Multicriteria fuzzy classification procedure PROCFTN: methodology and medical application

Belacel, Nabil; Boulassel, M.R.

National Research Council Canada   Conseil national de recherches Canada

Canada

National Research
Council Canada

Conseil national
de recherches Canada

Institute for
Information Technology

Institut de technologie
de l'information

NRC·CNRC

*Multicriteria Fuzzy Classification Procedure PROCFTN: Methodology and Medical Application ***

Belacel, N., and Boulassel, M.R.

Canadä

Multicriteria fuzzy classification procedure *PROCFTN*: Methodology and medical application

Nabil Belacel

National Research Council of Canada
Institute for Information Technology – e-Business
e-Health group
127 Carleton Street
Saint John, NB, E2L 2Z6
CANADA
E-mail: Nabil.Belacel@nrc-cnrc.gc..ca

Mohamed Rachid Boulassel

Montreal Chest Institute, Immunodeficiency Service,
Room 817, McGill
University Health Centre
CANADA
E-mail: rachid.boulasse@muhc.mcgill.ca

Send correspondence to:
Nabil Belacel
National Research Council of Canada
Institute for Information Technology – e-Business
e-Health group
127 Carleton Street
Saint John, NB E3B 6C2
CANADA
E-mail: Nabil.Belacel@nrc-cnrc.gc.ca
Tel: 1-506-636 4774   Fax: 1-506-636 4491

Abstract.

In this paper we introduce a new classification procedure for assigning objects to predefined classes, named *PROCFTN*. This procedure is based on a fuzzy scoring function for choosing a subset of prototypes, which represent the closest resemblance with an object to be assigned. It then applies the majority-voting rule to assign an object to a class. We also present a medical application of this procedure as an aid to assist the diagnosis of central nervous system tumours. The results are compared with those obtained by other classification methods, reported on the same data set, including decision tree, production rules, neural network, k-nearest neighbour, multilayer perceptron and logistic regression. Our results are very encouraging and show that the multicriteria decision analysis approach can be successfully used to help medical diagnosis.

Keywords: Multicriteria decision aid, Classification, Fuzzy sets, Fuzzy binary relations, Scoring function, Astrocytic tumour, Medical diagnosis.

1. Introduction

Classification methods are characterized by a learning phase, which consists in elaborating the classification rules from the available knowledge. This phase uses an inductive or deductive approach. With the inductive approach, the classification rules are acquired from examples and each example belongs to a well-known class. The aim of this algorithm is to produce classification rules for assigning new examples to classes. There are numerous methods which use the inductive approach, including the k nearest neighbor rule, Bayesian techniques, discriminant analysis, neural network and decision trees [25,27,32,37]. With the deductive approach, the classification rules are given a priori by the interaction with the decision-maker, or the expert. From these rules we determine the assignment of objects to classes. The expert system [3] and the rough set [24] methods belong to these kinds of approaches. In general, the methods mentioned above can only use inductive or deductive approach, but not both at the same time. In practice, some problems are such that their solution needs a method that can employ both kinds of approaches [6-7]. This is one of the reasons why the fuzzy assignment procedure *PROCFTN* was developed. This procedure uses the multicriteria decision aid (MCDA) approach. The latter is based on the preference relational system described by Roy in 1996 [34] and Vincke in 1992 [35]. In the MCDA approach, the decision problems require the comparison between the alternatives through the scores of different criteria or attributes using relative or absolute evaluations. The relative evaluation compares the alternatives in order to select the best one or to rank them in decreasing order of preference, while absolute evaluation compares the alternatives with the different prototypes of classes in order to assign the alternatives to specific classes. Furthermore, the MCDA approach avoids resorting to the use of distances and allows the use of qualitative and/or quantitative attributes without any transformations on the data. Besides, it helps to overcome some difficulties encountered when data are expressed in different units. These advantages offered by the MCDA approach constitute the second reason for developing a procedure, which resorts to this approach. On the other hand only a few methods using MCDA approach have been applied in medical diagnosis [4-8]. This fact constitutes another reason for developing our method. The MCDA literature is abound with numerous approaches to the choice problem but few works have been used to solve classification problems [22,23]. The *PROCFTN* procedure solves a choice problem in order to determine a subset of k, $k \geq 1$, of closest prototypes in terms of their resemblance, or

similarity, with an object to be assigned. Then, it applies the majority-voting rule to assign an object to a class.

The aims of the present study are (1) to elaborate a new fuzzy multicriteria classification method, which is based on a scoring function from a fuzzy preference relation; and (2) to evaluate the ability of this procedure to classify central nervous system tumours. The latter has been chosen in order to compare our classification results with those previously obtained by other classification methods.

The present paper is organized as follows: section 2 presents a general study of multicriteria classification problems. Section 3 describes the different stages of the *PROCFTN* method. Section 4 is devoted to the medical application of the proposed method.

2. Multicriteria classification problems

The multicriteria classification problem is known in multicriteria decision aid as a *sorting problematic* [4,30,36]. It consists in formulating the decision problem in terms of assigning each alternative or object to one or several classes. This assignment is achieved through the examination of the intrinsic value of the object by referring to pre-established norms. The classes are defined by a set of reference objects or prototypes. The classes may be ordered or not, depending on the formulation of the decisional problems:

- *Ordered classes*: They are characterized by a sequence of boundary reference objects representing upper and lower frontier classes. This type of problematic is known as an *ordinal sorting problematic* [30,36]. The assignment rule is formulated as follows: *each object, which is judged preferred to a reference object reflecting the lower limit of the class must be assigned to that class* [30]. The evaluation of students is an example, which can be treated by using this problematic [36].
- *Non-ordered classes*: They are characterized by one or several central reference objects or *prototypes*. This type of problematic is known as a *nominal sorting problematic* [4-8]. The assignment rule is formulated as follows: *each object, which is judged similar or indifferent to at least one prototype of a class, must be assigned to that class*. As an example, we can cite medical diagnosis where the objects i.e., patients, presented by different symptoms and the prototypes are represented by typical symptoms [4-8].

In general, to solve multicriteria classification problems, we proceed in two stages:

Stage 1: *Modelling of classes*: At this stage we fit the parameter values (weights, discrimination thresholds, etc.) by using the set of cases designated as the training set. These cases are partitioned into mutually exclusive classes (e.g. tumour grades) and

described by their values for a set of attributes (e.g., parameters generated by computer-assisted microscope analysis of cell images) [6-8]. After determining the initial parameter values using the available knowledge and the training set, we proceed to validate these parameters by one of two possible techniques:

- *Direct technique*: It consists of adjusting the parameters through the training set and under the decision-maker guidance.
- *Indirect technique*: It consists of fitting the parameters without assistance of the decision-maker. This technique requires less cognitive effort than the direct technique; it uses an automatic method to determine the optimal parameters, which minimizes classification errors.

Stage 2: *Assignment decision*: After designing the prototypes we proceed to assign the new objects to specific classes.

3. The developed method

In this section, we propose the fuzzy choice procedure called *PROCFTN*, "in French: PROcédure de Choix Flou dans le cadre de la problématique du Tri Nominal", to solve multicriteria classification problems. This procedure solves a choice problem in order to determine a subset of k, $k \geq 1$, nearest neighbors in terms of their resemblance, or similarity, with an object to be assigned. On the basis of the fuzzy scoring function [20-21] and the fuzzy indifference relations determined by the *PROAFTN* method [4-8], the *PROCFTN* procedure determines the k nearest neighbor prototypes. So, the *PROCFTN* method follows the k nearest neighbor (k-NN) procedure described by Cover and Hart [15]. The k-NN algorithm uses a training set to classify an unknown case by determining a number, k, of closest training cases. A majority-voting rule is applied to assign an unknown case to a class: an unclassified case is assigned to the class represented by the majority of its k nearest neighbors in the training set [16]. The parameter k in a k-NN algorithm is always given a priori. This is inconvenient insofar as the majority among the k nearest neighbors of an unknown case depends on the parameter k. So, if the number k changes, the assignment decision may also be changed. In order to remedy this disadvantage, our procedure solves a choice problem to determine the k closest prototypes.

The data and notations used by *PROCFTN* are:

- *A*: set of objects to be assigned to different classes (throughout this paper, *A* will be finite and non-empty)
- $F = \{g_1, g_2,..., g_n\}$: set of *n* attributes.

- Ω: set of p classes: $\Omega = \{C^1, \dots, C^p\}$, $p \geq 2$.

- B^h: prototype set of the hth class, where: $B^h = \{b_i^h;\ i=1,\dots,L_h\}$, $h=1,\dots,p$, with b_i^h designating the ith prototype of the hth class and L_h designating the cardinality of the set B^h.

- B: set of all prototypes, such that: $B = \bigcup_{h=1}^{p} B^h$ (the set B will be finite, non-empty and contains m prototypes).

- $\hat{A}$: set of objects A and B such as: $\hat{A} = A \cup B$.

The objects are described using a vector of n attributes $g = (g_1, g_2,\dots, g_n)$. The performance of object a on attribute g_j will be denoted by $g_j(a)$. Therefore, the scores of objects are evaluated on the attribute set F, such as:

$$\forall a \in A, \text{ we have:} \qquad g(a) = (g_1(a),\ g_2(a),\ \dots,\ g_n(a)),$$

In general, the prototype scores are given by intervals, so for each attribute g_j, we associate to each prototype b_i^h the interval $[S_j^1(b_i^h), S_j^2(b_i^h)]$, with $S_j^2(b_i^h) \geq S_j^1(b_i^h)$.

- w_j, $j=1,\dots,n$, are positive coefficients adding to one and reflecting the intrinsic relative importance attached by a decision maker to an attribute g_j (with $\sum_{j=1}^{n} w_j = 1$).

- $C_j(a, b_i^h)$, $j=1,\dots,n$, is the degree with which the attribute g_j is in favor of the indifference or resemblance relation between the object a and the prototype b_i^h. Figure 1 illustrates how it is calculated. In this figure, two positive discrimination thresholds $d_j^+(b_i^h)$ and $d_j^-(b_i^h)$, are used to take into account the imprecision of the data.

- $D_j(a, b_i^h)$, $j = 1,\dots,n$, is the degree with which the attribute g_j is against the indifference relation between the object a and the prototype b_i^h. Figure 2 illustrates how it is calculated. In this figure, veto thresholds $v_j^-(b_i^h)$ and $v_j^+(b_i^h)$, $j= 1,\dots,n$, are used to define the values for which a is considered as very different from b_i^h for the attribute g_j.

- $D_I(a,b_i^h)$ is the fuzzy comprehensive discordance relation for the assertion: "the object a is indifferent (or roughly equivalent) to the prototype b_i^h". To determine a comprehensive discordance index D_I, we can use a compromise aggregation operator that has a value 0 if at least one D_j, $j=1,\dots, n$, is equal to 1 [4-5,29-31]. Then,

$$D_I(a,b_i^h) = 1 - \prod_{j=1}^{n}(1 - D_j(a,b_i^h))^{w_j}$$

For a more detailed analysis of concordance and discordance indices, see references [4-5,10,28,30,33-35].

We denote "*the object a is assigned to class C^h*" by: $a \in C^h$.

Given an object a, described by n attributes, to assign the object a to the corresponding class, the *PROCFTN* proceeds in six stages:

1. Initialization
2. Performance matrix
3. Fuzzy preference relations between the prototypes
4. Scoring function from a fuzzy preference relation
5. Choice of prototypes
6. Assignment decision

3.1. *Initialization*

For each class C^h, $h=1,...,k$, we determine a set of L_h prototypes $B^h = \left\{ b_1^h, b_2^h, ..., b_{L_h}^h \right\}$. The prototypes are considered as good representatives of their class and are described by the score on each of the n attributes. More precisely, to each prototype b_i^h and each attribute g_j, $j=1,...,n$, an interval $[S_j^1(b_i^h), S_j^2(b_i^h)]$ is defined, with $S_j^2(b_i^h) \geq S_j^1(b_i^h)$, $j=1,...,n$, $h = 1,...,k$ and $i= 1,...,L_h$. For determining these intervals we follow the general scheme of the discretization technique described by Ching *et al.* [14] and Fayyad *et al.* [19] using the training set. The parameters such as thresholds and weights are determined by the interaction with the decision-maker. The strategy for adjusting the parameters follows the general scheme described in figure 3. An initialization step is used to propose the initial parameters (discrimination thresholds), which is updated during the optimization process. This strategy enables us to minimize the classification errors. The fact that the *PROCFTN* procedure uses a case set for which the assignment classes are known a priori in order to determine and adjust the parameters (interval boundaries, thresholds…), means that it can be considered as a supervised learning algorithm [37].

3.2. *Performance matrix*

The performance matrix is determined to evaluate the prototypes of classes on a set of attributes. The rows of this matrix represent the prototypes of the classes and the columns represent the attributes. The intersection between the row i and the column j corresponds to the resemblance degree $R_{jh}^i(a, b_i^h)$ between the prototype b_i^h and the object a to be assigned according to the attribute g_j (see Table 1). In order to calculate the value $R_{jh}^i(a, b_i^h)$ we determine the partial indifference $C_j(a, b_i^h)$, the partial discordance $D_j(a, b_i^h)$ and the overall discordance $D_I(a, b_i^h)$ indices using the *PROAFTN* procedure [4-5]. Once the

indices are given we calculate the fuzzy resemblance relations, which have the following properties:

1. If one or more attributes in *F* opposes a veto against the assertion "*a is indifferent to the prototype b_i^h*", then all partial resemblance relations $R_{jh}^i(a, b_i^h)$, *j=1,...,n*, between the object *a* and the prototype b_i^h will be null. So, the object *a* is considered as very different from the prototype b_i^h. Formally:

    If $(\exists j \in \{1,...,n\}, D_j(a, b_i^h) = 1)$, then $(\forall j = 1, ..., n, R_{jh}^i(a, b_i^h) = 0)$,

    $$h = 1, ..., p; \; i = 1, ..., L_h$$

2. The resemblance index takes into account the dissimilarity (i.e., comprehensive discordance index) between an object to assign and a given prototype. Thus, if the discordance index is very high, then the resemblance index will be very weak. Formally:

    If $(D_I(a, b_i^h) > C_j(a, b_i^h))$, then $R_{jh}^i(a, b_i^h) = C_j(a, b_i^h) \times (1 - D_I(a, b_i^h))$

    else $\quad R_{jh}^i(a, b_i^h) = C_j(a, b_i^h)$,

From properties 1 and 2, the resemblance degree between the object *a* and the prototype b_i^h is given as follows:

$$R_{jh}^i(a, b_i^h) = \begin{cases} C_j(a, b_i^h) & \text{if} \quad D_I(a, b_i^h) \leq C_j(a, b_i^h), \\ \\ C_j(a, b_i^h) \times (1 - D_I(a, b_i^h)) & otherwise \end{cases} \quad (1)$$

3.3 Fuzzy preference relation between the prototypes

By using the performance matrix determined in stage 1, the fuzzy preference relations between the various prototypes can be defined as follows:

Definition 1. *The prototype b_i^h is preferred to the prototype b_t^l (b_i^h P b_t^l) if and only if the resemblance degree between the object a and the prototype b_i^h is stronger than the resemblance between the object a and the prototype b_t^l on the whole set of attributes.*

The fuzzy preference relation *P* is based on the partial credibility indices P_j, *j=1,...,n*. Each index represents the credibility degree of the following situation: "*the degree of resemblance between object a and prototype b_i^h is stronger than the resemblance between object a and the other prototypes according to attribute g_j*".

The partial preference index P_j *between* the prototypes b_i^h and b_t^l is given as follows:

$$P_j(b_i^h, b_t^l) = max\{R_{jh}^i(a, b_i^h) - R_{jt}^l(a, b_t^l), 0\} \quad (2)$$

From these partial preference indices and by taking into account the relative importance of attributes, we determine for each pair of prototypes (b_i^h, b_t^l) the overall fuzzy preference relation $P(b_i^h, b_t^l)$.

$$P(b_i^h, b_t^l) = \sum_{j=1}^{n} \left\{ w_j \times P_j (b_i^h, b_t^l) \right\}, \qquad h=1,...,p \; ; \; l=1,...,p; \qquad (3)$$

$$i=1,...,L_h \, ; \; t=1,...,L_l,$$

3.4 Scoring functions from a fuzzy preference relation

The *PROCFTN* procedure is based on a scoring function from a fuzzy preference relation P and selects the best prototypes in terms of their resemblance with the object to be assigned. Before explaining in detail how *PROCFTN* proceeds to determine these prototypes, we point out a very important definition adapted to our context as follows [20]:

Definition 2. *Let B be a set of m prototypes and P be a fuzzy preference relation on B. A function f is said to be a scoring function on B for the relation P if it is a real-valued function defined on* $[0,1]^{2 \times m}$, *non-decreasing in terms of its m first arguments, non-increasing in terms of its m last arguments and such that:*

$$\forall i \in \{1,...,L_h\} \text{ and } \forall h \in \{1,...,p\}$$

$$s(b_i^h, B, P) = f(P(b_i^h, b_1^1),..., P(b_i^h, b_{Lp}^p), P(b_1^1, b_i^h),..., P(b_{Lp}^p, b_i^h))$$

where $s(b_i^h, B, P)$ *is the score of prototype* b_i^h *in B according to relation P.*

A few score functions reported in the literature are given below:

- *The leaving flow* (see [12]):

$$s_1(b_i^h, B, P) = \sum_{x \in B} P(b_i^h, x) \qquad (4)$$

- *The (complemented) entering flow* (see [12]):

$$s_2(b_i^h, B, P) = \sum_{x \in B} (1 - P(x, b_i^h)) \qquad (5)$$

- *The net flow* (see [12]):

$$s_3(b_i^h, B, P) = \frac{1}{m} \sum_{x \in B} \left(P(b_i^h, x) - P(x, b_i^h) \right) \qquad (6)$$

- *The min leaving flow* [10]:

$$s_4(b_i^h, B, P) = \min_{x \in B} P(b_i^h, x) \qquad (7)$$

- *The (complemented) max entering flow* (see [10,28]):

$$s_5(b_i^h, B, P) = 1 - \max_{x \in B} P(x, b_i^h) \qquad (8)$$

- *The Orlovski score* (see [1,28]):

$$s_6(b_i^h, B, P) = \min_{x \in B} \min\left(1 - P(x, b_i^h) + P(b_i^h, x); 1\right)$$ (9)

- *The Min difference score* (see [2]):

$$s_7(b_i^h, B, P) = \min_{x \in B}(P(b_i^h, x) - P(x, b_i^h))$$ (10)

Other scoring functions based on t-norms and t-conorms have been proposed by Roubens in [33].

3.5 Prototype choice

We introduce the fuzzy choice function C_s, which is used to select the prototypes. The latter presents the closest resemblance to the object a to be assigned. The fuzzy choice function is obtained from a scoring function s as follows:

$$\mu_{C_s(B)}(b_i^h) = s(b_i^h, B, P).$$

Any λ-cut of the fuzzy set $C_s(B)$ defines an ordinary choice function C_s and such that:

$$C_s(B) = \{ b_i^h \in B \text{ s.t. } \mu_{C_s(B)}(b_i^h) \geq \lambda \}, \text{ with } \lambda \geq \frac{1}{2} \text{ the cut value.}$$

Let us note that function C_s does not permit to define a choice function insofar as the $C_s(B)$ set can be empty. In order to remedy this difficulty, we use a choice function proposed by Orlovski (see [1,28-29]):

$$C_s(B) = \{b_i^h \in B \quad s.t. \quad \mu_{C_s(B)}(b_i^h) = \max_{x \in B} \mu_{C_s(B)}(x) \}$$ (11)

Considering the characteristics of the relation P, which is a cardinal relation, we can use the scoring functions given by Eq. (4-6, 9-10). In our application (see Section 4) we choose the Orlovski scoring function given by Eq. (9) because it is well adapted to our application. Actually, we have tested all scoring functions, i.e. the leaving flow, the (complemented) entering flow, the net flow, the Orlovski score, the min difference score, on training test sets. The results have shown that the Orlovski scoring function has outperformed other functions.

PROCFTN determines the degree with which each prototype dominates all the other prototypes in B according to their resemblance to the object a to be assigned. The choice set defined by Eq. (11) contains a subset of k ($k \geq 1$) prototypes bearing more resemblance to the object to be assigned.

3.6 Assignment decision

The classes, $C^h, h=1,...,p$, are represented by a B^h set of L_h prototypes, which are considered as good representatives of their class and are described by the score upon each of n

attributes. On the basis of the performance matrix (see Table 1 and Eq.(1)), to determine the plausible class for the object *a,* we proceed as follows:

1. *Computing the fuzzy preference relations between the prototypes:* The fuzzy preference relation represents the degree to which the resemblance between the object to be assigned and the prototype is stronger than the resemblance between the object and another prototype. This relation is calculated from Eq. (3).

2. *Fuzzy scoring function from a fuzzy preference relation:* The fuzzy scoring functions evaluate the extent to which each prototype dominates all the other elements in *B*. Among the scoring functions, which may be used by *PROCFTN,* we can mention the functions given in Eq. (4,5,6,9,10).

3. *Prototype choice:* The choice set of prototypes *C(B)* associating the scoring function from a fuzzy preference relation *P* is computed from Eq. (11). The subset of prototypes contains *k* prototypes, which represent the closest resemblance to an object to be assigned.

4. *Assigning the object 'a' to predefined classes:* Once the choice set *C(B)* is determined, a majority-voting rule is applied to assess an object in a class: an object is assigned to the class containing the majority of *k* prototypes of *C(B)* set. This assignment rule is inspired from the *k*-nearest neighbor method [15,16].

4. Medical application

4.1 Astrocytic tumours diagnosis

The method presented in this paper was tested with an experimental set of 250 cases of astrocytic tumours (AT). They are the most common among the primary intra-cranial neoplasms of the mature nervous system [9]. Several grading schemes have been proposed for AT classification and the most widely used is the classification proposed by the World Health Organization (WHO). According to WHO three distinct groups of AT are defined namely astrocytomas (AST), anaplastic astrocytomas (ANA) and glioblastomas (GBM) [24]. These three histopathological groups are divided into two levels of tumour aggressiveness. AST is considered to be low grade (benign) while ANA and GBM are considered as high grade (malignant) [13,24]. The determination of tumour aggressiveness (grading of malignancy) is based on the description by a pathologist of the morphological characteristics appearing on hematoxylin-eosin stained tissue sections. The recognition of more homogeneous grades of histologically similar cases is important as it allows an improved understanding of the tumour's progress and determines the future treatment.

The parameter values used to classify AT were kindly supplied by Dr. Decaestecker (Laboratory of Histology, Faculty of Medicine, Free University of Brussels, Belgium). They were generated by computer-assisted microscope analysis of cell image. The clinical characteristics related to this data set as well as the determination of the parameters can be found in Decaestecker [17] and in Decaestecker *et al.* [18]. The experimental set of 250 cases of AT was divided into three groups: 39 cases of AST, 47 cases of ANA and 164 cases of GBM. Each case was labeled according to its histopathological group as established previously by the clinical diagnosis. A total of 26 quantitative parameters (see Table 2) including ten for nuclear deoxyribonucleic acid content and sixteen for morphonuclear and chromatin texture parameters were submitted to the *PROCFTN* method, which determines the plausible assignment classes. The performance of this method was determined using the ten-fold cross-validation technique described by Weiss and Kulikowski [37]. Each group of AT was tested separately.

4.2 Results and discussion

We compared the classification results obtained by PROCFTN with those previously given by an experienced specialist in order to determine whether a case was correctly classified or not. We focused our analysis on the most important indicator of performance, i.e., accuracy of classification. The main reason for doing this is the structure of the available data. In the set of AST cases, 66 % were correctly classified and 44 % were incorrectly classified. The percentages of correct classification in the ANA and GBM groups were 68 % and 64 % respectively. It is important to point out that no case in the ANA group was classified as AST. So, it is possible to discriminate a high grade, i.e., ANA, from a low grade, i.e., AST, simply on the basis of the features measured by image analyzing systems. However, the fact that some cases in the AST group were classified as ANA or GBM is less convincing. The average percentage of the whole testing sets was 66 % for correct classification of cases. This percentage is unsatisfactory despite the fact that this result is similar to those reported by the other classifiers [17-18]. Figure 4 compares the different classifiers with the results obtained for the same data set studied. As we can see in figure 4, a performance similar to production rules, logistic regression and multilayer perceptron methods is obtained by our procedure [17-18]. Only the production rule, logistic regression and multilayer perceptron, realized a score of about 65 % of positively classified cases. These results further show that a classification method based on scoring function yields comparable results in terms of separation between different astrocytic tumour groups. In further research, we would like to

check whether the combination of the clinical decision rules with the parameters measured by the image analyzing system might improve the accuracy of the classification.

Compared to other classification approaches, our method offers several advantages. The first advantage is that our method uses the concordance and non-discordance principles to determine the preference relations. Essentially, our procedure uses veto thresholds and discordance indices allowing compensation to be avoided when attributes are strongly conflicting. Because of this important property, methods based on concordance and non-discordance principles are often called non-totally compensatory. This identifies their main advantage compared to more traditional methods based on a single overall criterion. These traditional methods usually assume a form of compensation between the contributions of particular attributes in the process of classification decision.

The second advantage is that prior knowledge (i.e. clinical decision rules) and data (i.e. clinical cases) can be combined without any difficulties by our method, so it should be quite easy to introduce new features in order to improve the accuracy of classification.

The third advantage is that the proposed procedure, *PROCFTN*, provides the possibility to have access to more detailed information concerning the classification decision. So, it gives clear guidance to the decision-maker (pathologist) for deciding whether an object belongs to a class, or not.

5. Conclusion

We have introduced a classification procedure, which is based on a scoring function from a fuzzy preference relation for solving classification problems. To the best of our knowledge, it is the first time that a method based on a scoring function from a fuzzy preference relation has been suggested as a tool for solving fuzzy classification problems. Preliminary results demonstrate the potential performance of this procedure when solving classification problems. They also show that the multicriteria decision aid approach will play an important role in clinical classification problems. Further developments of the procedure include the following research directions: (i) the extension of the *PROCFTN* procedure to more complex situations where the objects are only partly understood and are described by fuzzy subsets of attributes; (ii) the combination of *PROCFTN* and Variable Neighborhood Search metaheuristic for solving very large instances; (iii) the application of enhanced procedure to more real world problems in pattern recognition, medical diagnosis, data mining and e-business.

References

[1] A. Banerjee, Rational choice under fuzzy preferences: The Orlovski choice function, Fuzzy Sets and Systems 54 (1993) 295-300.

[2] C.R. Barrett, P.K. Pattanaik, M. Salles, On choosing rationally when preferences are fuzzy, Fuzzy Sets and Systems 34 (1990) 197-212.

[3] PH. Bartels, Weber JE., Expert systems in histopathology. I. Introduction and overview, Anal. Quant. Cytol. Histology,11 (1989) 1-7.

[4] N. Belacel, Méthodes de classification multicritère : Méthodologie et application à l'aide au diagnostic médical, Ph.D. thesis, SMG, Free University of Brussels, Belgium, 1999.

[5] N. Belacel, Multicriteria assignment method *PROAFTN*: Methodology and medical application, European Journal of Operational Research 125/1 (2000) 175-183.

[6] N. Belacel, M.R. Boulassel, Multicriteria fuzzy assignment method: a useful tool to assist medical diagnosis, Artificial Intelligence in Medicine 21 (2001) 201-207.

[7] N. Belacel, Ph. Vincke, J.M. Scheiff, M.R. Boulassel, Acute leukemia diagnosis aid using multicriteria fuzzy assignment methodology, Computer Methods and Programs in Biomedecine 64(2001) 145-151.

[8] N. Belacel, Ph. Vincke, M.R. Boulassel, Application of *PROAFTN* method to assist astrocytic tumors diagnosis using the parameters generated by computer-assisted microscope analysis of cell image, Innovation and Technology in Biology and Medicine, 20/4 (1999) 239-244.

[9] McL. P. Black, Brain tumours. Second of two parts, New England Journal of Medicine 324 (1991) 1555-1564.

[10] D. Bouyssou, A note on the min in favor ranking method for valued preference relations, in: M. Cerny, D. Gluckaufova and D. Loula (Eds.), Proc. International Workshop on Multicriteria Decision Making-Methods-Algorithms-Applications, Czechoslovak 1991, pp. 16-25.

[11] D. Bouyssou, On some properties of outranking relations based on a concordance-discordance principle, in: A. Goicoechea, L. Duckstein, S. Zionts (Eds.) Springer-Verlag, Multiple Criteria Decision Making. Berlin, 1992, pp. 93-106.

[12] J.P. Brans, Ph. Vincke, A preference ranking organization methods, Management Science 31 (1985) 647-656.

[13] P.C. Burger, F.S. Vogel, S.B. Green, T.A. Strike, Glioblastoma multiforme and anaplastic astrocytoma: Pathologic criteria and prognostic implications, Cancer 56 (1985) 1106-1111.

[14] J.Y. Ching, A.K.C. Wong, K.C.C. Chan, Class-Dependent descretization for inductive learning from continuous and mixed-mode data, IEEE Transaction on Pattern Analysis and Machine Intelligence 7 (1995) 641-651.

[15] T. Cover, P.G. Hart, Nearest neighbour pattern classification, IEEE Transaction on Information Theory, IT-13 (1967) 21-27.

[16] B.V. Dasarathy, Nearest Neighbour (NN) norms: NN Pattern classification technique, IEEE Computer Society Press, Los Alamitos, CA, 1991.

[17] C. Decaestecker, Développements méthodologiques pour la classification de données réelles. Application à l'aide au diagnostic et au pronostic de tumeurs gliales, thesis, Free University of Brussels, Belgium, Histology Laboratory of Medicine Faculty 1997.

[18] C. Decaestecker, I. Salmon, S. Camby, *et al.*, Identification of high versus lower risk clinical subgroups in a group of adult patients with supratentorial anaplastic astrocytomas, Journal of Neuropathology and Experimental Neurology, vol. 54, 3 (1995) 371-384.

[19] U.M. Fayyad, KB Irani, On the handing of continuous-valued attributes in decision tree, Machine Learning, 8 (1992) 87-102.

[20] J. Fodor, M. Roubens, Fuzzy preference modelling and multicriteria decision support, Kluwer Academic Publishers, Dordrecht, 1994.

[21] J. Fodor, S.A. Orlovski, P. Perny, M Roubens, The use of fuzzy preference models in multiple criteria choice, ranking and sorting, in: Fuzzy Sets in Decision Analysis, Operational Research and Statistics, Edited by Roman Slowinski, (Dubois D. and Prade H.) Kluwer Academic Publishers, London, 1998, pp. 69-101.

[22] S. Greco, B. Matarazzo, R. Slowinski, Rough sets theory for multicriteria decision analysis, European Journal of Operational Research 129/1 (2001) 1-47

[23] S. Greco, B. Matarazzo, R. Slwinki, The use of rough sets and fuzzy sets in MCDM, Chap. 14 [in] T. Gal, T. Stewart, T. Hanne (eds.), *Advances in Multiple Criteria Decision Making*, Kluwer Academic Publishers, Boston/Dordrecht/London, 1999, pp. 14.1-14.59.

[24] J. Jelonek, K. Krawiec, R. Slowinski, J. Stefanowski, J. Szymas, Neural networks and rough sets-comparison and combination for classification of histological pictures, in

Proceeding the first National Conference Neural Networks and their Applications, Kule, Czqstochawa, 1994, pp. 268-273.

[25] E. R. Kandel, J. H. Schwartz, T. M. Jesse, Principles of neural science, Elsevier, New York, 1991.

[26] P. Kleihues, P.C. Burger, B.W. Scheithauer, Histological typing of tumours of the central nervous system, World Health Organization International Histological Classification of Tumours. Second edition. Berlin, Springer Verlag, 1993.

[27] G.J. McLachlan, Discriminant analysis and statistical pattern recognition, Wiley et Sons, Inc,1992.

[28] S.A Orlovski, Decision-making with a fuzzy preference relation, Fuzzy Sets and Systems 1 (1978) 155-167.

[29] P. Perny, Modélisation, agrégation et exploitation de préférences floues dans une problématique de rangement, Ph.D. thesis, Paris, LAMSADE, Paris Dauphine University 1992.

[30] P. Perny, Multicriteria filtering methods based on concordance and non-discordance principles, Annals of Operations Research 80 (1998) 137-165.

[31] P. Perny, B. Roy, The use of fuzzy outranking relations in preference modeling, Fuzzy Sets and Systems 49 (1992) 33-53.

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

[33] M. Roubens, Some properties of choice functions based on binary relations, European Journal of Operational Research 40 (1989) 309-321.

[34] B. Roy, Multicriteria methodology for decision aiding, Kluwer Academic, 1996.

[35] Ph. Vincke, Multicriteria decision aid, J. Wiley (Eds.), New York, 1992.

[36] Y. Wei, Aide multicritère à la décision dans le cadre de la problématique du tri: concepts, méthodes et applications, Ph.D. thesis, Paris, LAMSADE, Paris Dauphine University 1992.

[37] SM. Weiss, CA. Kulikowski, Computer systems that learn, classification and prediction methods from statistics, neural nets, machine learning and expert systems, San Mateo, California Morgan Kaufmann Publishers, 1991.

Table 1. Performance matrix of prototypes according to their resemblance with an object a to be assigned.

	g_1	G_2	....	g_j	....	g_n
b_1^1	$R^1{}_{11}(a,b^1{}_1)$	$R^1{}_{21}(a,b^1{}_1)$	....	$R^1{}_{j1}(a,b^1{}_1)$	....	$R^1{}_{n1}(a,b^1{}_1)$
b_2^1	$R^2{}_{11}(a,b^1{}_2)$	$R^2{}_{21}(a,d^1{}_2)$	....	$R^2{}_{j1}(a,d^1{}_2)$	....	$R^2{}_{n1}(a,b^1{}_2)$
....	....	....	....	....	....	....
b_i^h	$R^i{}_{1h}(a,b^h{}_i)$	$R^i{}_{2h}(a,b^h{}_i)$	....	$R^i{}_{jh}(a,b^h{}_i)$	....	$R^i{}_{nh}(a,b^h{}_i)$
....	....	....	....	....	....	....
$b_{L_k}^k$	$R_{1k}^{L_k}(a,b_{L_k}^k)$	$R_{2k}^{L_k}(a,b_{L_k}^k)$	....	$R_{jk}^{L_k}(a,b_{L_k}^k)$	....	$R_{nk}^{L_k}(a,b_{L_k}^k)$

Table 2. Classification attributes used by *PROCFTN* procedure.

Attributes	Attributes Description
DI	DNA Index
%2C	Percentage of Diploid Cell Nuclei
%3C	Percentage of Triploid Cell Nuclei
%4C	Percentage of Tetraploid Cell Nuclei
%H2C	Percentage of Hyperdiploid Cell Nuclei
%H3C	Percentage of Hypertriploid Cell Nuclei
%H4C	Percentage of Hypertetraploid Cell Nuclei
%H5C	Percentage of Pentaploid Cell Nuclei
%ANEUP	Percentage of Aneuploid Cell Nuclei
CH3DI	DNA Index Hypertriploid
NA	Nuclear Area
SDNA	Standard Deviation Nuclear Area
IOD	Integrated Optical Density
MOD	Mean Optical Density
SK	Skewness
VOD	Variance of Optical Density
K	Kurtosis
SRL	Short Run Length
LRL	Long Run Length
GLD	Grey Level Distribution
RLD	Relative Distribution Frequencies
RLP	Relative Distribution Percentage
LM	Local Mean
E	Energy
CV	Coefficient Variance
C	Contrast

Figure 1. Illustrates the graphical representation of the partial indifference index between the object a and the prototype b_i^h. This graph assumes continuity and linear interpolation.
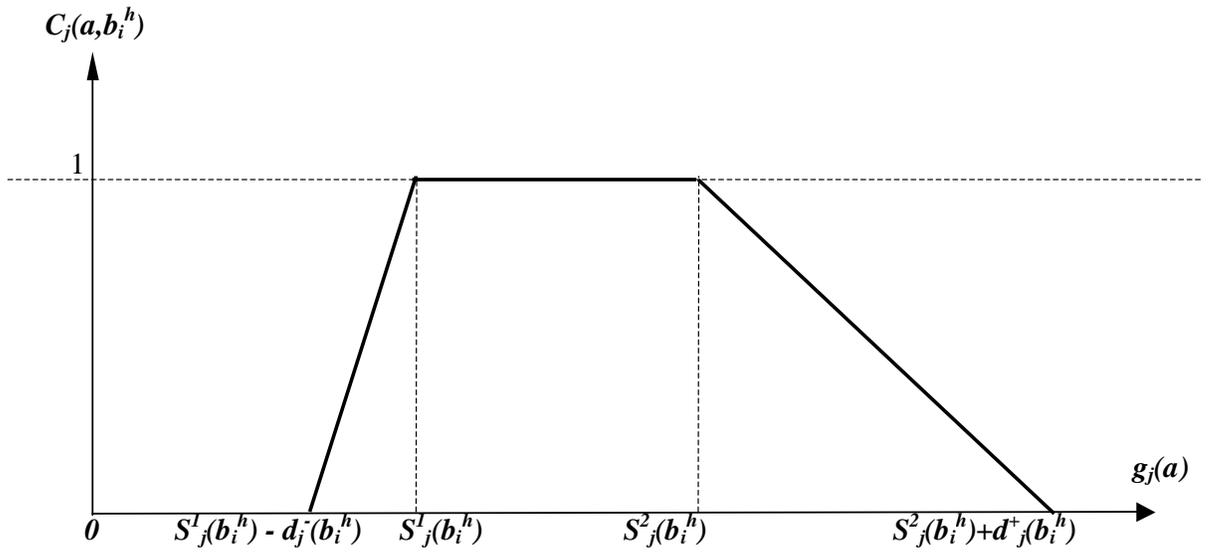
Figure 2. Illustrates the graphical representation of the partial discordance index with regard to the indifference relation between the object a and the prototype b^h_i. This graph assumes continuity and linear interpolation.
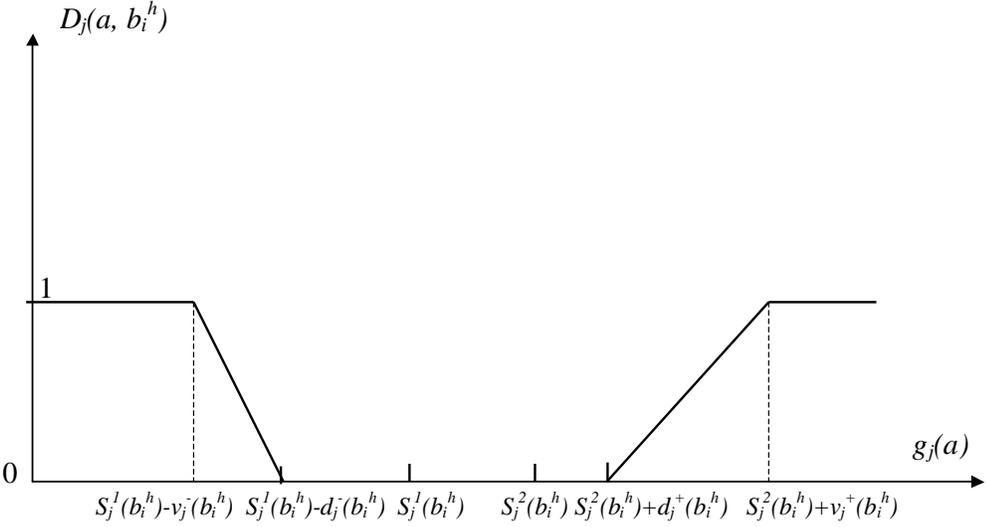

$$D_j(a, b^h_i)$$

$$g_j(a)$$

$S^1_j(b^h_i)-v^-_j(b^h_i)$ $S^1_j(b^h_i)-d^-_j(b^h_i)$ $S^1_j(b^h_i)$ $S^2_j(b^h_i)$ $S^2_j(b^h_i)+d^+_j(b^h_i)$ $S^2_j(b^h_i)+v^+_j(b^h_i)$

Figure 3. General scheme for parameter fitting.

Start

Set thresholds (q_j^+ and q_j^-, $j=1,\ldots,n$)
Set weights (w_j^h, $j=1,\ldots,n$ and $h=1,\ldots,k$)

Assign objects from training set to classes.

Stopping ?

Fit the parameters (weights and thresholds)
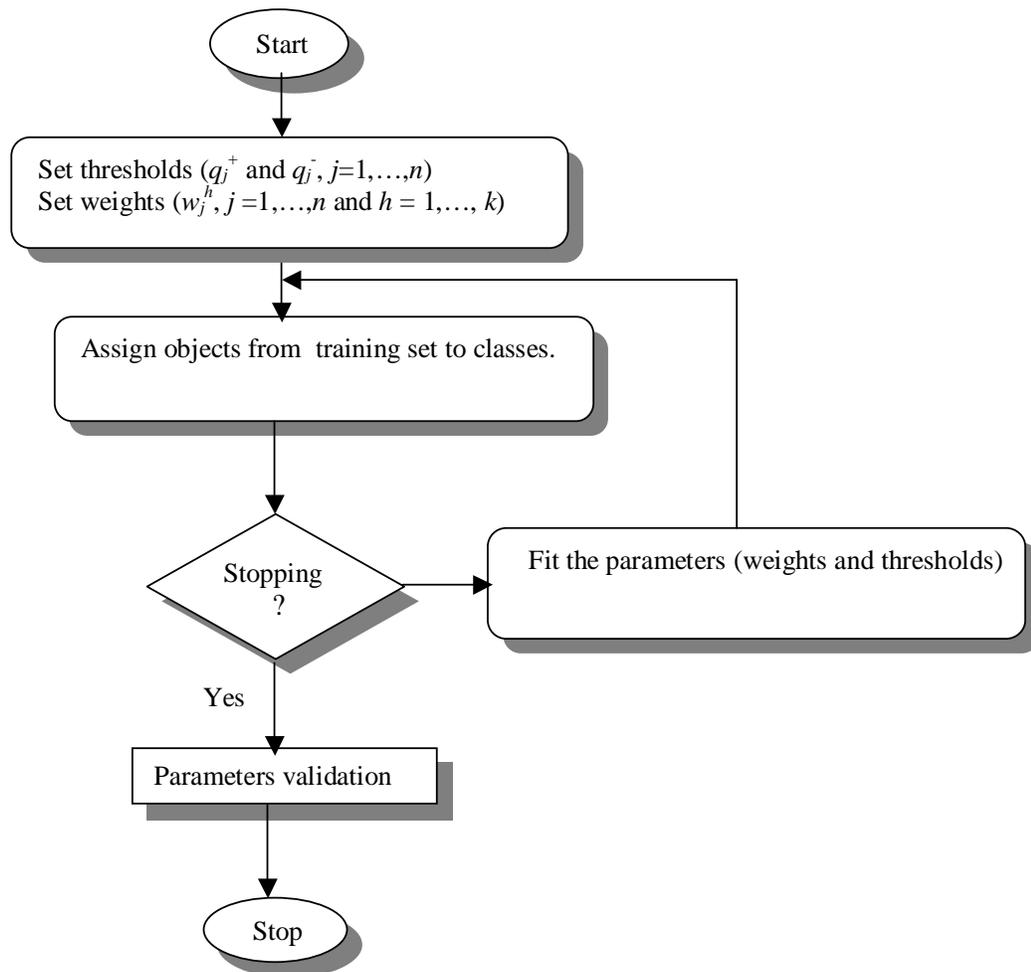
Yes

Parameters validation

Stop

Figure 4. Illustrates the performance of the different classifiers obtained from the three histopathological groups. For each classifier the figure gives the classification accuracy estimated by the 10-fold cross-validation technique.