
| | | | | | | **Total N** | 6 |
| **Zadeh** | 7.50 | **Zadeh_L** | 9.50 | **Zadeh_L2** | 4.33 | **Test Statistic** | 25.30 |
| **Algebraic Product/Sum** | 8.33 | **Algebraic Product/Sum_L** | 9.50 | **Algebraic Product/Sum_L2** | 4.67 | | |
| **Bounded Product/Sum** | 6.58 | **Bounded Product/Sum_L** | 7.83 | **Bounded Product/Sum_L2** | 2.17 | **Asymptotic Sig. (2 sided test)** | 0.008 |
| **Non Parametric Hamacher ($\lambda = 0$)** | 5.25 | **Non Parametric Hamacher_L ($\lambda = 0$)** | 6.92 | **Non Parametric Hamacher_L2 ($\lambda = 0$)** | 5.42 | | |

It is seen that there is no significant difference between the results of Zadeh operator applied with FID3-L-WABL and FID3-LR with p value, 0.075. In addition, the results of Zadeh operator applied with classical FID3 and FID3-LR with p value, 0.075 shows that there is no significant difference.

Hence, it is concluded that there is no significant difference for the pairs of Algebraic Product/Sum operator results obtained from FID3-L-WABL and FID3-LR, classical FID3 and FID3-LR. P values handled from the tests are given as 0.249 and 0.173, respectively. Non-Parametric Hamacher operator's results got from FID3-L-WABL and FID3-LR, classical FID3 and FID3-LR, respectively. These two pairs' comparisons have p-values as 0.463.

Yet, it is observed that Bounded Product/Sum operator's results got from FID3-L-WABL and FID3-LR, classical FID3 and FID3-LR are significant difference with p-value as 0.028, respectively. It is observed that FID3-LR achieved better results with Bounded Product/Sum reasoning operator than FID3-L-WABL and classical FID3 approaches.

It is seen that there is significant difference between the results of Non Parametric Hamacher and Bounded Product/Sum produced with FID3-LR (p-value, 0.043). It is seen that Bounded Product/Sum operator has better results with p value, 0.022. There is significant difference between the results obtained from Non Parametric Hamacher and Algebraic Product/Sum operators with FID3-L-WABL (p value, 0.043).

The results of Algebraic Product/Sum operator with FID3-L-WABL has better results than the results of Non Parametric Hamacher operator with FID3-L-WABL. Classical FID3 with Non Parametric Hamacher operators achieves better results than classical FID3 with Zadeh and Bounded Product/Sum operators with p value, 0.022.

As a summary, it is observed that there are no differences between the results of classical FID3, FID3-L-WABL and FID3-LR in general situation. Linguistic approaches have good performance no fewer than numerical approach. Yet, Fuzzy

ID3-LR has better performance than FuzzyID3-L-WABl and classical FuzzyID3 for Bounded Product/Sum operator’s results.

Table 6.14 Wilcoxon signed rank tests on different threshold for general situation.

| Comparison | Sum of Negative Ranks | Sum of Positive Ranks | Hypothesis | Z test statistic | Exact sig. (1-tailed) p value | Decision | Comparison | Sum of Negative Ranks | Sum of Positive Ranks | Hypothesis | Z test statistic | Exact sig. (1-tailed) p value | Decision |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Algebraic Product/Sum_L2 vs Zadeh_L2 ($M_D=M_{ZADEH\_L2}-M_{ALGEBRAIC/RODUCT-SUM\_L2}$) | 6.00 | 4.00 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.365 | 0.715 | It is not rejected | Non Parametric Hamacher vs.Zadeh ($M_D=M_{ZADEH}-M_{NON\ PARAMETRIC\ HAMACHER}$) | 15.0 | 0.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -2.023 | 0.043 | **It is rejected** |
| Bounded Product/Sum_L2 vs. Zadeh_L2 ($M_D=M_{ZADEH\_L2}-M_{BOUNDED/RODUCT-SUM\_L2}$) | 13.00 | 2.00 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -1.483 | 0.138 | It is not rejected | Bounded Product/Sum vs. Algebraic Product/Sum ($M_D=M_{ALGEBRAIC/RODUCT-SUM}-M_{BOUNDED/RODUCT-SUM}$) | 9.0 | 6.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.405 | 0.686 | It is not rejected |
| Non Parametric Hamacher_L2 vs. Zadeh_L2 ($M_D=M_{ZADEH\_L2}-M_{NON\ PARAMETRIC\ HAMACHER\_L2}$) | 9.0 | 6.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.405 | 0.686 | It is not rejected | Non Parametric Hamacher vs. Algebraic Product/Sum ($M_D=M_{ALGEBRAIC/PRODUCT}-M_{NON\ PARAMETRIC\ HAMACHER}$) | 15.0 | 0.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -2.023 | 0.043 | **It is rejected** |
| Bounded Product/Sum_L2 vs Algebraic Product/Sum _L2 ($M_D=M_{ALGEBRAIC/RODUCT-SUM\_L2}-M_{BOUNDED/RODUCT-SUM\_L2}$) | 10.0 | 0.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -1.826 | 0.068 | It is not rejected | Non Parametric Hamacher vs. Bounded Product/Sum ($M_D=M_{BOUNDED/PRODUCT}-M_{NON\ PARAMETRIC\ HAMACHER}$) | 10.0 | 5.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.674 | 0.500 | It is not rejected |
| Non Parametric Hamacher_L2 vs. Algebraic Product/Sum_L2 ($M_D=M_{ALGEBRAIC/PRODUCT\_L2}-M_{NON\ PARAMETRIC\ HAMACHER\_L2}$) | 3.0 | 7.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.730 | 0.465 | It is not rejected | Zadeh_L vs. Zadeh_L2 ($M_D=M_{ZADEH\_L2}-M_{ZADEH\_L}$) | 2.0 | 19.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | 1.782 | 0.075 | It is not rejected |
| Non Parametric Hamacher_L2 vs.Bounded Product/Sum_L2 ($M_D=M_{BOUNDED/PRODUCT\_L2}-M_{NON\ PARAMETRIC\ HAMACHER\_L2}$) | 0.0 | 15.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -2.023 | 0.043 | **It is rejected** | Zadeh vs. Zadeh_L2 ($M_D=M_{ZADEH\_L2}-M_{ZADEH}$) | 2.0 | 19.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -1.782 | 0.075 | It is not rejected |
| Algebraic Product/Sum_L vs. Zadeh_L ($M_D=M_{ZADEH\_L}-M_{ALGEBRAIC/RODUCT-SUM\_L}$) | 8.0 | 13.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.524 | 0.600 | It is not rejected | Algebraic Product/Sum_L vs Algebraic Product/Sum_L2 ($M_D=M_{ALGEBRAIC\ PRODUCT/SUM\_L2}-M_{ALGEBRAIC\ PRODUCT/SUM\_L}$) | 5.0 | 16.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -1.153 | 0.249 | It is not rejected |
| Bounded Product/Sum_L vs. zadeh_L ($M_D=M_{ZADEH\_L}-M_{BOUNDED/RODUCT-SUM\_L}$) | 18.0 | 3.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -1.572 | 0.116 | It is not rejected | Algebraic Product/Sum vs Algebraic Product/Sum_L2 ($M_D=M_{ALGEBRAIC\ PRODUCT/SUM\_L2}-M_{ALGEBRAIC\ PRODUCT/SUM}$) | 4.0 | 17.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -1.363 | 0.173 | It is not rejected |
| Non Parametric Hamacher_L vs. Zadeh_L ($M_D=M_{ZADEH\_L}-M_{NON\ PARAMETRIC\ HAMACHER\_L}$) | 19.0 | 2.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -1.782 | 0.075 | It is not rejected | Bounded Product/Sum_L vs. Bounded Product/Sum_L2 ($M_D=M_{BOUNDED\ PRODUCT/SUM\_L2}-M_{BOUNDED\ PRODUCT/SUM\_L}$) | 0.0 | 21.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -2.201 | 0.028 | **It is rejected** |
| Bounded Product/Sum_L vs. Algebraic Product/Sum_L ($M_D=M_{ALGEBRAIC/RODUCT-SUM\_L}-M_{BOUNDED/RODUCT-SUM\_L}$) | 10.0 | 5.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.674 | 0.500 | It is not rejected | Bounded Product/Sum vs. Bounded Product/Sum_L2 ($M_D=M_{BOUNDED\ PRODUCT/SUM\_L2}-M_{BOUNDED\ PRODUCT/SUM}$) | 0.0 | 21.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -2.201 | 0.028 | **It is rejected** |
| Non Parametric Hamacher_L vs. Algebraic Product/Sum_L ($M_D=M_{ALGEBRAIC/PRODUCT\_L}-M_{NON\ PARAMETRIC\ HAMACHER\_L}$) | 15.0 | 0.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -2.023 | 0.043 | **It is rejected** | Non Parametric Hamacher_L vs. Non Parametric Hamacher_L2 ($M_D=M_{NPHAMACHER\_L2}-M_{NPHAMACHER\_L}$) | 7.0 | 14.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.734 | 0.463 | It is not rejected |
| Non Parametric Hamacher_L vs. Bounded Product/Sum_L $M_D=M_{BOUNDED/PRODUCT\_L}-M_{NON\ PARAMETRIC\ HAMACHER\_L}$ | 9.0 | 6.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.405 | 0.686 | It is not rejected | Non Parametric Hamacher vs. Non Parametric Hamacher_L2 ($M_D=M_{NPHAMACHER\_L2}-M_{NPHAMACHER}$) | 7.0 | 14.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.734 | 0.463 | It is not rejected |
| Algebraic Product/Sum vs. Zadeh ($M_D=M_{ZADEH}-M_{ALGEBRAIC/RODUCT-SUM}$) | 6.0 | 9.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -0.405 | 0.686 | It is not rejected | Bounded Product/Sum vs. Zadeh ($M_D=M_{ZADEH}-M_{BOUNDED/RODUCT-SUM}$) | 13.0 | 2.0 | $H_0: M_D = 0$<br>$H_1: M_D \neq 0$ | -1.483 | 0.138 | It is not rejected |

### *6.3.4.2 Study of the behaviour of different t-operators on different threshold values for classical Fuzzy ID3, Fuzzy ID3-L-WABL, and Fuzzy ID3–LR*

In this study, classical Fuzzy ID3 (FID3), Fuzzy ID3-L-WABL, (FID3-L-WABL) and Fuzzy-LR with different threshold values between the range 0.60-0.90 are performed for the induction of fuzzy decision tree. Then, four non-parametric T-operators are worked on reasoning process. It is aimed to show that the performance of different T-operators on these three approaches.

*Iris data set*

Iris data set performance results for three approaches are given in Table 6.15, Table 6.16, and Table 6.17, respectively.

While classical FID3 approach with 0.82 threshold is performed for the induction process on Iris data set, Algebraic Product/Sum operator has the higest accuracy rate with 96.67% among the other non-parametric operators.

While FID3-L-WABL approach with 0.82 threshold is performed for the induction process on Iris data set, Algebraic Product/Sum operator has the higest accuracy rate with 96.67% among the other non-parametric operators.

While FID3-LR approach with 0.60 threshold is performed for the induction process on Iris data set, Non parametric operator has the higest accuracy rate with 76.00% among the other non-parametric operators. FID3-LR approach has the lowest accuracy rates among non-parametric T-operators.

Table 6.15 Iris data set performance results (%) different t-operators on different threshold results for classical Fuzzy ID3 (FID3).

| **Threshold** | **Zadeh** | **Algebraic Product/Sum** | **Bounded Product/Sum** | **NP_ Hamacher** |
|---|---|---|---|---|
| 0.60 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.61 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.62 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.63 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.64 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.65 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.66 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.67 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.68 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.69 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.70 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.71 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.72 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.73 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.74 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.75 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.76 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.77 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.78 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.79 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.80 | 94.00 | 94.00 | 94.00 | 94.00 |
| 0.81 | **95.33** | 94.67 | 94.67 | **94.67** |
| 0.82 | 94.67 | **96.67** | **95.33** | 94.00 |
| 0.83 | **95.33** | 93.33 | 93.33 | 88.00 |
| 0.84 | 94.67 | 95.33 | 94.67 | 88.00 |
| 0.85 | 94.67 | 95.33 | 94.67 | 88.00 |
| 0.86 | 94.67 | 95.33 | 94.67 | 88.00 |
| 0.87 | 94.67 | 95.33 | 94.67 | 88.00 |
| 0.88 | 94.67 | 95.33 | 94.67 | 88.00 |
| 0.89 | 94.67 | 95.33 | 94.67 | 88.00 |
| 0.90 | 94.67 | 95.33 | 94.67 | 88.00 |

Table 6.16 Iris data set performance results (%) different t-operators on different threshold results for Fuzzy ID3-L-WABL (FID3-L-WABL).

| **Threshold** | **Zadeh_L** | **Algebraic Product/Sum_L** | **Bounded Product/Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.61 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.62 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.63 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.64 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.65 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.66 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.67 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.68 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.69 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.70 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.71 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.72 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.73 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.74 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.75 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.76 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.77 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.78 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.79 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.80 | 95.33 | 94.67 | 94.67 | 94.67 |
| 0.81 | 94.00 | 96.00 | 95.33 | 94.00 |
| 0.82 | 96.00 | **96.67** | 95.33 | 91.33 |
| 0.83 | 96.00 | 95.33 | 93.33 | 88.00 |
| 0.84 | 96.00 | 95.33 | 93.33 | 88.00 |
| 0.85 | 96.00 | 95.33 | 93.33 | 88.00 |
| 0.86 | 96.00 | 95.33 | 93.33 | 88.00 |
| 0.87 | 96.00 | 95.33 | 93.33 | 88.00 |
| 0.88 | 96.00 | 95.33 | 93.33 | 88.00 |
| 0.89 | 96.00 | 96.00 | 93.33 | 88.00 |
| 0.90 | 96.00 | 96.00 | 93.33 | 88.00 |

Table 6.17 Iris data set performance results (%) different t-operators on different threshold results for FuzzyID3-LR (FID3-LR) approach.

| **Threshold** | **Zadeh_L** | **Algebraic Product/Sum_L** | **Bounded Product/Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 70.67 | 71.33 | 33.99 | **76.00** |
| 0.61 | 67.33 | 65.33 | 33.33 | 74.02 |
| 0.62 | 65.33 | 59.33 | 33.33 | 70.00 |
| 0.63 | 65.33 | 59.33 | 33.33 | 70.00 |
| 0.64 | 65.33 | 59.33 | 33.33 | 70.00 |
| 0.65 | 64.66 | 55.99 | 33.33 | 68.66 |
| 0.66 | 63.33 | 49.33 | 33.33 | 66.66 |
| 0.67 | 63.33 | 49.33 | 33.33 | 66.66 |
| 0.68 | 63.33 | 49.33 | 33.33 | 66.66 |
| 0.69 | 62.67 | 45.33 | 33.33 | 65.99 |
| 0.70 | 62.66 | 42.00 | 33.33 | 63.33 |
| 0.71 | 61.33 | 36.66 | 33.33 | 60.00 |
| 0.72 | 61.33 | 36.66 | 33.33 | 57.33 |
| 0.73 | 55.33 | 34.66 | 33.33 | 51.33 |
| 0.74 | 55.33 | 33.99 | 33.33 | 50.67 |
| 0.75 | 55.33 | 33.99 | 33.33 | 50.67 |
| 0.76 | 55.33 | 33.99 | 33.33 | 50.67 |
| 0.77 | 55.33 | 33.99 | 33.33 | 50.67 |
| 0.78 | 49.33 | 33.33 | 33.33 | 47.33 |
| 0.79 | 49.33 | 33.33 | 33.33 | 47.33 |
| 0.80 | 38.67 | 33.33 | 33.33 | 39.33 |
| 0.81 | 38.67 | 33.33 | 33.33 | 39.33 |
| 0.82 | 33.33 | 33.33 | 33.33 | 33.33 |
| 0.83 | 33.33 | 33.33 | 33.33 | 33.33 |
| 0.84 | 33.33 | 33.33 | 33.33 | 33.33 |
| 0.85 | 33.33 | 33.33 | 33.33 | 33.33 |
| 0.86 | 33.33 | 33.33 | 33.33 | 33.33 |
| 0.87 | 33.33 | 33.33 | 33.33 | 33.33 |
| 0.88 | 33.33 | 33.33 | 33.33 | 33.33 |
| 0.89 | 33.33 | 33.33 | 33.33 | 33.33 |
| 0.90 | 33.33 | 33.33 | 33.33 | 33.33 |

The Friedman aligned ranks as a non-parametric statistical procedure is applied to detect statistical differences among a group of results on 31 thresholds value between 0.60-0.90. for three approaches. The test results performed on both approaches have significant p-value, 0.000 as given in Table 6.18. It is seen that there are significant differences among the results ($\alpha$=0.05).

Wilcoxon signed rank tests are applied into the Iris data set to test the comparison among the reasoning methods. It is seen that the results of Non Parametric Hamacher operator are better than the results of Zadeh operator produced with classical FID3 (p-value, 0.001), significantly. Bounded Product/Sum operator has worse performance than Algebraic Product/Sum operator with classical FID3 (p value, 0.004).

The performance of Algebraic Product/Sum operator and Bounded Product/Sum operator is better than the performance of Non Parametric Hamacher operator with classical FID3 (p value, 0.002). The performance of Zadeh operator is higher than Algebraic Product/Sum operator, Bounded Product/Sum operator, and Non parametric Hamacher operator with FID3-L-WABL's performances (p value, 0.000). Bounded Product/Sum operator, and Non parametric Hamacher operator give better results than Algebraic Product/Sum operator with FID3-L-WABL (p value, 0.000). Hence, Non parametric Hamacher operator with FID3-L-WABL works better than Bounded Product/Sum results.

Table 6.18 Friedman aligned ranks test for Iris data set.

<table>
<tr><th rowspan="2">Algorithm</th><th rowspan="2">Rank</th><th rowspan="2">Algorithm</th><th rowspan="2">Rank</th><th rowspan="2">Algorithm</th><th rowspan="2">Rank</th><th colspan="2">Friedman aligned ranks</th></tr>
<tr><th>Total N</th><td>31</td></tr>
<tr><td>Zadeh</td><td>7.23</td><td>Zadeh_L</td><td>11.69</td><td>Zadeh_L2</td><td>3.11</td><td rowspan="2">Test Statistic (df)</td><td rowspan="2">303.9 (11)</td></tr>
<tr><td>Algebraic Product/Sum</td><td>7.63</td><td>Algebraic Product/Sum_L</td><td>10.31</td><td>Algebraic Product/Sum_L2</td><td>2.11</td></tr>
<tr><td>Bounded Product/Sum</td><td>7.11</td><td>Bounded Product/Sum_L</td><td>9.23</td><td>Bounded Product/Sum_L2</td><td>1.50</td><td rowspan="2">Asymptotic Sig. (2 sided test)</td><td rowspan="2">0.000</td></tr>
<tr><td>Non Parametric Hamacher ($\lambda = 0$)</td><td>6.27</td><td>Non Parametric Hamacher_L ($\lambda = 0$)</td><td>8.53</td><td>Non Parametric Hamacher_L2 ($\lambda = 0$)</td><td>3.27</td></tr>
</table>

It is observed that the performance of Algebraic Product/Sum operator, Bounded Product/Sum operator is higher than Zadeh operator with FID3-LR's performance (p value, 0.000) as FID3-L-WABL approach. Yet, Non parametric Hamacher operator with FID3-LR has better perfomance than Zadeh operator with FID3-LR (p value, 0.000).

Bounded Product/Sum operator has a worse performance than Algebraic Product/Sum operator with FID-LR (p value, 0.000). The performance of Non Parametric Hamacher operator is better than Algebraic Product/Sum and Bounded Product/Sum operators' performances with FID3-LR (p value, 0.000).

Zadeh, Bounded Product/Sum, and Algebraic Product/Sum with FID3-L-WABL's performances are better than Zadeh operator with FID3-LR performance (p value, 0.000). Zadeh, Bounded Product/Sum, Algebraic Product/Sum, and Non Parametric Hamacher operator with FID3-L-WABL's performances are better than Algebraic Product/Sum's with FID3-LR's peformance (p value, 0.000). In a similar manner, Zadeh, Bounded Product/Sum, Algebraic Product/Sum, and Non Parametric Hamacher operator with FID3-L-WABL's performances are better than Bounded Product/Sum with FID3-LR's peformance (p value, 0.000). Zadeh, Bounded Product/Sum, Algebraic Product/Sum, and Non Parametric Hamacher operator with FID3-L-WABL's performances are also better than Non Parametric Hamacher operator with FID3-LR's peformance (p value, 0.000).

It is observed that the performance of Zadeh operator with FID3-L-WABL is better than Zadeh operator, Algebraic Product/Sum, Bounded Product/Sum, and Non parametric Hamacher with classical FID3. Algebraic Product/Sum operator with FID3-L-WABL has also better results than Zadeh operator, Algebraic Product/Sum, Bounded Product/Sum, and Non parametric Hamacher with classical FID3. Hence, Bounded Product/Sum with FID3L-WABL's performance does not have any significant difference than Zadeh (p value, 0.372), Algebraic Product/Sum (p value,0.454), Bounded Product/Sum operator (p value, 0.454) with classical FID3's performances. Non-Parametric with FID3-L-WABL has better performance than it.

Zadeh, Algebraic Product/Sum, Bounded Product/Sum, and Non parametric Hamacher operators with classical FID3 have better performance than Zadeh operator with FID3-LR (p value, 0). Zadeh, Algebraic Product/Sum, Bounded Product/Sum, and Non parametric Hamacher operators with classical FID3 have better performance than Algebraic Product/Sum operator with FID3-LR (p value, 0).

Zadeh, Algebraic Product/Sum, Bounded Product/Sum, and Non parametric Hamacher operators with classical FID3 have better performance than Bounded Product/Sum operator with FID3-LR (p value, 0). Zadeh, Algebraic Product/Sum, Bounded Product/Sum, and Non parametric Hamacher operators with classical FID3 have better performance than Non parametric Hamacher operator with FID3-LR (p value, 0).

*Phoneme data set*

Phoneme data set performance results for three approaches are given in Table 6.19, Table 6.20, and Table 6.21, respectively.

While classical FID3 approach with the threshold range 0.60-0.63 is performed for the induction process on Phoneme data set, Non Parametric Hamacher operator has the higest accuracy rate with 75.22% among the other non-parametric operators.

While FID3-L-WABL approach in (0.72-0.74) threshold ranges is performed for the induction process on Iris data set, Algebraic Product/Sum operator has the higest accuracy rate with 75.54% among the other non-parametric operators.

While FID3-LR approach with 0.60 threshold is performed for the induction process on Phoneme data set, Non parametric Hamacher operator has the higest accuracy rate with 77.01% among the other non-parametric operators.

Table 6.19 Phoneme data set performance results (%) different t-operators on different threshold results for classical FuzzyID3(FID3)

| **Threshold** | **Zadeh** | **Algebraic Product/ Sum** | **Bounded Product/ Sum** | **NP_ Hamacher** |
|---|---|---|---|---|
| 0.60 | 75.09 | 75.26 | 74.41 | **75.22** |
| 0.61 | 75.09 | 75.26 | 74.41 | **75.22** |
| 0.62 | 75.09 | 75.26 | 74.41 | **75.22** |
| 0.63 | 75.09 | 75.26 | 74.41 | **75.22** |
| 0.64 | 74.48 | 74.63 | 74.15 | 73.98 |
| 0.65 | 74.48 | 74.46 | 74.07 | 73.89 |
| 0.66 | 74.50 | 74.48 | 74.07 | 73.72 |
| 0.67 | 74.50 | 74.67 | 74.09 | 73.74 |
| 0.68 | 74.50 | 74.67 | 74.09 | 73.74 |
| 0.69 | 74.46 | 74.57 | 74.09 | 73.65 |
| 0.70 | 74.32 | 74.43 | 74.11 | 73.70 |
| 0.71 | 74.46 | 74.48 | 74.13 | 73.79 |
| 0.72 | 74.50 | 74.50 | 74.15 | 73.39 |
| 0.73 | 74.52 | 74.48 | 74.15 | 73.89 |
| 0.74 | 74.52 | 74.48 | 74.15 | 73.89 |
| 0.75 | 73.67 | 73.59 | 73.33 | 73.37 |
| 0.76 | 71.93 | 71.84 | 71.89 | 71.65 |
| 0.77 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.78 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.79 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.80 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.81 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.82 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.83 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.84 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.85 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.86 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.87 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.88 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.89 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.90 | 70.65 | 70.65 | 70.65 | 70.65 |

Table 6.20 Phoneme data set performance results (%) different t-operators on different threshold results for FuzzyID3-L-WABL (FID3-L_WABL).

| **Threshold** | **Zadeh_L** | **Algebraic Product/ Sum_L** | **Bounded Product/ Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 74.49 | 75.22 | 74.37 | **75.02** |
| 0.61 | 74.49 | 75.22 | 74.37 | **75.02** |
| 0.62 | 74.49 | 75.22 | 74.37 | **75.02** |
| 0.63 | 74.49 | 75.22 | **74.41** | 74.94 |
| 0.64 | 74.30 | 74.30 | 74.15 | 73.98 |
| 0.65 | 74.35 | 74.35 | 74.06 | 73.72 |
| 0.66 | 74.39 | 74.46 | 74.07 | 73.38 |
| 0.67 | 74.42 | 74.54 | 74.07 | 73.82 |
| 0.68 | 74.43 | 74.54 | 74.07 | 73.82 |
| 0.69 | 74.56 | 74.65 | 74.09 | 73.91 |
| 0.70 | 74.89 | 74.98 | 74.13 | 74.20 |
| 0.71 | 75.09 | 75.27 | 74.13 | 74.57 |
| 0.72 | **75.28** | **75.54** | 74.15 | 74.79 |
| 0.73 | **75.28** | **75.54** | 74.17 | 74.79 |
| 0.74 | **75.28** | **75.54** | 74.17 | 74.80 |
| 0.75 | 74.70 | 74.91 | 73.58 | 73.48 |
| 0.76 | 74.17 | 74.41 | 73.24 | 73.87 |
| 0.77 | 74.07 | 74.28 | 72.96 | 73.89 |
| 0.78 | 74.07 | 74.28 | 72.96 | 73.89 |
| 0.79 | 74.04 | 74.24 | 72.96 | 73.87 |
| 0.80 | 72.39 | 72.43 | 71.80 | 72.26 |
| 0.81 | 71.11 | 71.21 | 70.95 | 71.11 |
| 0.82 | 71.11 | 71.21 | 70.95 | 71.11 |
| 0.83 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.84 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.85 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.86 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.87 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.88 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.89 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.90 | 70.65 | 70.65 | 70.65 | 70.65 |

Table 6.21 Phoneme data set performance results (%) different t-operators on different threshold results for Fuzzy-LR (FID3-LR).

| **Threshold** | **Zadeh_L2** | **Algebraic Product/Sum_L2** | **Bounded Product/Sum_L2** | **NP_ Hamacher_L2** |
|---|---|---|---|---|
| 0.60 | **76.02** | **75.46** | **70.97** | **77.01** |
| 0.61 | 75.81 | 75.15 | **70.97** | 76.04 |
| 0.62 | 75.41 | 74.17 | **70.97** | 74.83 |
| 0.63 | 75.41 | 74.17 | **70.97** | 74.83 |
| 0.64 | 74.47 | 73.24 | 70.65 | 73.91 |
| 0.65 | 74.46 | 73.24 | 70.65 | 73.91 |
| 0.66 | 74.46 | 73.24 | 70.65 | 73.91 |
| 0.67 | 74.46 | 73.24 | 70.65 | 73.91 |
| 0.68 | 74.46 | 73.24 | 70.65 | 73.91 |
| 0.69 | 73.29 | 72.50 | 70.65 | 73.01 |
| 0.70 | 73.29 | 72.50 | 70.65 | 73.01 |
| 0.71 | 72.50 | 71.85 | 70.65 | 72.21 |
| 0.72 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.73 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.74 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.75 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.76 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.77 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.78 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.79 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.80 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.81 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.82 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.83 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.84 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.85 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.86 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.87 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.88 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.89 | 70.65 | 70.65 | 70.65 | 70.65 |
| 0.90 | 70.65 | 70.65 | 70.65 | 70.65 |

The Friedman aligned ranks as a non-parametric statistical procedure is applied to detect statistical differences among a group of results on 31 thresholds value between 0.60-0.90. for three approaches. The test results performed on both approaches have significant p-value, 0.000 as given in Table 6.22. It is seen that there are significant differences among the results ($\alpha$=0.05).

Table 6.22 Friedman aligned ranks test for Phoneme data set.

| Algorithm | Rank | Algorithm | Rank | Algorithm | Rank | Friedman aligned ranks | |
|---|---|---|---|---|---|---|---|
| | | | | | | **Total N** | 31 |
| **Zadeh** | 7.50 | **Zadeh_L** | 8.48 | **Zadeh_L2** | **6.37** | **Test Statistic (df)** | 131.722 (11) |
| **Algebraic Product/Sum** | 7.97 | **Algebraic Product/Sum_L** | 9.79 | **Algebraic Product/Sum_L2** | 4.15 | | |
| **Bounded Product/Sum** | 5.71 | **Bounded Product/Sum_L** | 6.65 | **Bounded Product/Sum_L2** | 3.34 | **Asymptotic Sig. (2 sided test)** | 0.000 |
| **Non Parametric Hamacher ($\lambda = 0$)** | 5.60 | **Non Parametric Hamacher_L ($\lambda = 0$)** | 7.31 | **Non Parametric Hamacher_L2 ($\lambda = 0$)** | 5.15 | | |

Then, Wilcoxon signed rank tests are applied into the phoneme data set to test the comparison among the reasoning methods. It is seen that the results of Algebraic Product/Sum operator are better than the results of Zadeh operator produced with classical FID3 (p-value, 0.014), significantly. Bounded Product/Sum and Non Parametric Hamacher operators with classical FID3 have worse performance than Zadeh operator with classical FID3 (p value, 0.000).

The performance of Algebraic Product/Sum operator is better than Bounded Product/Sum operator with classical FID3 (p value, 0.000). Non parametric Hamacher operator with FID3-L-WABL’s performance is better than Algebraic Product/Sum operator with FID3-L-WABL’s performance (p value, 0.000).

Bounded Product/Sum operator and Algebraic Product/Sum operator with FID3-L-WABL’s works better than Zadeh’s operator with FID3-L-WABL. Non parametric Hamacher operator with FID3-L-WABL’s performance is better than Zadeh operator with FID3-L-WABL’s performance (p value, 0.000). Algebraic Product/Sum operator with FID3-L-WABL’s performance is better than Non parametric Hamacher operator with FID3-L-WABL’s performance. Nonetheless, Non parametric

Hamacher operator with FID3-L-WABL works better than Bounded Product/Sum operator (p value, 0.000).

The performance of Non Parametric Zadeh operator is better than Algebraic Product/Sum, Bounded Product/Sum, Non parametric Hamacher operators' performances with with FID3-LR (p value, 0.000). Algebraic Product/Sum works better than Bounded Product/Sum with FID3-LR (p value, 0.000). Non parametric Hamacher operators with FID3-LR works better than Algebraic Product/Sum and Bounded Product/Sum with FID3-LR.

Zadeh operator (p value, 0.000), Algebraic Product/Sum operator (p value, 0.000), Bounded Product/Sum operator (p value, 0.000) and Non parametric Hamacher operator (p value, 0.000) with FID3-L-WABL have better performance than Zadeh operator with FID3-LR.

Zadeh operator (p value, 0.000), Algebraic Product/Sum operator (p value, 0.000), Bounded Product/Sum operator (p value, 0.000) and Non parametric Hamacher operator (p value, 0.000) with FID3-L-WABL have better performance than Algebraic Product/Sum operator with FID3-LR.

Zadeh operator (p value, 0.000), Algebraic Product/Sum operator (p value, 0.000), Bounded Product/Sum operator (p value, 0.000) and Non parametric Hamacher operator (p value, 0.000) with FID3-L-WABL have better performance than Bounded Product/Sum operator with FID3-LR.

While Non parametric Hamacher operator (p value, 0.000) with FID3-LR have better results than Algebraic Product/Sum operator (p value, 0.000) and Bounded Product/Sum operator (p value, 0.000) with FID3-L-WABL. Yet, Non parametric Hamacher operator with FID3-LR has worse performance than Hamacher operator with FID3- L-WABL (p value, 0.000).

Zadeh, Bounded Product/Sum, Algebraic Product/Sum and Non Parametric Hamacher operators with classical FID3 performances are worse than Zadeh operator with FID3-L-WABL performance (p value, 0.000). In a similar manner, Zadeh, Bounded Product/Sum, Algebraic Product/Sum and Non Parametric Hamacher operators with classical FID3 performances are worse than Algebraic Product/Sum operator with FID3-L-WABL performance (p value, 0.000).

Bounded Product/Sum with FID3-L-WABl performance is worse than Bounded Product/Sum (p value, 0.019) and Non Parametric Hamacher (p value, 0.013) operators with classical FID3 performances'.

However, Non Parametric Hamacher operator with FID3-L-WABL has better performance than Bounded Product/Sum (p value, 0,001) and Non Parametric Hamacher operators (p value, 0.002) with classical FID3.

It is observed that Zadeh (p value, 0.013) and Algebraic Product/Sum (p value, 0.005) operators with classical FID3 performances' are better than Zadeh with Fuzzy ID3-LR performance. Zadeh (p value, 0.000), Bounded Product/Sum(p value, 0.000), Algebraic Product/Sum(p value, 0.000), and Non Parametric Hamacher (p value, 0.000) operators with classical FID3 performances are better than Algebraic Product/Sum with FID3-LR performance. None the less, Bounded Product/Sum (p value, 0.000) with FID3-LR performance is better than Zadeh, Bounded Product/Sum, Algebraic Product/Sum and Non Parametric Hamacher operators with classical FID3 performances (p value, 0).

Non Parametric Hamacher operators with FID3-LR perfoms better than Algebraic Product/Sum (p value, 0.000).

*Pima data set*

Pima data set performance results for three approaches are given in Table 6.23, Table 6.24, and Table 6.25, respectively.

While classical FID3 approach with the threshold 0.60 is performed for the induction process on Pima data set, Zadeh operator has the higest accuracy rate with 76.04% among the other non-parametric operators. While FID3-L-WABL approach with 0.69 threshold is performed for the induction process on Pima data set, Zadeh operator has the higest accuracy rate with 75.64% among the other non-parametric operators. While FID3-LR approach with 0.62 threshold is performed for the induction process on Pima data set, Non parametric Hamacher operator has the higest accuracy rate with 75.52% among the other non-parametric operators.

The Friedman aligned ranks as a non-parametric statistical procedure is applied to detect statistical differences among a group of results on 31 thresholds value between 0.60-0.90. for three approaches. The test results performed on both approaches have significant p-value, 0.000 as given in Table 6.26. It is seen that there are significant differences among the results ($\alpha$=0.05).

Then, Wilcoxon signed rank tests are applied into the pima data set to test the comparison among the reasoning methods. It is seen that the results of Algebraic Product/Sum operator are better than the results of Zadeh operator produced with classical FID3 (p-value, 0.000), significantly.

Bounded Product/Sum and Non Parametric Hamacher operators with classical FID3 have worse performance than Zadeh operator with classical FID3 (p value, 0.000). In a similar manner, Bounded Product/Sum and Non Parametric Hamacher operators with classical FID3 have worse performance than Algebraic Product/Sum operator with classical FID3 (p value, 0.000). On the other hand, Non Parametric Hamacher operators with classical FID3 has better performance than Bounded Product/Sum operator (p value, 0.007).

Table 6.23 Pima data set performance results (%) different t-operators on different threshold results for classical FuzzyID3 (FID3).

| **Threshold** | **Zadeh** | **Algebraic Product/ Sum** | **Bounded Product/Sum** | **NP_ Hamacher** |
|---|---|---|---|---|
| 0.60 | 74.48 | 74.35 | 74.35 | 74.35 |
| 0.61 | 74.48 | 74.35 | 74.35 | 74.35 |
| 0.62 | 74.48 | 74.35 | 74.35 | 74.35 |
| 0.63 | 74.22 | 73.69 | 73.96 | 73.44 |
| 0.64 | 74.35 | 73.70 | 73.83 | 73.05 |
| 0.65 | 75.39 | 75.26 | 74.74 | 73.44 |
| 0.66 | 75.78 | 75.52 | 74.87 | 73.70 |
| 0.67 | **76.04** | 75.39 | 75.13 | 73.31 |
| 0.68 | 75.26 | 74.74 | 74.22 | 71.88 |
| 0.69 | 75.52 | 75.39 | 74.35 | 72.27 |
| 0.70 | 74.74 | 74.47 | 73.44 | 71.35 |
| 0.71 | 74.34 | 74.47 | 72.78 | 71.48 |
| 0.72 | 74.48 | 75.12 | 71.61 | 70.83 |
| 0.73 | 74.34 | 74.60 | 71.61 | 72.52 |
| 0.74 | 74.08 | 74.60 | 71.22 | 72.65 |
| 0.75 | 73.56 | 74.86 | 70.70 | 72.78 |
| 0.76 | 73.56 | 75.64 | 70.18 | 72.78 |
| 0.77 | 73.56 | 75.64 | 70.18 | 72.78 |
| 0.78 | 73.43 | 75.64 | 70.05 | 72.91 |
| 0.79 | 72.91 | 74.99 | 69.92 | 73.17 |
| 0.80 | 73.17 | 74.86 | 70.05 | 73.17 |
| 0.81 | 73.04 | 74.47 | 69.92 | 73.17 |
| 0.82 | 72.52 | 73.69 | 69.40 | 73.04 |
| 0.83 | 72.39 | 73.56 | 69.27 | 72.78 |
| 0.84 | 72.65 | 74.08 | 69.66 | 71.87 |
| 0.85 | 69.79 | 72.00 | 67.32 | 68.48 |
| 0.86 | 68.10 | 70.05 | 66.41 | 67.45 |
| 0.87 | 68.49 | 69.66 | 66.15 | 67.58 |
| 0.88 | 68.36 | 69.79 | 65.76 | 67.32 |
| 0.89 | 68.23 | 69.92 | 65.76 | 67.97 |
| 0.90 | 68.10 | 69.53 | 65.76 | 67.19 |

Table 6.24 Pima data set performance results (%) different t-operators on different threshold results for FuzzyID3-L-WABL (FID3-L-WABL).

| **Threshold** | **Zadeh_L** | **Algebraic Product/Sum_L** | **Bounded Product/ Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 74.22 | 74.22 | 74.35 | 74.35 |
| 0.61 | 74.22 | 74.22 | 74.35 | 74.35 |
| 0.62 | 74.22 | 74.22 | 74.35 | 74.35 |
| 0.63 | 73.57 | 73.18 | 73.70 | 72.66 |
| 0.64 | 73.57 | 73.18 | 73.70 | 72.66 |
| 0.65 | 75.00 | 73.83 | 74.47 | 72.53 |
| 0.66 | 74.87 | 73.57 | 74.48 | 72.27 |
| 0.67 | 74.99 | 73.83 | 73.70 | 72.53 |
| 0.68 | 74.99 | 73.69 | 73.69 | 72.27 |
| 0.69 | **75.64** | 74.74 | 74.21 | 73.44 |
| 0.70 | 74.73 | 73.56 | 72.78 | 72.26 |
| 0.71 | 74.21 | 74.21 | 71.61 | 72.39 |
| 0.72 | 74.08 | 74.34 | 71.22 | 73.57 |
| 0.73 | 74.21 | 73.56 | 70.80 | 73.83 |
| 0.74 | 73.95 | 73.82 | 70.31 | 73.70 |
| 0.75 | 72.91 | 73.82 | 69.66 | 73.56 |
| 0.76 | 72.39 | 73.82 | 69.40 | 73.57 |
| 0.77 | 72.26 | 73.69 | 69.40 | 73.17 |
| 0.78 | 72.26 | 73.95 | 69.27 | 73.57 |
| 0.79 | 71.61 | 73.69 | 69.14 | 72.39 |
| 0.80 | 71.35 | 73.17 | 68.75 | 72.26 |
| 0.81 | 70.96 | 72.78 | 68.10 | 72.00 |
| 0.82 | 70.83 | 72.52 | 67.97 | 72.00 |
| 0.83 | 70.96 | 72.91 | 68.23 | 71.87 |
| 0.84 | 71.48 | 72.65 | 67.84 | 71.87 |
| 0.85 | 69.79 | 71.61 | 67.32 | 69.53 |
| 0.86 | 70.57 | 71.61 | 67.58 | 68.75 |
| 0.87 | 70.57 | 71.61 | 67.58 | 68.75 |
| 0.88 | 69.66 | 71.09 | 67.19 | 68.49 |
| 0.89 | 69.01 | 70.44 | 67.06 | 68.75 |
| 0.90 | 68.23 | 69.78 | 66.28 | 68.23 |

Table 6.25 Pima data set performance results (%) different t-operators on different threshold results for classical Fuzzy-LR (FID3-LR).

| **Threshold** | **Zadeh_L** | **Algebraic Product/Sum_L** | **Bounded Product/Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 75.13 | 73.83 | 65.50 | 74.47 |
| 0.61 | 74.73 | 73.30 | 65.24 | 74.73 |
| 0.62 | 74.47 | 73.56 | 65.10 | **75.52** |
| 0.63 | 74.47 | 72.91 | 65.10 | 75.51 |
| 0.64 | 74.60 | 72.53 | 65.10 | 75.39 |
| 0.65 | 74.08 | 71.21 | 65.10 | 75.12 |
| 0.66 | 74.60 | 71.48 | 65.10 | 75.38 |
| 0.67 | 74.60 | 71.09 | 65.10 | 75.25 |
| 0.68 | 74.34 | 70.08 | 65.10 | 75.25 |
| 0.69 | 73.95 | 70.56 | 65.10 | 74.99 |
| 0.70 | 72.78 | 69.79 | 65.10 | 74.47 |
| 0.71 | 72.39 | 69.39 | 65.10 | 73.56 |
| 0.72 | 69.66 | 67.31 | 65.10 | 70.18 |
| 0.73 | 67.32 | 66.41 | 65.10 | 67.32 |
| 0.74 | 66.53 | 65.88 | 65.10 | 66.79 |
| 0.75 | 65.75 | 65.75 | 65.10 | 66.01 |
| 0.76 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.77 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.78 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.79 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.80 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.81 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.82 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.83 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.84 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.85 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.86 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.87 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.88 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.89 | 65.10 | 65.10 | 65.10 | 65.10 |
| 0.90 | 65.10 | 65.10 | 65.10 | 65.10 |

Table 6.26 Friedman aligned ranks test for Pima data set.

| Algorithm | Rank | Algorithm | Rank | Algorithm | Rank | Friedman aligned ranks | |
|---|---|---|---|---|---|---|---|
| | | | | | | **Total N** | 31 |
| **Zadeh** | 10.06 | **Zadeh_L** | 8.11 | **Zadeh_L2** | 4.81 | **Test Statistic (df)** | 200.796 (11) |
| **Algebraic Product/Sum** | 10.65 | **Algebraic Product/Sum_L** | 8.66 | **Algebraic Product/Sum_L2** | 2.29 | | |
| **Bounded Product/Sum** | 6.52 | **Bounded Product/Sum_L** | 5.76 | **Bounded Product/Sum_L2** | 1.73 | **Asymptotic Sig. (2 sided test)** | 0.000 |
| **Non Parametric Hamacher ($\lambda = 0$)** | 6.65 | **Non Parametric Hamacher_L ($\lambda = 0$)** | 7.03 | **Non Parametric Hamacher_L2 ($\lambda = 0$)** | 5.74 | | |

While Algebraic Product/Sum (p-value, 0.004) and Bounded Product/Sum (p-value, 0.000) operators with FID3-L-WABL has better performance than Zadeh with FID3-L-WABL, Non Parametric Hamacher (p-value, 0.000) has worse performance than it. Algebraic Product/Sum with FID3-L-WABL has better performance than Bounded Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL. Hence, Non Parametric Hamacher with FID3-L-WABL (p-value, 0.000) has better performance than Bounded Product/Sum operator with FID3-L-WABL.

Zadeh with FID3-LR performs better than Algebraic Product/Sum (p-value, 0.000) and Bounded Product/Sum (p-value, 0.000) operators with FID3-LR. But, Non Parametric Hamacher with FID3-LR works better than Zadeh with FID3-LR.

Algebraic Product/Sum with FID3-LR has a better performance than Bounded Product/Sum (p-value, 0.000) operator with FID3-LR. Non Parametric Hamacher with FID3-LR works better than Algebraic Product/Sum (p-value, 0.000) and Bounded Product/Sum (p-value, 0.000) with FID3-LR.

Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL has better performance than Zadeh with FID3-LR. In a similar manner, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) with

FID3-L-WABL has better performance than Algebraic Product/Sum with FID3-LR. Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL also has better performance than Bounded Product/Sum with FID3-LR. Additionaly, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.001), and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL has better performance than Non Parametric Hamacher with FID3-LR.

Zadeh (p-value, 0.000) and Algebraic Product/Sum (p-value, 0.000) operator with classical FID3 have better performance than Zadeh operator with FID3-L-WABL. On the other hand, Zadeh operator with FID3-L-WABL has better performance than Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) with classical FID3.

Algebraic Product/Sum with FID3-L-WABL has better performance than Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) with classical FID3.

Yet, Bounded Product/Sum (p-value, 0.010) and Non Parametric Hamacher (p-value, 0.007) with classical FID3 have better performance than Bounded Product/Sum with FID3-L-WABL. However, Bounded Product/Sum (p-value, 0.001) with classical FID3 has better performance than Non Parametric Hamacherwith FID3-L-WABL.

Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) operators with classical FID3 have better performance than Zadeh operator with FID3-LR. In a similar manner, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) operators with classical FID3 have better performance than Algebraic Product/Sum operator with FID3-LR. Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) operators with classical FID3 also have better performance than Bounded Product/Sum operator with FID3-LR. Finally, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.002), and Non Parametric Hamacher (p-value, 0.002) operators with classical FID3 have better performance than Non Parametric Hamacher operator with FID3-LR.

*Ring data set*

Ring data set performance results for three approaches are given in Table 6.27, Table 6.28, and Table 6.29, respectively.

While classical FID3 approach with the threshold 0.72 is performed for the induction process on Ring data set, Algebraic Product/Sum operator has the higest accuracy rate with 75.07% among the other non-parametric operators.

While FID3-L-WABL approach with 0.69 threshold is performed for the induction process on Ring data set, Algebraic Product/Sum operator has the higest accuracy rate with 75.01% among the other non-parametric operators.

While FID3-LR approach with 0.60 threshold is performed for the induction process on Ring data set, Non parametric Hamacher operator has the higest accuracy rate with 62.54% among the other non-parametric operators.

The Friedman aligned ranks as a non-parametric statistical procedure is applied to detect statistical differences among a group of results on 31 thresholds value between 0.60-0.90. for three approaches. The test results performed on both approaches have significant p-value, 0.000 as given in Table 6.30. It is seen that there are significant differences among the results ($\alpha$=0.05).

Table 6.27 Ring data set performance results (%) different t-operators on different threshold results for classical FuzzyID3 (FID3).

| **Threshold** | **Zadeh** | **Algebraic Product/ Sum** | **Bounded Product/ Sum** | **NP_ Hamacher** |
|---|---|---|---|---|
| 0.60 | 70.20 | 70.41 | 65.97 | 65.08 |
| 0.61 | 72.36 | 72.34 | 66.11 | 65.96 |
| 0.62 | 72.57 | 72.38 | 66.15 | 65.69 |
| 0.63 | 72.50 | 72.41 | 66.14 | 65.53 |
| 0.64 | 72.78 | 72.78 | 66.39 | 65.12 |
| 0.65 | 73.01 | 73.09 | 66.49 | 65.05 |
| 0.66 | 73.21 | 74.01 | 66.70 | 65.68 |
| 0.67 | 73.43 | 73.44 | 66.70 | 65.59 |
| 0.68 | 73.84 | 73.95 | 66.64 | 66.88 |
| 0.69 | 74.84 | 74.95 | 66.19 | 68.31 |
| 0.70 | 74.82 | 74.95 | 65.59 | 68.68 |
| 0.71 | 74.82 | 75.01 | 75.01 | 69.18 |
| 0.72 | 74.61 | **75.07** | 64.72 | 69.22 |
| 0.73 | 73.20 | 74.43 | 63.72 | 68.77 |
| 0.74 | 73.55 | 74.81 | 63.82 | 69.01 |
| 0.75 | 73.00 | 74.55 | 63.49 | 66.99 |
| 0.76 | 71.20 | 73.18 | 62.41 | 68.00 |
| 0.77 | 71.74 | 74.42 | 61.55 | 68.14 |
| 0.78 | 68.47 | 73.88 | 58.93 | 65.53 |
| 0.79 | 65.05 | 69.36 | 57.23 | 62.69 |
| 0.80 | 64.20 | 69.66 | 56.66 | 62.32 |
| 0.81 | 63.26 | 69.27 | 55.99 | 61.65 |
| 0.82 | 62.15 | 68.51 | 55.30 | 60.55 |
| 0.83 | 59.41 | 64.27 | 54.20 | 58.28 |
| 0.84 | 56.53 | 60.00 | 52.95 | 55.93 |
| 0.85 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.86 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.87 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.88 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.89 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.90 | 49.51 | 49.51 | 49.51 | 49.51 |

Table 6.28 Ring data set performance results (%) different t-operators on different threshold results for FuzzyID3-L-WABL(FID3-L-WABL).

| **Threshold** | **Zadeh_L** | **Algebraic Product/ Sum_L** | **Bounded Product/ Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 69.82 | 70.20 | 66.03 | 65.30 |
| 0.61 | 71.85 | 72.31 | 66.30 | 66.26 |
| 0.62 | 71.99 | 72.54 | 66.34 | 66.27 |
| 0.63 | 71.96 | 72.54 | 66.32 | 65.86 |
| 0.64 | 72.00 | 72.59 | 66.34 | 65.70 |
| 0.65 | 72.11 | 72.58 | 66.45 | 65.50 |
| 0.66 | 72.24 | 72.74 | 66.59 | 65.32 |
| 0.67 | 73.04 | 73.55 | 66.97 | 66.49 |
| 0.68 | 73.53 | 74.03 | 66.91 | 67.77 |
| 0.69 | 74.24 | **75.01** | 70.43 | 69.07 |
| 0.70 | 74.30 | 74.73 | 66.11 | 68.95 |
| 0.71 | 74.27 | 74.63 | 65.56 | 69.05 |
| 0.72 | 74.31 | 74.39 | 65.58 | 68.97 |
| 0.73 | 74.41 | 74.96 | 65.39 | 69.11 |
| 0.74 | 74.51 | 74.95 | 65.04 | 68.78 |
| 0.75 | 74.14 | 74.88 | 64.61 | 68.88 |
| 0.76 | 73.45 | 74.41 | 63.69 | 68.70 |
| 0.77 | 64.14 | 67.38 | 56.73 | 62.07 |
| 0.78 | 53.43 | 55.07 | 50.86 | 52.65 |
| 0.79 | 52.35 | 53.47 | 50.68 | 51.73 |
| 0.80 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.81 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.82 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.83 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.84 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.85 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.86 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.87 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.88 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.89 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.90 | 49.51 | 49.51 | 49.51 | 49.51 |

Table 6.29 Ring data set performance results (%) different t-operators on different threshold results for Fuzzy ID3-LR (FID3-LR).

| **Threshold** | **Zadeh_L** | **Algebraic Product/Sum_L** | **Bounded Product/Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 60.34 | 63.38 | 53.55 | **62.54** |
| 0.61 | 60.04 | 62.86 | 53.55 | 62.09 |
| 0.62 | 59.60 | 63.08 | 53.58 | 62.08 |
| 0.63 | 59.06 | 63.16 | 53.58 | 62.08 |
| 0.64 | 58.65 | 62.78 | 51.49 | 61.72 |
| 0.65 | 58.22 | 62.16 | 51.00 | 61.02 |
| 0.66 | 57.86 | 61.96 | 50.99 | 60.86 |
| 0.67 | 57.86 | 61.96 | 50.99 | 60.86 |
| 0.68 | 57.30 | 61.99 | 50.98 | 60.73 |
| 0.69 | 54.72 | 57.29 | 50.57 | 56.31 |
| 0.70 | 54.91 | 57.34 | 50.57 | 56.42 |
| 0.71 | 51.58 | 52.51 | 50.01 | 51.86 |
| 0.72 | 53.15 | 52.76 | 50.14 | 52.11 |
| 0.73 | 51.08 | 49.76 | 49.64 | 49.76 |
| 0.74 | 51.08 | 49.76 | 49.64 | 49.76 |
| 0.75 | 51.08 | 49.76 | 49.64 | 49.76 |
| 0.76 | 51.08 | 49.76 | 49.64 | 49.76 |
| 0.77 | 51.08 | 49.76 | 49.64 | 49.76 |
| 0.78 | 51.08 | 49.76 | 49.64 | 49.76 |
| 0.79 | 51.08 | 49.76 | 49.64 | 49.76 |
| 0.80 | 51.08 | 49.76 | 49.64 | 49.76 |
| 0.81 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.82 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.83 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.84 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.85 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.86 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.87 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.88 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.89 | 49.51 | 49.51 | 49.51 | 49.51 |
| 0.90 | 49.51 | 49.51 | 49.51 | 49.51 |

Then, Wilcoxon signed rank tests are applied into the ring data set to test the comparison among the reasoning methods.

It is seen that the results of Algebraic Product/Sum operator are better than the results of Zadeh operator produced with classical FID3 (p-value, 0.000), significantly. Bounded Product/Sum and Non Parametric Hamacher operators with classical FID3 have worse performance than Zadeh operator with classical FID3 (p value, 0.000).

In a similar manner, Bounded Product/Sum and Non Parametric Hamacher operators with classical FID3 have worse performance than Algebraic Product/Sum operator with classical FID3 (p value, 0.000). Yet, Non Parametric Hamacher operators with classical FID3 has better performance than Bounded Product/Sum operator (p value, 0.004).

While Algebraic Product/Sum (p-value, 0.004) operator with FID3-L-WABL has better performance than Zadeh with FID3-L-WABL, Bounded Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) has worse performance than it. Algebraic Product/Sum with FID3-L-WABL has better performance than Bounded Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL. Hence, Non Parametric Hamacher with FID3-L-WABL (p-value, 0.000) has better performance than Bounded Product/Sum operator with FID3-L-WABL.

Zadeh with FID3-LR performs better than Bounded Product/Sum (p-value, 0.000) operators with FID3-LR. But, Algebraic Product/Sum (p-value, 0.014) and Non Parametric Hamacher (p-value, 0.015) with FID3-LR works better than Zadeh with FID3-LR.

Algebraic Product/Sum with FID3-LR is better performance than Bounded Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) operator with FID3-LR. Non Parametric Hamacher with FID3-LR works better than Bounded Product/Sum (p-value, 0.000) with FID3-LR.

Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL has better performance than Zadeh with FID3-LR.

In a similar manner, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL has better performance than Algebraic Product/Sum with FID3-LR.

Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL also has better performance than Bounded Product/Sum with FID3-LR. Additionally, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.001), and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL has better performance than Non Parametric Hamacher with FID3-LR.

Table 6.30 Friedman aligned ranks test for Ring data set.

| **Algorithm** | **Rank** | **Algorithm** | **Rank** | **Algorithm** | **Rank** | **Friedman aligned ranks** | |
|---|---|---|---|---|---|---|---|
| | | | | | | **Total N** | 31 |
| **Zadeh** | 9.81 | **Zadeh_L** | 7.69 | **Zadeh_L2** | 3.97 | **Test Statistic (df)** | 210.242 (11) |
| **Algebraic Product/Sum** | 10.48 | **Algebraic Product/Sum_L** | 8.74 | **Algebraic Product/Sum_L2** | 4.26 | | |
| **Bounded Product/Sum** | 6.98 | **Bounded Product/Sum_L** | 6.11 | **Bounded Product/Sum_L2** | 2.65 | **Asymptotic Sig. (2 sided test)** | 0.000 |
| **Non Parametric Hamacher** $(\lambda = 0)$ | 7.19 | **Non Parametric Hamacher_L** $(\lambda = 0)$ | 6.27 | **Non Parametric Hamacher_L2** $(\lambda = 0)$ | 3.84 | | |

Zadeh (p-value, 0.000) and Algebraic Product/Sum (p-value, 0.000) operator with classical FID3 have better performance than Zadeh operator with FID3-L-WABL. On the other hand, Zadeh operator with FID3-L-WABL has better performance than Bounded Product/Sum (p-value, 0.005) with classical FID3.

Algebraic Product/Sum with FID3-L-WABL has better performance than Zadeh (p-value, 0.021), and Algebraic Product/Sum (p-value, 0.005) with classical FID3.

Algebraic Product/Sum with FID3-L-WABL has worse performance than Bounded Product/Sum (p-value, 0.0001).

Yet, Bounded Product/Sum (p-value, 0.010) and Non Parametric Hamacher (p-value, 0.007) with classical FID3 have better performance than Bounded Product/Sum with FID3-L-WABL. However, Bounded Product/Sum (p-value, 0.001) with classical FID3 has better performance than Non Parametric Hamacherwith FID3-L-WABL.

Zadeh (p-value, 0.000) and Algebraic Product/Sum (p-value, 0.000) with classical FID3 have better performance than Non Parametric Hamacher with FID3-L-WABL.

Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) operators with classical FID3 have better performance than Zadeh operator with FID3-LR.

In a similar manner, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) operators with classical FID3 have better performance than Algebraic Product/Sum operator with FID3-LR.

Also, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) operators with classical FID3 have better performance than Bounded Product/Sum operator with FID3-LR.

Finally, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.002), and Non Parametric Hamacher (p-value, 0.002) operators with classical FID3 have better performance than Non Parametric Hamacher operator with FID3-LR.

*Sonar data set*

Sonar data set performance results for three approaches are given in Table 6.31, Table 6.32, and Table 6.33, respectively.

While classical FID3 approach with the threshold 0.88 is performed for the induction process on Sonar data set, Zadeh operator has the higest accuracy rate with 75.49% among the other non-parametric operators.

While FID3-L-WABL approach with 0.86 threshold is performed for the induction process on Sonar data set, Bounded Product/Sum operator has the higest accuracy rate with 77.39% among the other non-parametric operators.

While FID3-LR approach with 0.88 threshold is performed for the induction process on Sonar data set, Non parametric Hamacher operator has the higest accuracy rate with 77.42% among the other non-parametric operators.

The Friedman aligned ranks as a non-parametric statistical procedure is applied to detect statistical differences among a group of results on 31 thresholds value between 0.60-0.90. for three approaches. The test results performed on both approaches have significant p-value, 0.000 as given in Table 6.34. It is seen that there are significant differences among the results ($\alpha$=0.05).

Table 6.31 Sonar data set performance results (%) different t-operators on different threshold results for classical Fuzzy ID3 (FID3).

| **Threshold** | **Zadeh** | **Algebraic Product/ Sum** | **Bounded Product/ Sum** | **NP_ Hamacher** |
|---|---|---|---|---|
| 0.60 | 73.60 | 73.61 | 73.61 | 73.61 |
| 0.61 | 73.60 | 73.61 | 73.61 | 73.61 |
| 0.62 | 73.60 | 73.61 | 73.61 | 73.61 |
| 0.63 | 73.60 | 73.61 | 73.61 | 73.61 |
| 0.64 | 73.60 | 73.61 | 73.61 | 73.61 |
| 0.65 | 73.60 | 73.61 | 73.61 | 73.61 |
| 0.66 | 73.60 | 73.61 | 73.61 | 73.61 |
| 0.67 | 73.60 | 73.61 | 73.61 | 73.61 |
| 0.68 | 73.60 | 73.61 | 73.61 | 73.61 |
| 0.69 | 74.09 | 73.61 | 74.09 | 73.61 |
| 0.70 | 74.09 | 73.12 | 74.10 | 72.14 |
| 0.71 | 74.56 | 72.65 | 72.62 | 71.19 |
| 0.72 | 73.59 | 67.39 | 72.64 | 63.99 |
| 0.73 | 72.61 | 67.39 | 70.68 | 63.99 |
| 0.74 | 73.08 | 63.11 | 71.16 | 58.72 |
| 0.75 | 73.08 | 63.11 | 71.64 | 58.72 |
| 0.76 | 72.60 | 63.11 | 70.69 | 58.72 |
| 0.77 | 74.03 | 64.54 | 72.11 | 59.20 |
| 0.78 | 74.03 | 64.54 | 72.11 | 59.20 |
| 0.79 | 74.03 | 56.74 | 71.14 | 53.83 |
| 0.80 | 74.03 | 56.74 | 72.57 | 52.39 |
| 0.81 | 74.03 | 51.02 | 72.57 | 48.58 |
| 0.82 | 72.60 | 46.63 | 70.65 | 46.63 |
| 0.83 | 72.60 | 46.63 | 70.65 | 46.63 |
| 0.84 | 73.55 | 46.63 | 70.64 | 46.63 |
| 0.85 | 72.60 | 46.63 | 69.69 | 46.63 |
| 0.86 | 74.07 | 54.92 | 71.15 | 48.58 |
| 0.87 | 74.54 | 69.79 | 72.59 | 55.75 |
| 0.88 | **75.49** | 69.79 | 72.58 | 56.72 |
| 0.89 | 74.53 | 68.83 | 71.63 | 59.62 |
| 0.90 | 74.54 | 74.05 | 71.15 | 62.50 |

Table 6.32 Sonar data set performance results (%) different t-operators on different threshold results for FuzzID3-L-WABL (FID3-L-WABL).

| **Threshold** | **Zadeh_L** | **Algebraic Product/ Sum_L** | **Bounded Product/ Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 73.14 | 73.62 | 73.62 | 73.62 |
| 0.61 | 73.14 | 73.62 | 73.62 | 73.62 |
| 0.62 | 73.14 | 73.62 | 73.62 | 73.62 |
| 0.63 | 73.14 | 73.62 | 73.62 | 73.62 |
| 0.64 | 73.14 | 73.62 | 73.62 | 73.62 |
| 0.65 | 73.14 | 73.62 | 73.62 | 73.62 |
| 0.66 | 73.14 | 73.62 | 73.62 | 73.62 |
| 0.67 | 73.14 | 73.62 | 73.62 | 73.62 |
| 0.68 | 73.14 | 73.62 | 73.62 | 73.62 |
| 0.69 | 73.62 | 73.62 | 74.11 | 72.65 |
| 0.70 | 73.62 | 73.62 | 74.09 | 72.15 |
| 0.71 | 73.61 | 73.14 | 73.11 | 70.19 |
| 0.72 | 72.62 | 72.16 | 72.62 | 67.74 |
| 0.73 | 73.10 | 72.16 | 72.62 | 65.84 |
| 0.74 | 73.10 | 73.59 | 73.58 | 64.41 |
| 0.75 | 73.10 | 74.08 | 73.09 | 64.41 |
| 0.76 | 73.58 | 73.12 | 73.09 | 63.94 |
| 0.77 | 75.01 | 74.55 | 74.99 | 64.41 |
| 0.78 | 75.01 | 74.55 | 74.99 | 65.39 |
| 0.79 | 75.01 | 67.72 | 74.99 | 58.08 |
| 0.80 | 75.01 | 56.77 | 74.52 | 50.94 |
| 0.81 | 74.05 | 46.63 | 74.52 | 46.63 |
| 0.82 | 75.02 | 46.63 | 75.01 | 46.63 |
| 0.83 | 75.49 | 52.82 | 75.48 | 51.87 |
| 0.84 | 75.97 | 52.82 | 75.96 | 51.87 |
| 0.85 | 75.97 | 46.63 | 75.96 | 46.63 |
| 0.86 | 76.45 | 50.44 | **77.39** | 47.58 |
| 0.87 | 76.45 | 62.64 | 76.42 | 49.05 |
| 0.88 | 76.93 | 62.15 | 76.42 | 50.99 |
| 0.89 | 75.49 | 67.86 | 76.42 | 55.31 |
| 0.90 | 75.98 | 74.05 | 75.46 | 57.68 |

Table 6.33 Sonar data set performance results (%) different t-operators on different threshold results for FuzzyID3-LR (FID3-LR).

| **Threshold** | **Zadeh_L** | **Algebraic Product/Sum_L** | **Bounded Product/Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 65.89 | 70.73 | 63.94 | 69.77 |
| 0.61 | 65.89 | 70.73 | 63.94 | 69.77 |
| 0.62 | 66.86 | 69.76 | 63.94 | 68.30 |
| 0.63 | 66.86 | 69.76 | 63.94 | 68.79 |
| 0.64 | 66.86 | 69.76 | 63.94 | 68.79 |
| 0.65 | 65.89 | 62.44 | 63.94 | 60.99 |
| 0.66 | 65.59 | 62.44 | 63.94 | 60.99 |
| 0.67 | 68.27 | 62.52 | 65.37 | 63.02 |
| 0.68 | 69.70 | 65.38 | 66.32 | 63.48 |
| 0.69 | 70.19 | 65.38 | 66.32 | 63.48 |
| 0.70 | 70.19 | 65.38 | 66.32 | 62.51 |
| 0.71 | 69.23 | 72.52 | 61.08 | 69.65 |
| 0.72 | 70.19 | 71.20 | 60.60 | 65.91 |
| 0.73 | 72.09 | 71.70 | 57.74 | 68.28 |
| 0.74 | 73.06 | 77.39 | 57.74 | 73.08 |
| 0.75 | 74.03 | 77.89 | 57.75 | 74.99 |
| 0.76 | 74.97 | 76.92 | 57.75 | 73.56 |
| 0.77 | 76.92 | 77.43 | 56.77 | 75.02 |
| 0.78 | 75.98 | 76.93 | 56.77 | 75.97 |
| 0.79 | 76.96 | 77.41 | 56.77 | 75.49 |
| 0.80 | 76.96 | 77.41 | 57.25 | 73.10 |
| 0.81 | 76.48 | 77.41 | 57.25 | 74.53 |
| 0.82 | 76.95 | 77.40 | 57.72 | 75.48 |
| 0.83 | 76.46 | 77.40 | 57.72 | 75.48 |
| 0.84 | 76.93 | 77.40 | 57.72 | 74.54 |
| 0.85 | **77.42** | 78.85 | 57.72 | 76.96 |
| 0.86 | 76.93 | 78.37 | 56.76 | 75.04 |
| 0.87 | 78.62 | 79.33 | 55.30 | 77.89 |
| 0.88 | 76.92 | 76.45 | 54.82 | 75.96 |
| 0.89 | 75.49 | 75.97 | 54.82 | 76.46 |
| 0.90 | 76.92 | 75.51 | 53.87 | 76.46 |

Then, Wilcoxon signed rank tests are applied into the sonar data set to test the comparison among the reasoning methods.

It is seen that the results of Zadeh operator with classical FID3are better than the results of produced by Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) operators with classical FID3. Bounded Product/Sum operator with classical FID3 have better performance than Algebraic Product/Sum operator with classical FID3 (p value, 0.000).

In a similar manner, Non Parametric Hamacher operators with classical FID3 have worse performance than Bounded Product/Sum(p value, 0.000) and Algebraic Product/Sum (p value, 0.000) operators with classical FID3.

Algebraic Product/Sum (p-value, 0.019) and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL has worse performance than Zadeh with FID3-L-WABL. Bounded Product/Sum (p-value, 0.000) with FID3-L-WABL has better performance than Algebraic Product/Sum with FID3-L-WABL.

Hence, Non Parametric Hamacher with FID3-L-WABL (p-value, 0.000) has worse performance than Algebraic Product/Sum and Bounded Product/Sum operator with FID3-L-WABL.

Zadeh with FID3-LR is better performance than Bounded Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.004) operators with FID3-LR. And, Algebraic Product/Sum (p-value, 0.000) and Bounded Product/Sum (p-value, 0.000) with FID3-LR Zadeh with FID3-LR work better than Non Parametric Hamacher with FID3-LR Zadeh with FID3-LR.

Non Parametric Hamacher (p-value, 0.000) with FID3-LR has better performance than Bounded Product/Sum (p-value, 0.000) with FID3-LR .

Table 6.34 Friedman aligned ranks test for Sonar data set.

| **Algorithm** | **Rank** | **Algorithm** | **Rank** | **Algorithm** | **Rank** | **Friedman aligned ranks** | |
|---|---|---|---|---|---|---|---|
| | | | | | | **Total N** | 31 |
| **Zadeh** | 7.27 | **Zadeh_L** | 8.16 | **Zadeh_L2** | 7.34 | **Test Statistic (df)** | 101.591 (11) |
| **Algebraic Product/Sum** | 5.13 | **Algebraic Product/Sum_L** | 7.42 | **Algebraic Product/Sum_L2** | 8.00 | | |
| **Bounded Product/Sum** | 6.89 | **Bounded Product/Sum_L** | 9.60 | **Bounded Product/Sum_L2** | 2.42 | **Asymptotic Sig. (2 sided test)** | 0.000 |
| **Non Parametric Hamacher** $(\lambda = 0)$ | 4.21 | **Non Parametric Hamacher_L** $(\lambda = 0)$ | 5.34 | **Non Parametric Hamacher_L2** $(\lambda = 0)$ | 6.23 | | |

While Zadeh (p-value, 0.031) and Bounded Product/Sum (p-value, 0.041) with FID3-L-WABL has better performance than Zadeh with FID3-LR, Non Parametric Hamacher (p-value, 0.001) with FID3-L-WABL has a worse performance than Zadeh with FID3-LR.

Algebraic Product/Sum (p-value, 0.000) with FID3-LR has a better performance than Non Parametric Hamacher (p-value, 0.001) with FID3-L-WABL.

Also, Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.024) with FID3-L-WABL has better performance than Bounded Product/Sum with FID3-LR.

While Zadeh (p-value, 0.001), and Bounded Product/Sum (p-value, 0.001) with FID3-L-WABL has better performance than Non Parametric Hamacher with FID3-LR, Non Parametric Hamacher (p-value, 0.001) with FID3-LR has a better performance than Non Parametric Hamacher with FID3-L-WABL.

Zadeh (p-value, 0.011), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) operator with classical FID3 have worse performance than Zadeh operator with FID3-L-WABL. On the other hand, while Algebraic Product/Sum operator with FID3-L-WABL works worse than Zadeh (p-value, 0.007) with classical FID3, it works better than Algebraic Product/Sum (p-value, 0.004) and Non Parametric Hamacher (p-value, 0.000) with classical FID3.

Bounded Product/Sum with FID3-L-WABL has a better performance than Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.000), Bounded Product/Sum (p-value, 0.000), and Non Parametric Hamacher (p-value, 0.000) with classical FID3.

Non Parametric Hamacher with FID3-L-WABL has a better performance than Zadeh (p-value, 0.000), Bounded Product/Sum (p-value, 0.000) with classical FID3.

Algebraic Product/Sum (p-value, 0.006) and Non Parametric Hamacher (p-value, 0.000) operators with classical FID3 have better performance than Zadeh operator with FID3-LR.

Algebraic Product/Sum with FID3-L-LR has a better performance than Algebraic Product/Sum (p-value, 0.005) and Non Parametric Hamacher (p-value, 0.001) with classical FID3.

Bounded Product/Sum with FID3-LR has a worse performance than Zadeh, Algebraic Product/Sum (p-value, 0.000) and Bounded (p-value, 0.003) with classical FID3.

While Non Parametric Hamacher with FID3-LR has a worse performance than Zadeh (p-value, 0.023) with classical FID3, it has a better performance than Algebraic Product/Sum (p-value, 0.030) and Non Parametric Hamacher (p-value, 0.001) with classical FID3.

*Wdbc data set*

Wdbc data set performance results for three approaches are given in Table 6.35, Table 6.36, and Table 6.37, respectively.

While classical Fuzzy ID3 approach with the threshold 0.88 is performed for the induction process on Wdbc data set, Algebraic Product/Sum operator has the higest accuracy rate with 93.67% among the other non-parametric operators.

While Fuzzy ID3-L-WABL approach with 0.90 threshold is performed for the induction process on Wdbc data set, Bounded Product/Sum operator has the higest accuracy rate with 94.55% among the other non-parametric operators.

While Fuzzy ID3-LR approach with threshold range 0.88-0.90 is performed for the induction process on Wdbc data set, Algebraic Product/Sum operator has the higest accuracy rate with 94.38% among the other non-parametric operators.

The Friedman aligned ranks as a non-parametric statistical procedure is applied to detect statistical differences among a group of results on 31 thresholds value between 0.60-0.90 for three approaches. The test results performed on both approaches have significant p-value, 0.000 as given in Table 6.38. It is seen that there are significant differences among the results ($\alpha$=0.05).

Then, Wilcoxon signed rank tests are applied into the Wdbc data set to test the comparison among the reasoning methods. It is seen that the results of Zadeh operator with classical FID3 are worse than the results of produced by Bounded Product/Sum (p-value, 0.000) with classical FID3. Zadeh with classical FID3 has better performance than Non Parametric Hamacher (p-value, 0.000) with classical FID3.

While Bounded Product/Sum operator with classical FID3 has a better performance than Algebraic Product/Sum operator with classical FID3 (p value, 0.000), Non Parametric Hamacher operator has a worse performance than Algebraic Product/Sum operator with classical FID3 (p value, 0.000). In a similar manner, Non Parametric Hamacher operator with classical FID3 has a worse performance than Bounded Product/Sum (p value, 0.000) operator with classical FID3.

Table 6.35 Wdbc data set performance results (%) different t-operators on different threshold results for classical Fuzzy ID3 (FID3).

| Threshold | Zadeh | Algebraic Product/ Sum | Bounded Product/ Sum | NP_ Hamacher |
|---|---|---|---|---|
| 0.60 | 88.23 | 89.46 | 89.81 | 85.95 |
| 0.61 | 88.23 | 89.46 | 89.81 | 88.40 |
| 0.62 | 88.23 | 89.46 | 89.81 | 88.23 |
| 0.63 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.64 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.65 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.66 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.67 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.68 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.69 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.70 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.71 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.72 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.73 | 89.98 | 90.51 | 90.86 | 85.95 |
| 0.74 | 90.51 | 91.04 | 91.39 | 84.18 |
| 0.75 | 90.51 | 91.04 | 91.39 | 84.18 |
| 0.76 | 90.34 | 90.86 | 91.39 | 83.82 |
| 0.77 | 91.39 | 91.74 | 92.09 | 83.12 |
| 0.78 | 91.39 | 91.74 | 92.09 | 83.12 |
| 0.79 | 91.03 | 80.77 | 92.09 | 74.63 |
| 0.80 | 91.74 | 69.54 | 92.62 | 64.80 |
| 0.81 | 91.74 | 69.54 | 92.62 | 64.80 |
| 0.82 | 91.38 | 46.73 | 92.79 | 46.73 |
| 0.83 | 91.74 | 57.96 | 92.79 | 54.80 |
| 0.84 | 92.44 | 69.89 | 92.09 | 62.17 |
| 0.85 | 92.09 | 69.89 | 92.09 | 62.87 |
| 0.86 | 92.09 | 69.89 | 92.27 | 62.70 |
| 0.87 | 91.92 | 81.22 | 91.92 | 68.18 |
| 0.88 | 92.97 | 82.62 | 92.09 | 65.90 |
| 0.89 | 92.79 | 93.67 | 92.09 | 72.58 |
| 0.90 | 92.79 | **93.67** | 92.09 | 73.63 |

Table 6.36 Wdbc data set performance results (%) different t-operators on different threshold results for Fuzzy ID3-L-WABL (FID3-L-WABL).

| Threshold | Zadeh_L | Algebraic Product/ Sum_L | Bounded Product/ Sum_L | NP_ Hamacher_L |
|---|---|---|---|---|
| 0.60 | 88.23 | 89.46 | 89.81 | 88.40 |
| 0.61 | 90.16 | 90.33 | 90.51 | 85.95 |
| 0.62 | 90.16 | 90.33 | 90.51 | 85.95 |
| 0.63 | 90.16 | 90.33 | 90.51 | 85.95 |
| 0.64 | 90.16 | 90.33 | 90.51 | 85.95 |
| 0.65 | 90.16 | 90.33 | 90.51 | 85.95 |
| 0.66 | 90.16 | 90.33 | 90.51 | 85.95 |
| 0.67 | 89.98 | 89.98 | 90.51 | 85.95 |
| 0.68 | 89.98 | 89.98 | 90.51 | 85.95 |
| 0.69 | 89.98 | 89.98 | 90.51 | 84.89 |
| 0.70 | 89.98 | 89.98 | 90.51 | 84.89 |
| 0.71 | 89.98 | 89.98 | 90.51 | 85.89 |
| 0.72 | 89.98 | 78.75 | 90.51 | 75.60 |
| 0.73 | 89.98 | 78.75 | 90.51 | 75.60 |
| 0.74 | 90.87 | 79.82 | 91.57 | 73.83 |
| 0.75 | 90.87 | 79.82 | 91.57 | 73.83 |
| 0.76 | 92.10 | 69.36 | 92.62 | 65.15 |
| 0.77 | 92.10 | 69.36 | 92.62 | 65.15 |
| 0.78 | 92.10 | 69.36 | 92.62 | 65.15 |
| 0.79 | 92.10 | 69.36 | 92.62 | 65.15 |
| 0.80 | 92.10 | 69.36 | 92.62 | 65.15 |
| 0.81 | 92.62 | 57.96 | 92.62 | 54.98 |
| 0.82 | 92.45 | 46.91 | 92.45 | 46.91 |
| 0.83 | 92.45 | 69.54 | 92.62 | 60.24 |
| 0.84 | 92.62 | 69.89 | 92.62 | 60.07 |
| 0.85 | 92.80 | 69.89 | 92.62 | 60.94 |
| 0.86 | 92.62 | 71.47 | 92.27 | 56.03 |
| 0.87 | 92.62 | 71.82 | 92.27 | 56.21 |
| 0.88 | 92.62 | 59.89 | 92.27 | 51.47 |
| 0.89 | 94.20 | 94.38 | 93.67 | 73.44 |
| 0.90 | **94.55** | 94.38 | 93.85 | 75.74 |

Table 6.37 Wdbc data set performance results (%) different t-operators on different threshold results for Fuzzy ID3-LR (FID3-LR).

| **Threshold** | **Zadeh_L** | **Algebraic Product/Sum_L** | **Bounded Product/Sum_L** | **NP_ Hamacher_L** |
|---|---|---|---|---|
| 0.60 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.61 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.62 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.63 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.64 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.65 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.66 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.67 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.68 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.69 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.70 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.71 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.72 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.73 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.74 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.75 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.76 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.77 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.78 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.79 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.80 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.81 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.82 | 90.16 | 90.86 | 90.86 | 91.04 |
| 0.83 | 90.16 | 89.98 | 89.63 | 88.58 |
| 0.84 | 91.39 | 90.16 | 89.10 | 88.05 |
| 0.85 | 92.62 | 90.33 | 88.05 | 87.67 |
| 0.86 | 92.62 | 90.33 | 88.05 | 87.99 |
| 0.87 | 92.44 | 93.14 | 47.17 | 90.86 |
| 0.88 | 93.15 | **94.38** | 37.26 | 91.92 |
| 0.89 | 93.15 | **94.38** | 37.26 | 91.92 |
| 0.90 | 92.97 | **94.38** | 37.26 | 92.09 |

Algebraic Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL have worse performance than Zadeh with FID3-L-WABL. Bounded Product/Sum (p-value, 0.001) with FID3-L-WABL has a better performance than Zadeh operator with FID3-L-WABL. Hence, Non Parametric Hamacher with FID3-L-WABL (p-value, 0.000) has a worse performance than Algebraic Product/Sum operator with FID3-L-WABL. But, Bounded Product/Sum with FID3-L-WABL has a better performance than Algebraic Product/Sum operator (p-value, 0.000) and Non-Parametric Hamacher (p-value, 0.000) with FID3-L-WABL.

Zadeh with FID3-LR is worse performance than Algebraic Product/Sum (p-value, 0.000) operator with FID3-LR. And, Algebraic Product/Sum with FID3-LR works better than Bounded Product/Sum (p-value, 0.000) with FID3-LR Zadeh. Also, Non Parametric Hamacher with FID3-LR has a better perfomance than Bounded Product/Sum (p-value, 0.000) with FID3-LR Zadeh.

While Zadeh (p-value, 0.004) and Bounded Product/Sum (p-value, 0.000) with FID3-L-WABL have better performance than Zadeh with FID3-LR, Algebraic Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL have worse performance than Zadeh with FID3-LR.

Table 6.38 Friedman aligned ranks test for Wdbc data set.

| **Algorithm** | **Rank** | **Algorithm** | **Rank** | **Algorithm** | **Rank** | **Friedman aligned ranks** | |
|---|---|---|---|---|---|---|---|
| | | | | | | **Total N** | 31 |
| **Zadeh** | 5.77 | **Zadeh_L** | 7.77 | **Zadeh_L2** | 6.29 | **Test Statistic (df)** | 189.11 (11) |
| **Algebraic Product/Sum** | 6.19 | **Algebraic Product/Sum_L** | 4.32 | **Algebraic Product/Sum_L2** | 8.79 | | |
| **Bounded Product/Sum** | 9.26 | **Bounded Product/Sum_L** | 9.35 | **Bounded Product/Sum_L2** | 7.31 | **Asymptotic Sig. (2 sided test)** | 0.000 |
| **Non Parametric Hamacher** ($\lambda = 0$) | 2.23 | **Non Parametric Hamacher_L** ($\lambda = 0$) | 1.56 | **Non Parametric Hamacher_L2** ($\lambda = 0$) | 9.15 | | |

While Zadeh (p-value, 0.000) and Bounded Product/Sum (p-value, 0.000) with FID3-L-WABL have better performance than Zadeh with FID3-LR, Algebraic

Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL have worse performance than Algebraic Product/Sum with FID3-LR.

Zadeh (p-value, 0.016) and Bounded Product/Sum (p-value, 0.002) with FID3-L-WABL have better performance than Bounded Product/Sum with FID3-LR. Yet, Algebraic Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) with FID3-L-WABL also has worse performance than Bounded Product/Sum with FID3-LR.

While Bounded Product/Sum (p-value, 0.011) with FID3-L-WABL has better performance than Non Parametric Hamacher with FID3-LR, Algebraic Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) with FID3-LR has a better performance than Non Parametric Hamacher with FID3-L-WABL.

Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.009), and Non Parametric Hamacher (p-value, 0.000) operator with classical FID3 have worse performance than Zadeh operator with FID3-L-WABL. On the other hand, Bounded Product/Sum operator with classical FID3 works better than Zadeh (p-value, 0.000) with FID3-L-WABL.

Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.006), and Bounded Product/Sum (p-value, 0.000) with classical FID3 have better performance than Algebraic Product/Sum with FID3-L-WABL. But, Algebraic Product/Sum with FID3-L-WABL has better performance than Non Parametric Hamacher (p-value, 0.000) with classical FID3.

In a similar manner, Zadeh (p-value, 0.000), and Algebraic Product/Sum (p-value, 0.006) with classical FID3 have better performance than Bounded Product/Sum with FID3-L-WABL. But, Bounded Product/Sum with FID3-L-WABL has a better performance than Non Parametric Hamacher (p-value, 0.000) with classical FID3.

Zadeh (p-value, 0.000), Algebraic Product/Sum (p-value, 0.009), Bounded Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) operator with classical FID3 have better performance than Non Parametric Hamacher operator with FID3-L-WABL.

Bounded Product/Sum (p-value, 0.000) with classical FID3 has a better performance than Zadeh with FID3-LR. But, Zadeh with FID3-LR has a better performance than Non Parametric Hamacher (p-value, 0.000) with classical FID3.

Zadeh (p-value, 0.029), Algebraic Product/Sum (p-value, 0.000) and Non Parametric Hamacher (p-value, 0.000) with classical FID3 have worse performance than Algebraic Product/Sum with FID3-LR.

Bounded Product/Sum with FID3-LR has a better performance than Algebraic Product/Sum (p-value, 0.035) and Non Parametric Hamacher (p-value, 0.002) with classical FID3. Yet, Bounded Product/Sum with classical FID3 has a better performance than Bounded Product/Sum with FID3-LR.

Non Parametric Hamacher (p-value, 0.000) with FID3-LR performs better than Algebraic Product/Sum with classical FID3.

#### *6.3.4.2 Conclusion*

In this study, novel two fuzzy decision tree approaches for linguistic data are proposed. In daily life, words are used for the communication. Especially, this century is the evolution of the information. This information is stored as words. It is necessary to solve the relations among the words and sentences. This study aims to seek novel ways to find the rules among the data stored as words. In addition, it gives a methodology in order to solve the classification problem for fuzzy data warehouses.

In the first approach, L-R fuzzy data is used then WABL method for Defuzzification is adapted and it is applied. Fuzzy c-means algorithm is also used in order to handle membership degrees. After the fuzzification phase, fuzzy ID3 approach is applied. It is seen that FID3-L-WABL approach has a good performance not less than classical FID3. In second approach, L-R fuzzy data is adapted into FuzzyID3 algorithm, directly. FID3-LR uses FkM-F algorithm for the fuzzification phase. It is observed that the comparisons have shown that this approach has a better performance in some different reasoning approaches.

The behaviour of this approach by using 31 threshold value changes in the range 0.60-0.90 is analyzed for 6 well known data sets. It is seen that FID3-L-WABL and FID3-LR with different T-operators have better behaviour on some data sets (Iris data set, Phoneme data set) than classical FID3.

It is observed that Iris data set has the maximum accuracy rates of 96.67% for Algebraic Product/Sum with classical FID3 and Algebraic Product/Sum FID3-L-WABL while $\theta_r = 0.75$. Wdbc data set has the highest accuracy rate with 94.55% for Bounded Product/Sum with FID3-L-WABL while $\theta_r = 0.75$. While Sonar data set has the highest accuracy rate 77.42% for Non-parametric Hamacher with FID3-LR($\theta_r = 0.75$), Pima data set has the highest accuracy rate 76.04% for Non-parametric Hamacher with classical FID3($\theta_r = 0.75$).

Phoneme data set has the maximum accuracy rate of 77.01% for Non-parametric Hamacher with FID3-LR. Lastly, Ring has the highest performance for Non-Parametric Hamacher with FID3-LR.

Four small data sets (Iris, Wdbc, Sonar, Pima) have good performances on classical FID3, Fuzzy-L-WABL, and Fuzzy-LR on different T-operators. Hence, Fuzzy-LR has better performance than the other approaches for large data sets (Phoneme, and Ring).

In the future, several works remains to be addressed. This study can directly be applied to the data set which is defined as linguistics. Moreover, a synergy can be adapted into the study between overlap functions and decomposition strategies for linguistic data approach. Finally, linguistic summary can be adapted in to the reasoning procedure in order to find the important rules.

# CHAPTER SEVEN
# CONCLUSION

In this work, a fundamental solution is proposed to solve the geographic classification problem and the effects of different T-operators on its reasoning procedure on numeric data is investigated in this problem solution. In addition, two novel fuzzy ID3 approaches working on linguistic data have been proposed. The first one is FID3-L-WABL (Fuzzy ID3 Algorithm Based on Linguistic Data by Using WABL Defuzzification Method) which is a novel version of the known Fuzzy Interactive Dichotomizer 3 (Fuzzy ID3) classification algorithm working on linguistic data. The second one is FID3-LR (Fuzzy ID3 Algorithm for L-R Fuzzy Data) which is a mixture of FkM-F (Fuzzy k-means Clustering Model for Fuzzy data) clustering algorithm working on L-R fuzzy data and Fuzzy Interactive Dichotomizer 3 (Fuzzy ID3) classification algorithms.

Fuzzy c-means algorithm and FkM-F were performed in MATLAB 2014a. The codes for the experiments, FuzzyID3 by using T-operators, FID3-L-WABL, and FID3-LR, have been developed in the MS Visual Studio C# IDE for the experimental study (intel i7, 2.4 GHz, 4 Gb RAM). OliveDeSoft is designed for current and future studies to analyze the olive oil quality and geographic characterization. In addition, fuzzy ID3 algorithm has been designed as an integrated software system called as Fuzzy Artemis.

The fundamental idea of the FID3-L-WABL is to work on fuzzy data. In this approach, L-R (Left-Right) fuzzy number is used. Each fuzzy number is defuzzified by using WABL (Weighted Averaging Based on Levels). Then, it is adapted with Fuzzy c-means algorithm to achieve the fuzzification. Consequently, Fuzzy ID3 algorithm is applied. In other words, FID3-L-WABL is worked on linguistic dataset to provide a classification system. It is flexible. WABL is used to calculate the average representative of a fuzzy number. WABL approach is the most robust mathematical model among the defuzzificaiton methods. It supports the performance of FID3 algorithm to obtain the rules within the linguistic dataset.

The other proposed algorithm, FID3-LR, works directly on L-R (Left-Right) fuzzy data. The fuzzificaiton is done by FkM-F. FkM-F is a fuzzy clustering algorithm which is presented for L-R (Left-Rigt) fuzzy data. It uses a weighted dissimilarity measure to compute the distances between two fuzzy L-R data. It handles membership degrees for each cluster whose number is specified before. This approach is convenient if the database is defined as linguistic.

These two novel approaches are supported with the different non-parametric T-opeators on reasoning procedure. Computational experiments are performed and, these experiments are encouraged by statistical analyses. After experiments with various T-operators on six different datasets, the proposed approaches that give better results have been observed.

To summarized, in this thesis;

Fuzzy ID3 algorithm (FID3) has been discoursed and a geographic classification problem for virgin olive oil is analyzed by using different T-operators on reasoning phase.

*A software called as OliveDeSoft is proposed* in order to classify the olive oil samples.

Fuzzy ID3 algorithm (FID3) has been achieved and a novel FID3 algorithm, called FID3-L-WABL, of which is linguistic variant *has been proposed*. WABL defuzzification method is used to defuzzify L-R fuzzy data.

*FID3-LR algorithm is suggested on the basis of Fuzzy ID3 algorithm* which works directly on L-R fuzzy data. It works fundamentally on linguistic databases to solve the classification problems.

A software called *FuzzyArtemis is presented* to succeed in the experimental study for Fuzzy ID3, Fuzzy ID3-L-WABL, and Fuzzy ID3-LR.

Fuzzy c-means (FCM) algorithm and FkM-F are performed in MATLAB 2014a.

The codes for the experiments, FuzzyID3 by using T-operators, FID3-L-WABL, and FID3-LR, have been developed in the MS Visual Studio C# IDE for the experimental study (intel i7, 2.4 GHz, 4 Gb RAM). They have been designed and integrated into a software.

In the future, OliveDeSoft can be improved to determine the quality of virgin olive oil and the charaterization of olive oil as a tool of geographic indications for olive oil sector in Turkey. The proposed algorithms can also be applied directly into linguistic databases. Moreover, a synergy can be adapted into the study between overlap functions and decomposition strategies for linguistic data approach. Finally, linguistic summary can be adapted into the reasoning procedure to make the summarization of the linguistic data sets by evaluatig the rule base. In the light of these aims, FuzzyArtemis can be improved.

## REFERENCES


Abu-halaweh, N.M., & Harrison, R.W. (2010). An improved fuzzy decision tree induction tool. *Fuzzy Information Processing Society (NAFIPS), 2010 Annual Meeting of the North American*.

Alcala-Fdez, J., Fernandez, A., Luengo, J., Derrac, J., Garcia, S., Sanchez, L., & Herrera F. (2009). KEEL data-mining software tool: Data set repository, integration, of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic and Soft Computing,* 13 (1), 959-977.

Aroca-Santos, R., Cancilla, C.C., S. Pariente & E. S., Torrecilla, J. S. (2016). Neural networks applied to characterize blends containing refined and extra virgin olive oils. *Talanta*, 161, 304-308.

Bandler, W., & Kohout, L. (1980). Fuzzy power sets and fuzzy implication operators. *Fuzzy Sets and Systems,* 4, 13-30.

Bezdek, J.C. (1974a). Cluster validity with fuzzy numbers. *Journal Cybernetics*, 58-73.

Bezdek, J.C. (1974b). Numerical taxonomy with fuzzy sets. *Journal of Mathematical Biology*, 1, 57-71.

Bezdek, J.C. (1981). *Pattern recognition with fuzzy objective function algorithms*. New York: Springer US.

Binetti, G., Coco L.D., Ragone R., Zelasco, S., Perri, E., Montemurro, C., Valentini, R., Naso, D., Fanizzi, F., & Schena, F.P. (2017). Cultivar classification of Apulian olive oils: Use of artficial neural networks for comparing NMR, NIR and merceological data. *Food Chemistry*, 219, 131-138.

Bordogna, G., & Passi, G. (1993). A fuzzy linguistic approach generalizing boolean information retrieval: a model and its evaluation. *Journal of the American Society of Information Science and Technology*, 44, 70-82.

Breiman, L., Friedman, J., Olshen, R. & Stone, C. (1984). *Classification and regression trees:* London: Chapman & Hall.

Chang, R.L.P., & Pavladis, T. (1977). Fuzzy decision tree algorithms. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-7, 28-35.

Chang, P.C., Fan, C.Y., & Lin, J.L. (2011). Trends discovery in financial time series data using a case based fuzzy decision tree. *Expert Systems with Applications*, 38 (5), 6070-6080.

Chiang, I. & Hsu, J. (2002). Fuzzy classification trees for data analysis. *Fuzzy Sets and Systems*, 130, 87-99.

Coppi, R, D'Urso, P., & Giordani, P. (2012) Fuzzy and possibilistic clustering for fuzzy data, *Computational Statistics and Data Analysis*, 56, 915-927.

*Datasets*. (n.d.). Retrieved January 1, 2014, from http://sci2s.ugr.es/keel/datasets.php.

Delgado, M., Verdegay, J.L., & Vila, M.A. (1992). Linguistic decision making models. *International Journal of Intelligent Systems*, 7 (5), 479-492.

Dietterich, T.G., Hild, H., & Bakiri, G. (1990). A comparative study of ID3 and backpropagation for English text-to-speech mapping. *Proceedings of the 1990 Machine Learning Conference*, 24-31.

Dombi, J. (1982). A general class of fuzzy operators, the De Morgan class of fuzzy operators and fuzziness induced by fuzzy operators. *Fuzzy Set and Systems*, 8, 149-163.

Dubois, D., & Prade, H. (1986). *New results about properties and semantics of fuzzy set-theroetic operators*. Fuzzy Sets, New York: In: P.P. Wang and S.K. Chang, Edition, Plenum Press.

Dunn J.C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated Clusters. *Journal of Cybernetics*, 3, 32-57.

Dunn, J. C. (1974) Well separated clusters and optimal fuzzy partitions. *Journal of Cybernetics*, 4, 95-104.

D'Urso, P., & Giordani, P. (2006) A weighted fuzzy c-means clustering model for fuzzy data, *Computational Statistics and Data Analysis, 50* (6), 1496-1523.

Elkano, M., Galar, M., Sanz, J.A., Fernández, A., Barrenechea, E., & Herrera, F. (2015). Enhancing multiclass classification in FARC-HD Fuzzy Classifier: On the Synergy Between n-dimensional overlap functions and decomposition strategies. *IEEE Transactions on Fuzzy Systems*, 23 (5), 1562-1580.

*European Commision Regulation*. (1991). Retrieved June, 2012, from http://publications.europa.eu/en/publication-detail/-/publication/a55c747d-7176-4776-9b16-6eb913087fc1/language-en.

*European Commision Regulation*. (1997). Retrieved June, 2012, from http://publications.europa.eu/en/publication-detail/-/publication/725d3e60-bc78-4342-aec8-91745b7714ee/language-en.

Farahbod, F. & Eftekhari, M. (2012) Comparison of different T-norm operators in classification problems. *International Journal of Fuzzy Logic Systems, 2* (3).

Giles, R. (1976) Lukasiewicz logic and fuzzy set theory. *International Journal of Man-Machine Studies*, 8, 313-327.

Gumuşkesen, S. & Yemiscioglu F. (2007) Project Name: Türkiye'deki zeytin çeşitlerinin ve zeytinyağlarının bölgesel olarak Karakterizasyonu, Project Number: 2005/BİL/020. Retrieved February 24, 2016, from http://food.ege.edu.tr/d-83/akademikyapi.html.

Gupta, M. M., & Qi, J. (1991). Theory of T-norms and fuzzy inference methods. *Fuzzy Sets and Systems,* 40, 431-450.

Harrington, P. B. (1991). Fuzzy multivariate rule-building expert systems: minimal neural networks. *Journal of Chemometrics*, 5, 467-486.

Harrington, P. B. (1993). Minimal neural networks: differentiation of classification entropy. *Chemometrics and Intelligent Laboratory Systems*, 19, 143-154.

Harrington, P. B. (2017). Support vector machine classification trees based on fuzzy entropy of classification. *Analytica Chimica Acta*, 954, 14-21.

Harrington, P. B., Kister, J., Artaud, J., & Dupuy, N. (2009). Automated Principal Component-Based Orthogonal Signal Correction Applied to Fused Near Infrared-Mid Infrared Spectra of French Olive Oils. *Analytical Chemistry*, 81 (17) 7160-7169.

Hsu, S., Hsu, J. Y., & Chiang, I. (1995). Automatic generation of fuzzy control rules by machine learning methods. *In Proceedings of 1995 IEEE International Conference on Robotics and Automation*, 287–292.

Ishibuchi, H., Nakashima, T., & Nii, M. (2004). *Classification and modeling with linguistic information granules: advanced approaches to linguistic data mining*. New York: Springer-Verlag.

Jang, J.S:R., Sun, C.T., Mizutani, E. (1997). *Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence.* Upper Saddle River: Prentice Hall.

Janikow, C. Z. (1996). Exemplar learning in fuzzy decision trees. *In Proceedings of IEEE International Conference on Fuzzy Systems & International Fuzzy Engineering Symposium*, 1500–1505.

Janikow, C. Z. (1998). Fuzzy decision trees: issues and methods. *IEEE Transactions on Systems, Man, Cybernetics*, 28, 1–14.

Kantarcı, S. (2010). *Multi expert decision making by using 2 tuple fuzzy linguistic representation and its application to olive oil sensory evaluation*. Master Thesis, Dokuz Eylül University, İzmir.

Kantarcı, S., Turanoğlu, E., & Ulutagay, G. (2013). A comparison of decision tree algorithms for the classification of thyroid disease. *Seventh International*

*Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control*, 141-148.

Kantarcı, S., Vahaplar, A., Kınay, A.Ö., Nasiboğlu, E. (2015a). Influence of different T-operators in fuzzy decision trees. *16. Uluslararası Ekonometri, Yöneylem Araştırması ve İstatistik Sempozyumu (EYİ2015)*.

Kantarcı, S., Vahaplar, A., Kınay, A.Ö., Nasiboğlu, E. (2015b). Influence of different T-norm and T-conorm operators in fuzzy decision trees. *In Proceddings of 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE),* 1-6.

Kantarcı, S., Vahaplar, A., Kınay, A. O., & Nasibov, E. (2013). A software (OliveDeSoft) system for geographic characterization of Turkish virgin olive oil via fuzzy decision tree. *In 3rd International Fuzzy Systems Symposium*, 198-202.

Kinoshita, J. (1994). Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems. *In Proceedings of the third IEEE Conference on Fuzzy Systems,* 3, 2113-2118,

Larose, D. T. (2005). *Discovering knowledge in data: An introduction to data mining.* Haboken: John Wiley &Sons Inc.

Lee, K., Lee, K., Lee, J., & Lee-Kwang, H. (1999). A fuzzy decision tree induction method for fuzzy data. *In Proceedings IEEE Conference on Fuzzy Systems*, 16-25.

Liu, X., & Pedrycz, W. (2007). The development of fuzzy decision trees in the framework of axiomatic fuzzy set logic. *Applied Soft Computing*, 7, 325-342.

Martínez, L., Espinilla, M., & Perez, L. G. (2008). A linguistic multigranular sensory evaluation model for olive oil. *International Journal of Computational Intelligence Systems*, 1 (2), 148-158.

Menger, K. (1942). Statistical metrics. *Proceedings of the National Academy of Sciences*, U.S.A. 28, 535-537.

Miller, G. A. (1955). The magical number seven or minus two: some limits on our capacity of processing information. *Psychological Review*, 101 (2), 343-352.

Nakashima, T., Schaefer, G., Yokota, Y., & Ishubuchi, H. (2007). A weighted fuzzy classifier and its application to image processing tasks. *Journal of Fuzzy Sets and Systems*, 158 (3), 284-294.

Nasiboğlu, E. Kantarci S., Kınay A.Ö., & Vahaplar A. (2013). Geographic Classification of Turkish Natural Olive Oil via CART Decision Tree Algorithm Achieved from Chemical Data. *XIV. Ekonometri, İstatistik ve Yöneylem Araştırması ve İstatistik Sempozyumu*, 210.

Nasibov, E.N. (2002) Certain Integral Characteristics of Fuzzy Numbers and a Visualinteractive Method for Choosing the Strategy of their Calculation, *Journal of Computer and Systems Sciences Internation*al, 41, 4, 584-590.

Nasibov, E.N., Baskan, O., & Mert, A. (2005) A learning algorithm for level sets weights in weighted level-based averaging method. *Fuzzy optimization and decision making*, 4, 279-291.

Nasibov, E.N, & Mert, A. (2007). On methods of defuzzification of parametrically Represented Fuzzy Numbers. *Automatic Control and Computer Sciences,* 41 (5), 265-273.

Nasibov, E. & Kantarci, S. (2010). Çok tanecikli bulanık sözel bilgiye dayalı karar verme modeli ve zeytinyağının duyusal analizine uygulanması. *11. Ekonometri ve İstatistik Sempozyumu*.

Nasibov, E. Kantarci, S., Vahaplar, A., & Kınay, A.Ö. (2013). Geographic classification of Turkish natural olive oil via decision tree algorithms: a comparative analysis. *3rd International Symposium on Computing in Science & Engineering*, 206-210.

Nasibov, E., Kantarcı Savaş, S., Vahaplar, A., & Kınay, A.Ö. (2016). Geographic classification of virgin olive oil with using T-operators in fuzzy decision tree approach. *Chemometrics and Intelligent Laboratory Systems*, 155, 86-96.

Oussallah, M. (2003). On the use of Hamacher’s t-norms family for information aggregation. *Information Sciences*, 153, 107-154.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, (1), 81-106.

Quinlan, J. R. (1990). Decision trees and decision making. *IEEE Transactions on Systems., Man, Cybernetics,* 20, 339-346.

Quinlan, J. R. (1993). *C4.5: Programs for machine learning*: San Mateo: Morgan Kaufmann.

Ross, T. J. (2010). *Fuzzy logic with engineering applications* (3rd ed.). New York: John Wiley & Sons.

Sanz, J.A., Bustince, H., Fernández, A., & Herrera, F. (2012). IIVFDT: Ignorance functions based interval-valued fuzzy decision tree with genetic tuning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems,* 20 (2), 1-30.

Sanz, J. A., Galar, M., Jurio, A., Brugos, A., Pagola, M. & Bustince, H. (2014). Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule based classification system. *Applied Soft Computing*, 20, 103-111.

Sap, M. N. M., & Khokhar, R. H. (2004). Fuzzy decision tree for data mining of time series stock market databases critical assessment of microarray data analysis. *Critical Assessment Microarray Data Analysis (CAMDA).*

Sarle, W.S., & Warren, S. (1994). Neural networks and statistical models. *In Proceedings of* the *Nineteenth Annual SAS Users Group International Conference,* 1-13.

Schweizer, B., & Sklar, A. (1983). Probabilistic metric spaces (North-Holland, Amsterdam).

Tokumaru, M., & Muranaka, N. (2009). Product-impression analysis using fuzzy C4.5 decision tree. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 13, 6.

Ulutagay, G., & Kantarci S. (2014). Classification with fuzzy OWA distance. *2014 International Conference on Fuzzy Theory and Its Applications (iFUZZY2014),* 195-198.

Ulutagay, G., & Kantarci, S. (2013). Classification of LR fuzzy data using OWA distance. *In 3$^{rd}$ International Fuzzy Systems Symposium*, 186-192.

Ulutagay, G., & Kantarcı, S. (2015). An extension of fuzzy L-R data classification with fuzzy OWA distance. *International Journal of Intelligent Systems,* 30 (9), 1006-1020.

Umano, M., Okamoto, H., Hatono, I., Tamura, H., Kawachi, F., Umedzu, S., & Kinoshita J. (1994). Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems. *In Proceedings of the third IEEE Conference on Fuzzy Systems*, 2113-2118.

Vahaplar, A., Kantarci, S., Kınay, A.Ö., & Nasiboğlu, E. (2013). Classification of virgin olive oil by using fuzzy ID3 algorithm. *Seventh International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control,* 275-283.

Weber, S. (1983). A general concept of fuzzy connectives. negations and implications based on t-norms and t-conorms. *Fuzzy Sets and Systems,* 11, 115-134.

Yager, R.R. (1980). On a general class of fuzzy connectives. *Fuzzy Sets and Systems,* 4, 235-242.

Yuan, Y., & Shaw, M. J. (1995). Induction off fuzzy decision trees. *Fuzzy Sets and Systems,* 69, 125-139.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8, 338-353.

Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning -I. *Information Sciences*, 8, 199-249.

Zimmerman, H.J. (2001). Fuzzy set theory and its applications. Kluwer Academic Press. Dordrecht.