Chemometrics and Intelligent Laboratory Systems 155 (2016) 86–96

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

A survey on geographic classification of virgin olive oil with using T-operators in Fuzzy Decision Tree Approach

Efendi Nasibov [a,b,1], Suzan Kantarcı Savaş [a,*], Alper Vahaplar [a], Ayşe Övgü Kınay [a]

[a] Department of Computer Science, Dokuz Eylul University, İzmir 35160, Turkey
[b] Institute of Control Systems, National Academy of Sciences, Baku AZ-1141, Azerbaijan

ARTICLE INFO

Article history:
Received 7 November 2014
Received in revised form 4 April 2016
Accepted 8 April 2016
Available online 18 April 2016

Keywords:
Pattern classification
Machine learning
Fuzzy ID3 algorithm
T-operators
Geographic classification

ABSTRACT

Olive oil is a crucial agricultural food product from past to present. The quality control of this product is too difficult. The geographic classification of it has an importance for the countries in order to provide the traceability. This paper aims to present a classification system for the geographic classification of virgin olive oil based on chemical parameters which contain uncertainty. Proposed system constructs the rules by using fuzzy decision tree algorithm. This algorithm builds rules by using ID3 algorithm with fuzzy entropy on the fuzzified data. The reasoning procedure based on rule based classification is handled with different T-operators. Fuzzy *c*-means algorithm is used in order to fuzzify the olive oil data set. The cluster numbers of each variable are decided according to partition coefficient validity criteria. The model is examined by using different decision tree approaches (C4.5 and standard version Fuzzy ID3 algorithm). The quality of proposed FID3 reasoning method with nine different T-operators is analyzed by using accuracy rates handled with 20 threshold values. Also, the conclusions are supported by statistical analysis. Experimental results show that fuzzy reasoning method has a crucial manner for the geographic classification. The classification system can perform better performance via different parameters for parametric T-operators.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Olive oil is an important agricultural food product as a fruit juice. It is too difficult to control the quality of this product because it is influenced by several matters such as olive type, climate conditions, soil structure, maturation index, and the techniques for extraction of olive oil. The standard procedures are published by European Union and International Olive Council (COI) in order to follow the quality of it. It is seen that there is a need in improving the analytical methods in order to control the geographic origin of olive oils for supporting the protection of denominated protected origin (DPO) policy. The pedoclimatic and process conditions influence the quality of this product. Also, the different cultivars hold different chemical structures. Geographic classification problem is an important issue for the traceability of the olive oil quality. It is necessary to investigate the relationship among the chemical parameters for each region.

Machine learning is a discipline that is interested in the design and development of algorithms for computers that derive behaviors based on empirical data. The aim of the machine learning is to recognize patterns and get some intelligent knowledge based on data without assumptions. Relations among observed variables can be illustrated by using machine learning algorithms. The measurements used for chemistry in several studies have uncertainty [1–5]. Decision trees can be used as a machine learning tool for the geographic classification problem which helps to assign a new olive oil sample into a region. This method is highly interpretable nonparametric analysis method [6,7]. Nowadays, fuzzy decision trees have been still studied actively by many researchers [8–22]. In [9], a new algorithm called Fuzzy ID3 was proposed to generate a fuzzy decision tree. Fuzzy decision tree is constructed in an inductive way. It consists of nodes for testing attributes, edges for branching by test values of fuzzy sets and leaves for deciding class according to class membership.

The main aim of this paper is to report a methodological approach for the classification of olive samples based on Fuzzy ID3 approach. Also, this study is interested in fuzzy reasoning method of Fuzzy ID3 approach. 9 different T-operators are examined in order to make reasoning by using fuzzy rule based system of fuzzy decision tree. This study is examined on 101 virgin olive oils samples collected from four different regions (North Aegean, South Aegean, Mediterranean, and South East) by using measurements of chemical parameters. The data set is normalized by using min–max normalization. The parametric methods are not preferred for the statistical analysis because of the data structure. The explanatory data analysis was performed by using Principal Component Analysis (PCA). The performance of nine different

* Corresponding author.
E-mail addresses: efendi.nasibov@deu.edu.tr (E. Nasibov), suzankantarci@gmail.com (S. Kantarcı Savaş), alper.vahaplar@deu.edu.tr (A. Vahaplar), ovgu.kinay@deu.edu.tr (A.Ö. Kınay).
[1] Present address: Department of Computer Science, Dokuz Eylul University, İzmir 35160, Turkey.

http://dx.doi.org/10.1016/j.chemolab.2016.04.004
0169-7439/© 2016 Elsevier B.V. All rights reserved.

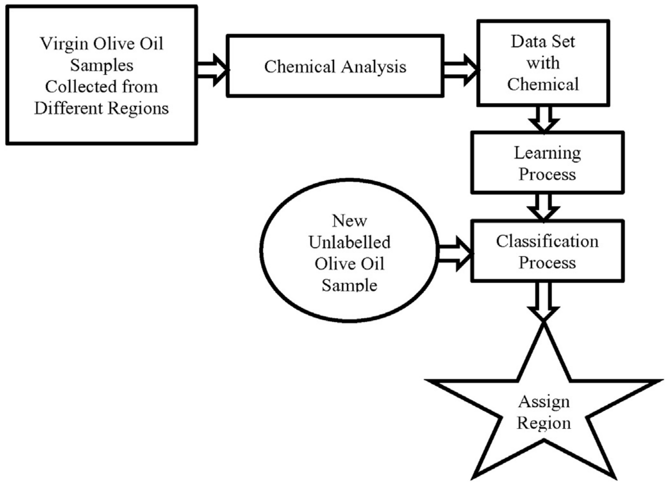

Fig. 1. Geographic classification problem scheme.

T-operators was compared with C4.5, the standard version of Fuzzy ID3 algorithm. Leave-one-out procedure is performed in order to measure the performances of the algorithms. The Friedman aligned rank test and pairwise comparisons were performed to evaluate fuzzy reasoning method based on different T-operators. The rest of the paper is organized as follows: Section 2 presents the problem definition and related works. The methods are given in Section 3. Experimental study and the conclusion are given in Section 4, Section 5, respectively.

2. Problem definition and related works

Geographic classification problem has multivariate structure. This problem aims to find the class of an unassigned sample. It is necessary to research the geographic origin of olive oils for supporting the protection of denominated protected origin (DPO) policy. The principal aim is to support the tackling authenticity controls of olive oil qualities by using modern analytical tools. In the line of this aim, many investigators concern with the geographic classification of olive oils [23–33]. The types of extra virgin olive oil have been confirmed by using Linear Discriminant Analysis (LDA) and Back Propagation Artificial Neural Networks (BP-ANN) in [24]. In [25], the adulteration in olive oil was defined by near-infrared spectroscopy and using chemometric techniques such as principal component analysis, partial least squares regression (PLS) and applied methods for data pretreatments such as signal detection correction. The scholars were interested in the classification of olive oils using high throughout flow ^{1}H NMR fingerprinting with Principal Component Analysis, Linear Discriminant Analysis and Probabilistic Neural Networks in [26]. In [27], fatty acid composition, phenolic compounds, and total sterols are evaluated in order to estimate the value of a new sample by using Classification Binary Tree.

Principal Component Analysis and SIMCA Classification Model are used for the statistical analysis of the data in [28]. The data sets included fatty acid composition and data provided from Fourier transformed infrared spectroscopy. As a result of this study, it was seen that fatty acid composition was an effective composition for classifying the olive oil samples.

In literature, it is seen that the decision tree algorithms are used in chemometrics [3–5]. In [3], the use of partial derivatives by minimal neural networks improves both efficiency and efficacy over the numerical computations. The expressions for the partial derivatives of fuzzy entropy of classification have been presented. Hence, an expert system is defined in [4]. These rules are constructed using ID3 algorithm with a fuzzy expression of classification entropy. This expert system is called as FURES (Fuzzy Multivariate Rule-Building Expert System). This system uses local processing which furnishes qualitative information in the rule structure of the classification trees and variable loadings of the weight vectors. Also, fuzzy decision tree approach was used in [5] in order to compare with PLSD-DA for classification of French olive oils for geographic regions. Fuzzy logic developed by Zadeh in mid-1960's is an useful tool in order to deal with the information with including uncertainty [34]. In chemometrics, it can be useful to find answers in different kind of classification problems which contain itself uncertainty. In literature, different kinds of fuzzy decision tree approaches are as given in [10]. In this approach, fuzzy decision tree is constructed by reducing the classification ambiguity. Another fuzzy decision tree approach with a number of reasoning procedures based on conflict resolution in rule based systems and efficient approximate reasoning methods were presented in [11]. An adaptable software system called as Salammbô to construct fuzzy decision trees and a model was introduced in order to study discrimination measures [12]. C-fuzzy decision trees and this new version works with intrusion detection system [13]. Fuzzy decision tree presented in [14] used a type of Fuzzy ID3 algorithm. In [15], an implementation about fuzzy decision tree for user modeling was described. The system helps the inexperienced user with the construction of fuzzy decision trees. Hence, there are many more real life applications of fuzzy decision trees in different kinds of area such as

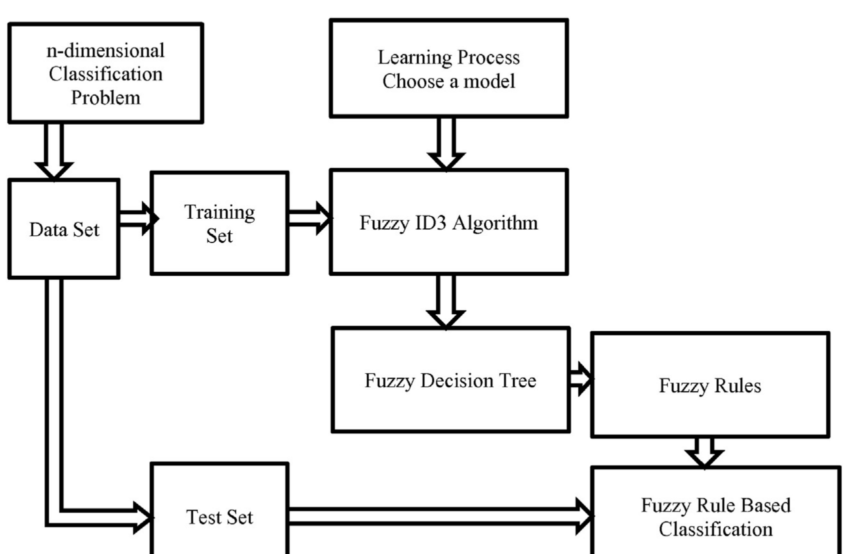

Fig. 2. A classification problem with Fuzzy ID3 algorithm combined with FRBC.

Table 1
T-Operators used in fuzzy reasoning method.

Non-parametric operators

Ref	T-norm operators	T-conorm operators
Zadeh [34]	$T_1(x,y) = \min(x,y)$	$T_1^*(x,y) = \max(x,y)$
Product Sum [52,53]	$T_2(x,y) = x.y$	$T_2^*(x,y) = x+y-xy$
Giles [54]	$T_3(x,y) = \max(x+y-1,0)$	$T_3^*(x,y) = \min(x+y,1)$
Nonparametric Hamacher [55]($\lambda=0$)	$T_4(x,y) = \frac{xy}{(x+y-xy)}$	$T_4^*(x,y) = \frac{x+y-2xy}{1-xy}$

Parametric operators

Ref	T-norm operators	T-conorm operators	Parametric range
Hamacher [55]	$T_5(x,y) = \frac{xy}{\lambda+(1-\lambda)(x+y-xy)}$	$T^*{}_5(x,y) = \frac{x+y-(2-\lambda)xy}{\lambda+(1-\lambda)(1-xy)}$	$\lambda \geq 0$
Yager [56]	$T_6(x,y) = \max(1-((1-x)^p+(1-y)^p)^{1/p},0)$	$T_6^*(x,y) = \min((x^p+y^p)^{1/p},1)$	$p=(0,1)$
Dombi [57]	$T_7(x,y) = \frac{1}{1+((\frac{1}{x}-1)^{\lambda}+(\frac{1}{y}-1)^{\lambda})^{1/\lambda}}$	$T_7^*(x,y) = \frac{1}{1+((\frac{1}{x}-1)^{-\lambda}+(\frac{1}{y}-1)^{-\lambda})^{-1/\lambda}}$	$\lambda=(0,1)$
Dubois&Prade [50]	$T_8(x,y) = \frac{xy}{\max(x,y,\lambda)}$	$T_8^*(x,y) = 1-\frac{(1-x)(1-y)}{\max(1-x,1-y,\lambda)}$	$\lambda=(0,1)$
Weber [52]	$T_9(x,y) = \max(\frac{x+y-1+\lambda xy}{1+\lambda},0)$	$T_9^*(x,y) = \min(x+y+\lambda xy,1)$	$\lambda=(0,1)$

medical, multimedia, and chemical applications in [16]. In fuzzy decision tree approach, reasoning procedure has very crucial manner. Geographic classification problem scheme is given in Fig. 1. In our problem, we are interested in Geographic classification of virgin olive oils samples collected from four different regions in Turkey. In learning process, we use Fuzzy ID3 algorithm proposed in [9]. Hence, we propose to adapt different T-operators into the fuzzy reasoning method based on rules generated from fuzzy decision tree in [9]. Fuzzy rule based classification system given in [35] is combined with fuzzy decision tree's reasoning approach.

3. Methods

We briefly explain Principal Component Analysis (PCA), C4.5 Algorithm. Then, we describe Fuzzy logic and Fuzzy c-means algorithm as fuzzification tool. Also, we review briefly Fuzzy ID3 builder combined with Fuzzy Rule Based Classification and its reasoning method. We talk about T-operators. At last, we propose fuzzy ID3 reasoning method by using different T-operators.

3.1. Principal Component Analysis (PCA)

PCA is a kind of projection method. It reduces the dimension in order to transform the original measurement variables into novel and uncorrelated variables. These novel variables are called as principal components. They preserve as much as possible of the information existed in the original data. Each component is a linear combination of the original variables. The direction of greatest variance in the data set is represented by using the orthogonal axes. The score plots helps in order to show this situation [36].

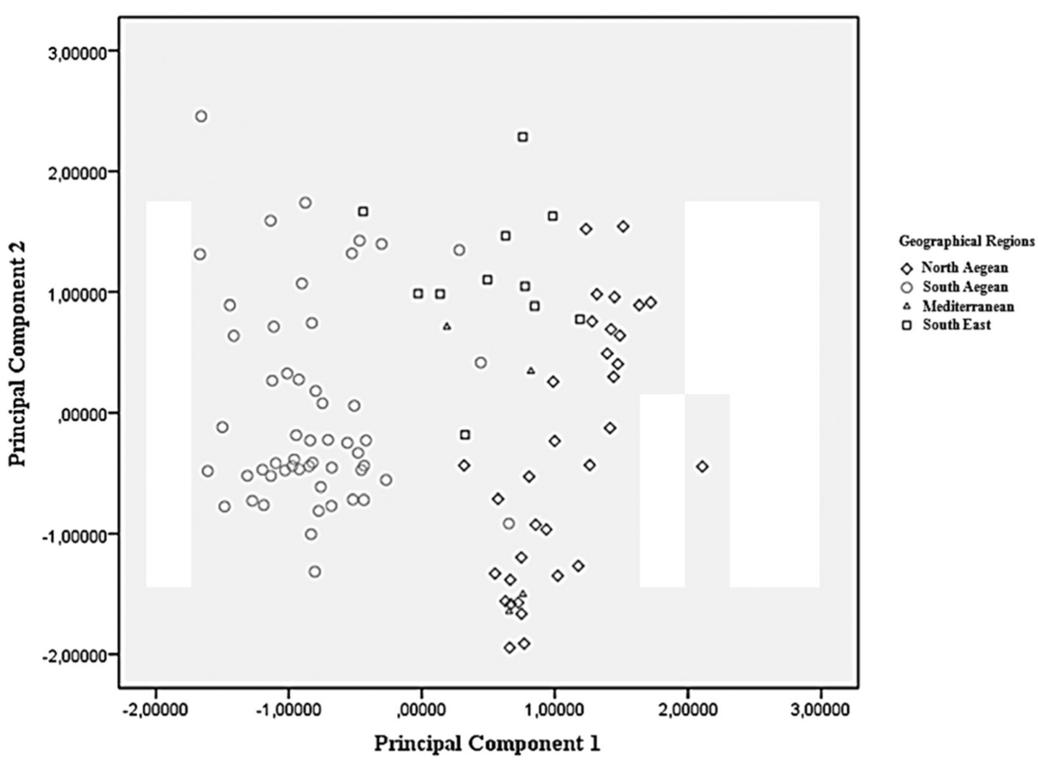

Fig. 3. Principal components plot on the virgin olive oil samples.

3.2. C4.5 Algorithm

C4.5 aims to use divide-and-conquer strategy [37]. This algorithm is an improved version of ID3 algorithm [38]. Maximum gain information is used as a splitting criterion by the algorithm. It calculates overall entropy of the training data and entropy for partitioned dataset according to each attribute. Then it takes the differences between the entropies. This value is called as the gain information. Then, the attribute which has the highest gain information is selected for splitting. The formulas for entropy (Eq. (1)) and gain information (Eq. (2)) are given below, respectively:

$$Entropy(S) = -\sum_{j=1}^{m} \frac{freq(C_j, S)}{|S|} . \log_2 \frac{freq(C_j, S)}{|S|} \quad (1)$$

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^{k_A} \frac{|S_i|}{|S|} Entropy(S_i) \quad (2)$$

where S is a training set and $C = \{C_1, C_2, \ldots, C_m\}$ defines m classes. S_i, $i = 1, \ldots, k_A$, are subsets of S according to domain of the attribute A. It uses pruning for the construction of the tree in order to avoid over-fitting problem.

3.3. Fuzzy logic and fuzzy c-means algorithm as fuzzification tool

A fuzzy set is a class of objects with a continuum of grades of membership. In fuzzy set theory, a fuzzy subset of the universe of discourse U is described by a membership function $\mu_v(V) : U \rightarrow [0,1]$, which represents the degree to which $u \in U$ belongs to the set v. Fuzzification can be explained as the process of transformation crisp values into membership degrees for the linguistic terms of fuzzy sets. Generally, fuzzy membership functions are used for the fuzzification process, such as triangular membership functions. Trapezoidal membership functions, etc. On the other hand, Fuzzy c-means (FCM) is another way in order to get the membership degrees for fuzzy variables. The idea of this algorithm was suggested in [39] and it was improved in [40]. The algorithm aims to determine a fuzzy c partition matrix U. In order to achieve this aim, it minimizes an objective function. The objective function J_m is defined as follows for fuzzy partition (Eq. (3)):

$$J_m(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_{ik})^m (d_{ik})^2 \quad (3)$$

Where

$$d_{ik} = d(x_k, v_i) = \left[\sum_{j=1}^{p} (x_{kj} - v_{ij})^2 \right]^{1/2}, k = 1, \ldots, n; i = 1, \ldots, c. \quad (4)$$

and μ_{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 $m \in (1, \infty)$ reflects sharpness of the fuzzification process. In Eq. (4), d_{ik} shows any distance measure (usually the Euclidean distance) between k^{th} data point and i^{th} cluster center in p dimensional space. Also, v_i shows the i^{th} cluster center. Then, Eq. (5) is used in order to calculate each of the clusters centers for each class:

$$v_{ij} = \frac{\sum_{k=1}^{n} \mu_{ik}^m x_{kj}}{\sum_{k=1}^{n} \mu_{ik}^m}, i = 1, \ldots, c; j = 1, \ldots, p. \quad (5)$$

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

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

Membership degrees are calculated according to the Eq. (6):

$$\mu_{ik} = \frac{1}{\sum_{z=1}^{c} \left(\frac{\|x_k - v_i\|}{\|x_k - v_z\|} \right)^{\frac{2}{m-1}}}, \quad i = 1, \ldots, c; k = 1, \ldots, n \quad (6)$$

The determination of the correct number of clusters (c) for the FCM algorithm has a crucial issue. In literature, there are some scalar measures of partitioning fuzziness, called *validity indicators* [41–43]. Partition coefficient is a scalar measure as formulized as below (Eq. (7)):

$$V_{PC} = \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^2 \quad (7)$$

whereas optimal cluster number is max (V_{PC}, U, c).

The cluster number of each attribute shows the number of linguistic terms of each fuzzy attribute. At the end of this process, we handle fuzzy attributes with fuzzy linguistic terms whose number decided by using optimal cluster number as Eq. (7).

Table 3
The performance results of each algorithm for non-parametric operators.

Algorithms	Accuracy rate (%)
C4.5	86.14
FuzzyID3_reasoning with Product Sum_Umano	86.14
FuzzyID3_ reasoning with Zadeh T-Opeators T_1 & T_1^*	86.14
FuzzyID3_ reasoning with Product-Sum T_2 & T_2^*	86.14
FuzzyID3_ reasoning with Non Parametric Hamacher(λ = 0)T_3 & T_3^*	86.14

Table 4
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. valuesT_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

3.4. Fuzzy ID3 builder adapted with fuzzy rule based classification system

Fuzzy Rule Based Classification System (FRBCS) has a crucial manner in the field of pattern recognition and classification problems. These systems have computational flexibility in order to use linguistic labels in the antecedents of their rules. They have been implemented different type of problems in real life, such as image processing [44], medical problems [45], etc.

A classification problem needs a set of training samples. These samples are called as a training set. It is necessary a mapping function called as classifier in order to learn the model from the training samples. The model provides the class of a new sample.

Let a training set consists of p samples. $x_p = (x_{p1}, \ldots, x_{pn})$ be the p^{th} sample of the training set where x_{pi} is the value of the i^{th} attribute $(i = 1, 2, \ldots, n)$ of the p^{th} training sample. Classification problem is a supervised learning problem. So each sample belongs to a class shown as $y_p \in C = \{C_1, C_2, \ldots, C_m\}$, where m is the number of classes of the problem [35].

Two main components of FRBCSs are summarized in [35]:

3.4.1. Knowledge base

It contains the rule base (RB) and the database. The rules and the membership functions are stored in this base.

3.4.2. Fuzzy reasoning method

In this mechanism, the samples are classified by using the information stored in the knowledge base.

In this paper, we focus on Fuzzy Interactive Dichotomizer 3 (Fuzzy ID3) algorithm as a classifier. Its' structure uses a kind of fuzzy rule based learning algorithm for making the reasoning. In order to generate rules, Fuzzy ID3 algorithm constructs a tree in learning process. To do so, fuzzy entropy is applied to find the attributes which has the maximum information whereas minimum uncertainty. Each path of the tree shows the rules. There are Rule Weights (RW) for each classes at each leaf node. RW_j shows j^{th} rule weight handled from fuzzy confidence value CF_j which is equal to RW_j.

The solution of a classification problem with Fuzzy ID3 algorithm combined with Fuzzy Rule Based Classification System can be summarizedin Fig. 2. as follows:

3.4.3. Fuzzy Interactive Dichotomizer 3

Fuzzy decision tree is a kind of classification tool which is well suited to the geographic characterization problem of olive oil. A tree is generated and the decision rules are achieved by using each path from the root to the leaves of the tree. In this context, the tree builder researches the paths among the chemical parameters of the olive oil samples collected from different regions. Fuzzy Interactive Dichotomizer 3 (Fuzzy ID3) defined in [9] is used as a tree builder algorithm. It is fuzzified version of ID3 algorithm proposed by Quinlan in [38]. It deals with crisp and fuzzy variables defined by the user. This algorithm divides the data set according to a data attribute. This data attribute is selected by using a measure called as information gain based on fuzzy entropy. It means that it searches the attributes which has the information with the highest quality. The algorithm is given below:

Assume there are N labeled fuzzified patterns and n attributes $A = \{A_1, A_2, \ldots, A_n\}$. For each k assume that $(1 \le k \le n)$. The attribute A_k takes m_k values of fuzzy subsets $(A_{k1}, A_{k2} \ldots, A_{km_k})$. C denotes the classification target attribute, taking m values $C_1, C_2, \ldots, C_m$. The symbol $M(.)$ is used to denote the cardinality of a given fuzzy set, that is, the sum of the membership values of the fuzzy set [9,21].

The induction process of fuzzy ID3 is given as follows:

Step 1: Generate a root node which has a set of all data. This data set is fuzzified data set, 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 labelA_{ki}.$(i = 1, 2, \ldots, m_k)$, compute its relative frequencies with respect to class C_j $(j = 1, 2, \ldots, m)$.

$$p_{ki}(j) = \frac{M(A_{ki} \cap C_j)}{M(A_{ki})} \tag{8}$$

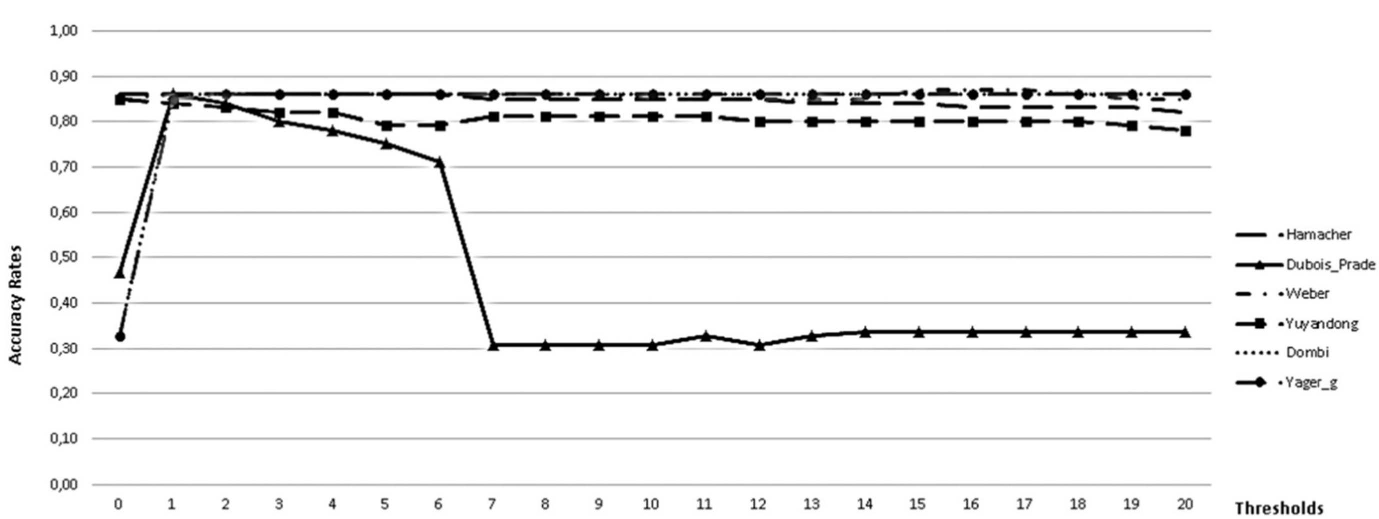

Fig. 4. Accuracy rates handled from different parametric operators. Range = (0–20) and $\theta_r = 0.75$.

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

$$Entr_{ki} = \sum_{j=1}^{m} -p_{ki}(j)\log(p_{ki}(j)) \quad (9)$$

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

$$E_k = \sum_{i=1}^{m_k} \frac{M(A_{ki})}{\sum_{j=1}^{m_k} M(A_{kj})} Entr_{ki} \quad (10)$$

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

$$Atr = \arg\max_{1 \le k \le n}(G_k) \quad \text{where} \quad G_k = E_k - Entr_{ki} \quad (11)$$

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. Delete the branch if it is empty. If the branch is non-empty, compute the relative frequencies by using (Eq. (8)) of all objects within the branch into each class. If the relative frequency of each class is above the given threshold θ_r or all the attributes have been expanded for this branch. Terminate the branch as a leaf. Otherwise, select the attribute from among those which have not been expanded yet in this branch with the smallest average fuzzy classification entropy. (Eq. (11)) as a new decision node for the branch and add its linguistic labels as candidates branches to analyze. At each leaf, each class will have its relative frequency.

Step 4: Repeat Step 3 while there are branches to analyze. If there are no candidate branches the decision tree is complete.

3.4.3.1. The rule structure generated from each branch of the fuzzy decision tree. After the fuzzy decision tree induction, the rules are generated from each branch. Each branch behaves as 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_j, where R_j is the label of the jth rule. $x = (x_1, ..., x_n)$ is an n-dimensional pattern vector that represents the example. 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, at each leaf node has rule weights which are computed as the relative frequency for each class (as in Step 3).

3.4.4. Reasoning (classification)

Let $x_p = (x_{p1}, ..., x_{pn})$ be the p^{th} example of the training set, which is composed of p examples, where x_{pi} is the value of the i^{th} attribute $(i = 1,2...,n)$ of the p^{th} sample. Each example belongs to a class $y_p \in C = \{C_1, C_2, ..., C_m\}$, where m is the number of classes of the problem. Assume that x_p be a new example to be classified FID3 reasoning procedure suggested in [9] applies a fuzzy reasoning method as given for FARC-HD in [35] computed in four steps. The steps are given below adapted with Fuzzy ID3 reasoning structures (This structure is called as Umano in Tables 3 and 6):

Step 1: *Matching degree*: In this step, the strenght of activation of the if-part for all rules handled from each path of the fuzzy decision tree in the RB with the pattern x_p is computed

$$\mu_{A_j}(x_p) = T\left(\mu_{A_{j1}}(x_{p1}), ..., \mu_{A_{jn_j}}\left(x_{pn_j}\right)\right) \quad (12)$$

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 (it is product in this case), and n_j is the number of antecedents of the rule.

Table 5
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
Friedman aligned ranks.

Algorithm	Rank	Friedman aligned ranks	
Zadeh	7.68		
Umano	4.72		
Product-Sum	4.72	Total N	20
Giles	4.42		
Non parametric Hamacher (λ = 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
Weber	8.85		

Step 2: *Association degree*: The association degree of the pattern x_p with each rule in the RB is computed as follows where RW_j is handled from each leaf node which is at the end of each path.

$$b_j(x_p) = \mu_{A_j}(x_p).RW_j \tag{13}$$

Step 3: *Confidence degree*: In this stage, the confidence degree for each class l is computed to obtain the confidence degree of a class, the association degrees of the rules of that class are summed [9].

$$conf_l(x_p) = \sum_{R_j \in RB; C_j = l} b_j(x_p) \qquad l = 1, 2, \ldots, m \tag{14}$$

Step 4: *Classification*: The class is obtained with the highest confidence degree assign as the predicted one [9].

$$Class = \arg\max_{l=1,\ldots,m} \left(conf_l(x_p)\right) \tag{15}$$

3.5. FuzzyID3 reasoning method based on T-operators

In the line of this study, different types of T-operators will be adapted into fuzzy ID3 reasoning method proposed in [9].

3.5.1. Overview of T-operators

T-norm and T-conorm Operators, which were developed from the triangular inequalities, are also called as T-Operators. They were originated from the studies of probabilistic metric spaces [46,47]. Fuzzy set theory combined with T-norm and T-conorm operators [48,49]. These operators are used in order to find intersection and union of two fuzzy sets. In literature, it was seen that some different types of T-operators work better in some decision making situations [50]. 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. There are different types of T-operators that may be better suited for the problems interested in.

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

$$\mu_C(x) = (\mu_A(x) \vee \mu_B(x)) \tag{16}$$

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

$$\mu_C(x) = (\mu_A(x) \wedge \mu_B(x)) \tag{17}$$

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 1. : Let $T: [0,1]x[0,1] \rightarrow [0,1]$. T is a T-norm if and only if (iff) for all $x, y, z, t \in [0,1]$:

$$T(x,y) = T(y,x) \quad \text{(commutativity)}$$

$$T(x,y) \leq T(z,t) \text{ if } x \leq z \text{ and } y \leq t \text{ and (monotonicity)}$$

$$T(x, T(y,z)) = T(T(x,y),z) \text{ (associativity)}$$

$$T(0,0) = 0;\ T(x,1) = T(1,x) = x \text{ (boundary)}$$

Definition 2. : Let $T^*: [0,1]x[0,1] \rightarrow [0,1]$. T* is a T*-conorm if and only if (iff) for all $x, y, z, t \in [0,1]$:

$$T*(x,y) = T*(y,x) \text{ (commutativity)}$$

$$T*(x,y) \leq T*(z,t) \text{ if } x \leq z \text{ and } y \leq t \text{ (monotonicity)}$$

$$T*(x, T*(y,z)) = T*(T*(x,y),z) \text{ (associativity)}$$

$$T^*(1,1) = 1;\ T(x,0) = T(0,x) = x \text{ (boundary)}$$

3.5.2. Proposed reasoning method for FuzzyID3 based on T-operators

Assume that x_p be a new example to be classified Fuzzy ID3 reasoning method suggested in [9] is given in Section 3.4.4. Now, we will give here the steps of our approach in order to make reasoning by using the rules generated from a fuzzy decision tree. The steps are given below combined with FID3 reasoning: (in our case, we use nine different types of T-norm operators [51] listed in Table 1.)

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

$$\mu_{A_j}(x_p) = T\left(\mu_{A_{j1}}(x_{p1}), \ldots, \mu_{A_j n_j}(x_{pnj})\right) \tag{18}$$

Table 7
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	Non parametric Hamacher (λ = 0)	Product Sum	Umano	Giles	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	
Giles	0.000	0.031	0.034	1.000	1.000	1.000	1.000		
Umano	0.001	0.093	0.101	1.000	1.000	1.000			
Product Sum	0.001	0.093	0.101	1.000	1.000				
Hamacher (λ = 0)	0.001	0.093	0.101	1.000					
Hamacher	0.001	0.093	0.101						
Yager	1.000	1.000							
Zadeh	1.000								

Table 8
Accuracy rates handled for different thresholds (%).

θ_r	Zadeh	Umano	Product-Sum	Giles	Non Parametric 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

where $\mu_{A_j}(x_{pi})$is the matching degree of the example with i^{th} antecedent of the rule R_j. T is a T-norm (listed in Table 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 RB is computed as follows where RW_j is handled from each leaf node which is at the end of each path, R_j. T is a T-norm (listed in Table 1.)

$$b_j(x_p) = T\left(\mu_{A_j}(x_p), RW_j\right) \tag{19}$$

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 a T-conorm (listed in Table 1.)

$$conf_l(x_p) = T^*(b_1(x_p), b_2(x_p) \ldots, b_R(x_p)), \tag{20}$$

where $b_j(x_p), j = 1, 2, \ldots, R$, is the association degree of the pattern x_p to the class l according to the j^{th} rule.

Step 4: Classification: The class is obtained with the highest confidence degree assign as the predicted one [9].

$$Class = \underset{l=1,\ldots,m}{\mathrm{argmax}}(conf_l(x_p)) \tag{21}$$

4. Experimental study

In this section, we talk about the experimental study in order to analyze the performance of our proposal given in Section 3.5.2. Firstly, we give the description of the olive oil samples and the methodology used in chemical analyses of olive oil samples. Secondly, we talk about PCA results. Thirdly, data normalization is given. Fourthly, the information is given about the implementation of fuzzy c-means algorithm.

4.1. Olive oil samples

Olives were collected from certain trees of the cultivars which were determined subject matter of this work: Ayvalik, Memecik, Kilis Yaglik, Nizip Yaglik. The samples collected in 2002–2003, 2004–2005 and

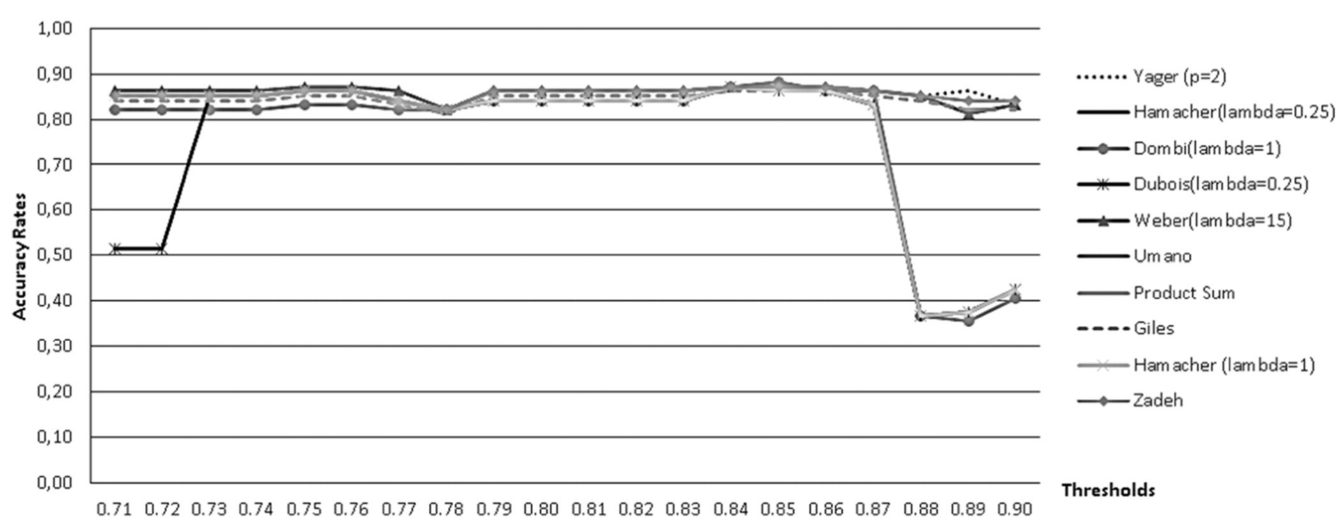

Fig. 5. Accuracy rates handled for different thresholds.

2005–2006 harvest seasons. 101 olive oil samples [30] collected from different regions (North Aegean (33), South Aegean (53), Mediterranean (4), and South East (11)) were chosen for the experimental study. The analyses of fatty acids were performed according to the official method of the European Community Regulation (1991). The olive oil samples were esterified in a methanol solution of 2N KOH for 30 min at 50 °C. 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 wirh CP-WAX 52CB (Varian) length 25 m. 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 140 °C. an initial programmed rate of 1o C/min up to 160 °C. then a second rate of 2 °C/min up to 220 °C and a final isotherm for 15 min. The injector and flame ionization detector temperatures were 250 °C. 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 check of 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 (Diario oficial de las Comunidades Europeas L 341. 12.12.1997. p.25). The apparatus was a Hewlett Packard HPLC instrument model 1100 consisted by a degasser. Quaternary pump. 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 (Diario oficial de las Comunidades Europeas L 248. 5.9.1991. p.1). 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. The software is programmed called as OliveDeSoft in the Visual C# for the experimental study (intel i7, 2.4 GHz, 4 Gb RAM).

4.2. Implementation of PCA

Principal component analysis is performed on this data set in order to explore the data structure. The principal components plot is given in Fig. 3. It is clear that there is information related to the geographic origin of virgin olive oils on the results handled 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.

4.3. Min–max normalization and fuzzy c-means algorithm

The data set was 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 our 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. (22)):

$$v' = \frac{v - \min_A}{\max_A - \min_A} \tag{22}$$

Firstly, the data fuzzification process was applied by using Fuzzy *c*-means (FCM). Partition coefficient was used in order to determine the number of clusters [42–43]. The calculated partition coefficient value for each cluster is given in Table 2.

4.3.1. Performance measure and statistical tests

The chemical measurements have imprecise information. In our study, we choose fuzzy ID3 algorithm based on fuzzy logic. 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 algorithm. Each fuzzy variable has fuzzy terms inside of it as described in Section 3.3. The clusters are determined by using partition coefficient value in Section 4.4. Our approach uses nine different T-operators into the reasoning procedure. Also, the standard version of Fuzzy ID3 represented in [9] and C4.5 [37] algorithms are 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 technique widely used in order to test different methods. This metric is defined as percentage of correctly classified samples [35]. 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.

4.3.1.1. Studying fuzzy reasoning method with non-parameteric operators. C4.5 algorithm also uses entropy as splitting criteria, like ID3 algorithm. It was improved by Quinlan in 1994 in order to deal with the numerical data [37]. The observed performance of this algorithm is 86.14%. Then, it is seen that the performance of Fuzzy ID3 algorithm with reasoning method in [9] has same performance with 86.14%. The performance results of non-parametric approaches given in Table 3 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 Giles T-operators has the minimum performance value with 85.15%.

4.3.1.2. Studying fuzzy reasoning method with parameteric operators. We consider that we want to control the performance of fuzzy reasoning method within different parameters. We believe that we can reach a better classification accuracy rate by changing the parameters value. The performance results for each Fuzzy ID3 reasoning with parametric T-operators (listed in Table 1.) are given in Table 4. 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% as same as non parametric operators.

The graph of accuracy rates handled from different parametric operators within $\theta_r = 0.75$ (0–20) are given in Fig. 4. It is supported that Fuzzy ID3 reasoning with Weber has good performance average 85.90% within range (0–20) and Hamacher has with average 84.82%.

4.3.1.2.1. Study of the Behavior of Fuzzy ID3 reasoning method based on different T-operators.. We have employed the Friedman aligned ranks as a non-parametric statistical procedure in order to detect statistical differences among a group of results for 20 threshold (θ_r) values in Table 6. This test obtains *p*-value as equal to zero, which shows that there are significant differences among the results.

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 handled given in Table 7. 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.
- Giles vs. Yager with adj. p-value = 0.034.
- Giles vs. Zadeh with adj. p-value = 0.031.
- Giles vs. Weber with adj. p-value = 0.000.
- Umano vs. Weber with adj. p-value = 0.001.
- 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 5, it is also seen that the highest average is handled from Weber. It is seen from pairwise comparisons that Weber has better results than Umano which is standart version. Weber has also better results than Hamacher. Yager has better results than Dombi and Dubois. Yager, Zadeh, and Weber have better results than Giles. As a result, Weber, which is a parameteric operator given bold as above, has better results than all non-parameteric operators.

Also, the graph of the accuracy rates are handled for different thresholds within all approaches in Fig. 5. Accuracy rates handled for different thresholds within different fuzzy reasoning method are given in Table 8. 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 behavior of threshold values will be researched.

5. Conclusion

In this study, we are interested in geographic classification of olive oil. It is one of the basic agricultural product 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 too difficult. In accordance with this paper, chemical measurements were used in order to make experimental study. Chemical measurements contain uncertainty. In order to deal with uncertain information, Fuzzy ID3 classifier was chosen for the construct the classification of olive oil samples. Additively, Fuzzy ID3 reasoning method based on T-operators has been proposed. We have aimed to see the performances of proposed fuzzy reasoning method in order to solve geographic classification problem. We have observed that the results handled from four nonparametric versions have the same performance value with handled from C4.5 algorithm. Then, we have checked the performance of parametric operators. 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 (θ_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 Umano which is standart version. Hence, Weber has better results than all non-parameteric operators' results. So, we claim that by using different parameters, we can handle better reasoning performance for fuzzy ID3. In future research, there are several works to be addressed related with the adaptation of n-dimensional overlap functions [35].

Acknowledgements

This work has been supported by Dokuz Eylul University under Scientific Research Project No. 2011.KB.FEN.035. The authors would like to thank reviewers for their valuable comments and Mrs. Ummuhan Tibet and Dr. Aytac Gumuskesen for allowing us to use the data set.

References

[1] R. Aparicio, R. Aparicio-Ruiz, Chemometrics as an aid in authentication, in: M. Jee (Ed.), Oils and Fats Authentication, Blackwell Publishing, Oxford. United Kingdom, and CRC Press, Boca Raton. FL. 2002, pp. 156–180.

[2] F. Marini, Artificial neural networks in foodstuff analyses: trends and perspectives a review, Anal. Chim. Acta 635 (2) (2009) 121–131.

[3] P.B. Harrington, Fuzzy multivariate rule-building expert systems: minimal neural networks, J. Chemom. 5 (1991) 467–486.

[4] P.B. Harrington, Minimal neural networks: differentiation of classification entropy, Chemom. Intell. Lab. Syst. 19 (1993) 143–154.

[5] P.B. Harrington, J. Kister, J. Artaud, N. Dupuy, Automated principal component-based orthogonal signal correction applied to fused near infrared-mid infrared spectra of French olive oils, Anal. Chem. 81 (17) (2009) 7160–7169.

[6] J.R. Quinlan, Induction of decision trees, Mach. Learn. 1 (1986) 81–106.

[7] L. Breiman, J. Friedman, R. Olshen, et al., Classification and Regression Trees, Chapman & Hall, London, 1984.

[8] R.L.P. Chang, T. Pavladis, Fuzzy decision tree algorithms, IEEE Trans. Syst. Man Cybern. SMC-7 (1) (1977).

[9] M. Umano, H. Okamoto, I. Hatono, H. Tamura, F. Kawachi, S. Umedzu, J. Kinoshita, Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems, Proceedings of the Third IEEE Conference on Fuzzy Systems, Orlando, 3 1994, pp. 2113–2118.

[10] Y. Yuan, M.J. Shaw, Induction off fuzzy decision trees, Fuzzy Sets Syst. 69 (1995) 125–139.

[11] C.Z. Janikow, Fuzzy decision trees: issues and methods, IEEE Trans. Syst. Man Cybern. 28 (1998) 1–14.

[12] C. Marsala, B. Bouchon-Meunier, Choice of a method for the construction of fuzzy decision trees (Published in conference proceedings style.), Fuzzy Systems (FUZZ'03), The 12th IEEE International Conference, 1 2003, pp. 584–589 (May, 23–28).

[13] W. Pedrycz, Z.A. Sasnowski, C-Fuzzy decision trees, IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. 35 (4) (2005) 498–511.

[14] D. Marc, T.H. Dang, C. Marsala, B. Bouchon-Meunier, Fuzzy decision tree for the user modeling from human–computer interactions, Presented at 5th International Conference on Human System Learning (ICHSL'05) 2005, pp. 287–302.

[15] X. Liu, W. Pedrycz, The development of fuzzy decision trees in the framework of axiomatic fuzzy set logic, Appl. Soft Comput. 7 (2007) 325–342.

[16] B. Bouchon-Meunier, M. Detyniecki, M.J. Lesot, C. Marsala, N. Rifqi, Real-world fuzzy logic applications in data mining and information retrieval, Fuzzy Logic Stud. Fuzziness Soft Comput. 215 (2007) 219–247.

[17] K. Crockett, Z. Bandar, D. Mclean, On the optimization of T-norm parameters within Fuzzy Decision Trees, Fuzzy Systems Conference, FUZZ-IEEE 2007, IEEE International, 23–26 July 2007 2007, pp. 1–6.

[18] K. Makkithaya, N.V.S. Reddy, U.D. Acharya, Improved C-Fuzzy Decision Tree for intrusion detection, World Acad. Sci. Eng. Technol. 18 (2008) 273–277.

[19] C. Marsala, Gradual fuzzy decision trees to help medical diagnosis, Presented at IEEE World Congress on Computational Intelligence, June. 10–15, 2012 (Brisbane, Australia).

[20] J.A. Sanz, H. Bustince, A. Fernández, F. Herrera, On the cooperation of interval-valued fuzzy sets and genetic tuning to improve the performance of Fuzzy Decision Trees, IEEE International Conference on Fuzzy Systems, June 27–30 2011, pp. 1247–1254 (Taipei, Taiwan).

[21] J.A. Sanz, H. Bustince, A. Fernández, F. Herrera, IIVFDT: ignorance functions based interval-valued fuzzy decision tree with genetic tuning, Int. J. Uncertainty Fuzziness Knowledge Based Syst. 20 (2) (2012) 1–30.

[22] M. Zeinalkhani, M. Eftekhari, Comparing different stopping criteria for fuzzy decision tree induction through IDFID3, Iran. J. Fuzzy Syst. 11 (1) (2014) 27–48.

[23] L. Mannina, M. Patumi, N. Proietti, D. Bassi, A.L. Segre, Geographical characterization of Italian extra virgin olive oils using high-field HNMR spectroscopy, J. Agric. Food Chem. 49 (6) (2001) 2687–2696.

[24] F. Marini, F. Balestrieri, R. Bucci, A.D. Magrý, A.L. Magrý, D. Marini, Supervised pattern recognition to authenticate Italian extra virgin olive oil varieties, Chemom. Intell. Lab. Syst. 73 (2004) 85–93.

[25] A. Cichelli, G.P. Pertesana, High performance liquid chromotographic analysis of chlorophylls, pheophytins and catotenoids in virgin olive oils: chemometric approach to variety classification, J. Chromotogr. A 1046 (2004) 141–146.

[26] S. Rezzi, D.E. Axelson, K. Hèberger, F. Reniero, C. Marini, C. Guillou, Classification of olive oils using high throughput flow "H″ NMR fingerprinting with principal component analysis, linear discriminant analysis and probabilistic neural networks, 2005 13–24.

[27] P.V. Petrakis, A. Agiomyrgianaki, S. Christophoridou, A. Spyros, P. Dais, Geographical characterization of Greek virgin olive oils (Cv. Koroneiki) using "H" and "P NMR" fingerprinting with canonical discriminant analysis and classification binary trees, J. Agric. Food Chem. 56 (2008) 3200–3207.

[28] G. Gurdeniz, B. Ozen, F. Tokatlı, Comparison of fatty acid profiles and mid-infrared spectral data for classification of olive oils, Eur. J. Lipid Sci. Technol. 112 (2010) 218–226.

[29] P. Zunin, R. Boggia, P. Salvadeo, P. Evangelisti, Geographical traceability of West Liguria extravirgin olive oils by the analysis of volatile terpenoid hydrocarbons, J. Chromatogr. A 1089 (2005) 243–249.

[30] O. Galtier, N. Dupuy, Y. Le Dr'eau, D. Ollivier, C. Pinatel, J. Kister, J. Artaud, Geographic origins and compositions of virgin olive oils determinated by chemometric analysis of NIR spectra, Anal. Chim. Acta 595 (2007) 136–144.
[31] A.S. Gumuşkesen, F. Yemiscioglu, Project Name: Türkiye'deki Zeytin Çeşitlerinin ve Zeytinyağlarının Bölgesel Olarak Karakterizasyonu, 2007 (Project Number: 2005/BİL/020. [Online. retrieved February 24. 2016. from http://food.ege.edu.tr/d-83/akademikyapi.html).
[32] A. Bakhouche, J. Lozano-Sánchez, R. Beltrán-Debón, J. Joven, A. Segura-Carretero, A. Fernández-Gutiérrez, Phenolic characterization and geographical classification of commercial Arbequina extra-virgin olive oils produced in southern Catalonia, Food Res. Int. 50 (2013) 401–408.
[33] L.G. Dias, N. Rodrigues, A.C.A. Veloso, J.A. Pereira, A.M. Peres, Monovarietal extra-virgin olive oil classification: a fusion of human sensory attributes and an electronic tongue, Eur. Food Res. Technol. 242 (2016) 259–270.
[34] L.A. Zadeh, Fuzzy sets, Inf. Control. 8 (1965) 338–353.
[35] M. Elkano, M. Galar, J.A. Sanz, A. Fernández, E. Barrenechea, F. Herrera, Enhancing multiclass classification in FARC-HD Fuzzy classifier: on the synergy between n-dimensional overlap functions and decomposition strategies, IEEE Trans. Fuzzy Syst. 23 (5) (2015) 1562–1580.
[36] I.T. Jolliffe Principal, Component Analysis, Springer-Verlag, 1986 487.
[37] J.R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kaufmann, San Mateo. California, 1993.
[38] J.R. Quinlan, Induction of decision trees, Mach. Learn. 1 (1986) 81–106.
[39] J.C. Dunn, A fuzzy relative of the ISODATA Process and its use in detecting compact well-separated clusters, J. Cybern. 3 (1973) 32–57.
[40] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum, New York, 1981 256.
[41] J.C. Bezdek, Cluster validity with fuzzy numbers, J. Cybern. (1974) 58–73.
[42] J.C. Bezdek, Numerical taxonomy with fuzzy sets, J. Math. Biol. 1 (1974) 57–71.
[43] J. Dunn, Well separated clusters and optimal fuzzy partitions, J. Cybern. 4 (1974) 95–104.
[44] T. Nakashima, G. Schaefer, Y. Yokota, A weighted fuzzy classifier and its application to image processing tasks, Fuzzy Sets Syst. 158 (3) (2007) 284–294.
[45] J. Sanz, M. Galar, A. Jurio, A. Brugos, M. Pagola, H. Bustince, Medical diaognosis of cardiovascular diseases using an interval-valued fuzzy rule based classification system, Appl. Soft. Comput. J. 20 (2014) 103–111.
[46] K. Menger, Statistical metrics, Proc. Natl. Acad. Sci. U. S. A. 28 (1942) 535–537.
[47] B. Schweizer, A. Sklar, Probabilistic Metric Spaces, (North-Holland, Amsterdam, 1973.
[48] U. Höhle, Probabilistic uniformization of fuzzy topologies, Fuzzy Sets Syst. 1 (1978) 311–332.
[49] C. Alsina, E. Trillas, L. Valverde, On some logical connectives for fuzzy set theory, J. Math. Anal. Appl. 93 (1983) 15–26.
[50] D. Dubois, H. Prade, New results about properties and smeantics of fuzzy set-theroetic operators, in: P.P. Wang, S.K. Chang (Eds.), Fuzzy Sets, Plenum Press, New York 1986, pp. 59–75.
[51] M.M. Gupta, J. Qi, Theory of T-norms and fuzzy inference methods, Fuzzy Sets Syst. (40) (1991) 431–450.
[52] S. Weber, A general concept of fuzzy connectives. Negations and implications based on T-norms and T-conorms, Fuzzy Sets Syst. 11 (1983) 115–134.
[53] W. Bandler, L. Kohout, Fuzzy power sets and fuzzy implication operators, Fuzzy Sets Syst. 4 (1980) 13–30.
[54] R. Giles, Lukasiewicz logic and fuzzy set theory, Int. J. Man–Mach. Stud. 8 (1976) 313–327.
[55] M. Oussallah, On the use of Hamacher's T-norms family for information aggregation, Inf. Sci. 153 (2003) 107–154.
[56] R.R. Yager, On a general class of fuzzy connectives, Fuzzy Sets Syst. 4 (1980) 235–242.
[57] J. Dombi, A general class of fuzzy operators, the De Morgan class of fuzzy operators and fuzziness induced by fuzzy operators, Fuzzy Sets Syst. 8 (1982) 149–163.