Mathematical & Computational Applications, Vol. 8, No. 2, pp. 253-262, 2003

RULE-BASED FUZZY CLASSIFICATION USING QUERY PROCESSING

Mübariz EMİNOV
Department of Statistics and Computer Science
Mugla University, Mugla ,Turkey
emubariz@hotmail.com

Abstract- This paper describes the derivation of fuzzy classification rules based on c-means fuzzy clustering algorithm as results that are induced of fuzzy clusters. Each fuzzy cluster is associated with a fuzzy classification rule in which fuzzy sets are obtained by projecting the cluster to one-dimensional domains. In order to provide a unique assignment of data to a defined class it is suggested to use the fuzzy query processing executed on the base of induced linguistic fuzzy classification rules. This approach has been applied to fuzzy classification of population where fast and efficient assignment as well as the rank of a data in the same class is supplied.

Keywords- fuzzy cluster, classification, fuzzy set, projection, fuzzy query processing

1. INTRODUCTION

In many application areas such as banking, business, biology, economics, engineering and medical diagnosis etc. sometimes there is huge unstructured numerical data in the form a database. The data analysis is one of new trends in data management in order to gain information from the data being conducted. At present, the well known methods such as statistics, machine learning, neural networks and fuzzy data analysis are being used for exploratory data analysis.

In this paper the issue regarding the construction of rule-based fuzzy classifier is considered. Fuzzy classifier concerns to the accomplishment and classification of the data as per fuzzy rules derived from the fuzzy data analysis. The advantage of the use of fuzzy rules lies in their linguistic interpretability.

Sometimes IF-THEN fuzzy rules containing linguistic definitions can be obtained from experts. Though it may be processed by applying fuzzy set theory proposed by Zadeh [5]. This theory provides the possibility of transforming linguistic description into a mathematical framework in which suitable computation for processing numerical data and formal inference can be carried out. Knowledge acquisition however, is often a very tedious task and the representation of linguistic rules by the fuzzy sets, i.e. the choice of adequate fuzzy set to respective linguistic labels is a severe problem. In many cases, only unstructured data to be classified is available so that neither linguistic classification rules nor their fuzzy sets representation can be defined. For this reason, it is applied to the fuzzy cluster technique that utilizes numerical (crisp) data to be classified in the form of real valued vectors to obtain fuzzy clusters (classes) such as fuzzy sets. Linguistic classification rules are obtained from the derived fuzzy clusters (fuzzy sets) by making the projection of these clusters onto the axes of all domains in the multidimensional space. Extraction of fuzzy linguistic rules from measured data, which contain all object (record) to be the vector $x = (x_1, x_2,.., x_p)$ in database, has received a lot of attention for building of fuzzy classifier by means of such kind of the approach. It is essential, however, to obtain suitable execution of the classification itself according to the delivered fuzzy rules in sense of its solid and quick solution, as well.

In section 2, basic c-means fuzzy clustering algorithm (FLC) to optimal assign of data to clusters (classes) based on a given objective function is briefly reviewed. Section 3 describes shortly the induction of fuzzy rules from fuzzy clusters by projecting these on one-dimensional input domains. After the derivation of linguistic fuzzy classification, rules from these fuzzy rules is discussed. Section 4 is devoted to a final unique assignment of a data to class based on the fuzzy inference. To accomplish classification of a numerical data it is suggested the use of fuzzy query processing for which proposed SQL extension utilizes. Section 5 deals with the implementation of fuzzy classifier concerning classification of individuals according to their performance capacity.

2. FUZZY CLUSTER ANALYSIS

Fuzzy cluster analysis can essentially be categorized as one domain of the data analysis. Its main aim is partitioned a given set of data or objects into clusters (subsets, group, classes). This partition should have two properties: homogeneity within the clusters and heterogeneity between clusters [2]. Since it will be concerned with data in the form of crisp measurements only, i.e. they would be real valued vectors $x = (x_1, x_j, \ldots x_p) \in R^p$, the Euclidean distance between data can be used as a measure of the dissimilarity.

In according to a fuzzy cluster partition, set of an objects $X=\{x_1, x_2, \ldots, x_n\} \in R^p$ is assigned into c fuzzy clusters. Every no exclusive fuzzy cluster is dealt as fuzzy subset of the objects. This means that a partition of set n objects (patterns) into c clusters $1 \leq i \leq c$, is expressed by an $n \times c$ matrix U or (u_{ik}) [4], where $u_{ik} \in [0,1]$ is the membership degree of datum x_k to cluster i. Such kind of partition referred to as c- means fuzzy (probabilistic) clustering should satisfy the following conditions:

$$\sum_{k=1}^{n} u_{ik} > o \qquad \text{for all } i \in \{1,\ldots,c\} \tag{2.1}$$

and

$$\sum_{i=1}^{c} u_{ik} = 1 \qquad \text{for all } k \in \{1,\ldots,n\} \tag{2.2}$$

Fuzzy c-means algorithm (FCM) as a fuzzy version of hard c-means clustering is introduced by Dunn [9] and improved by introduction the fuzzifier m by Bezdek [3]. Thus FCM recognizes spherical clouds of points (datum) in p-dimensional space. Each cluster here is represented by its centre. This is also called a prototype, since it is regarded as a representative of all data assigned to the cluster.

Main issue in fuzzy cluster analysis is to obtain the optimal assignment of data to clusters, in other words, the choice of the optimal cluster centre points (prototypes) for given belongingness of the data to the clusters. This happens usually, by means of the cluster algorithm aimed at minimising the objective function [1,8]

$$J(X, U, V) = \sum_{i=1}^{c} \sum_{k=1}^{n} (u_{ik})^{m} d^{2} (v_i, \chi_k) \tag{2.3}$$

under the constraints (2.1) and (2.2), corresponding to the guarantees that no cluster is completely empty.

$X = \{\chi_1, \chi_2, \ldots \chi_n\} \in R^p$ is the data, c be the number of fuzzy clusters, $u_{ik} \in [0,1]$ is membership degree of datum χ_k to cluster i, $v_i \in R^p$ is the prototype for cluster i and d (v_i, χ_k) be the Euclidian distance between prototype v_i and datum χ_k. The parameter $1 < m$ is called fuzziness index and, usually $m = 2$ is chosen.

The quadratic distance of the data to the prototypes $d_{ik} = \| \chi_k - v_i \|$ weighted with their membership degrees is used for minimizing (2.3). For this reason, the prototypes of the cluster centers v_I, are calculated by the following equation:

$$v_i = \frac{\sum_{k=1}^{n} (u_{ik})^{m} \chi_k}{\sum_{k=1}^{n} (u_{ik})^{m}} \tag{2.4}$$

as a necessary condition for (2.3) to have a local minimum. After randomly initialization of the partition matrix (u_{ik}), the prototypes v_i and new matrix (u_{ik}) at each optimization step are updated according to (2.4) and followed as

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(d^2(v_i,\chi_k) / d^2(v_j,\chi_k) \right)^{2/(m-1)}} \tag{2.5}$$

This iteration procedure is proceed until successive approximation $\| v^{(t-1)} - v^{(t)} \| \leq \varepsilon$ is stabilized.

The most important problem in clustering is to determine the optimal number of clusters, when the number of clusters is not known, in advance. Hence the number of classes is unknown, as well.

This paper is concerned with unsupervised classification where the just mentioned knowledge is unavailable. In this case, for each $c \in \{2,3,..c_{max}\}$, has to carry out the fuzzy cluster analysis in order to find an optimal partition of data with respect to the new correspondent objective function. Beginning from c = 2 for each partition, a value such that the results of the clustering analysis can be estimated with respective objective function according to (2.3). Since this function is regarded as a validity function because it is decreasing for an increasing c. The found optimal number of cluster then coincides with the number of classes that will be considered in the following section.

3. DERIVATION RULES FROM THE FUZZY CLUSTER

In the previous section the way to determine the membership matrix U obtained from fuzzy clustering analysis have been discussed. This matrix contains the memberships not all of the objects of the data set (only training part of the data) in each of the found clusters. Therefore to represent all possible data the discrete membership matrix has to be extended to a continuous membership function. These membership functions are used to describe fuzzy If-Then classification rules. The aspects of derivation of these rules will be given in present section.

The essential idea of derivation of classification rules from fuzzy clusters is the following. Each fuzzy cluster is assumed to be assigned to one class. The membership degrees of the data to the clusters determine the extent to which classes they belong, as a member of the corresponding class. Therefore obtained fuzzy clusters may be associated with a linguistic classification rules in classifiers. To this aim, each of the fuzzy clusters defined in multidimensional domains is projected into one-dimensional domains leading to a fuzzy set on the real numbers. To constitute the fuzzy set that is the γ^{th} projection of the cluster from corresponding to optimal discrete membership matrix (u_{ik}) may be used in the following equation [1]:

$$\mu_\gamma(\chi_j) = \sup \{u_{ik} \mid x = (\chi_1,\dots\chi_j\dots\chi_p) \in R^p \} \tag{3.1}$$

As seen from (3.1), this $\mu_\gamma(\chi_j)$ fuzzy set is non-convex set or fuzzy number because in projecting the training data are used only. Therefore the convex membership function has to be computed after projecting or to approximating it by a trapezoidal or triangular continuous membership function as proposed in [7].

To fuzzy set in the projection space defined with corresponding continuous membership functions are assigned linguistic labels like small, weight, tall, etc. It is very easily in comparison to assignment linguistic labels to membership functions [6] with high dimensional domains because projections offer a higher transparency and interpretability. For this reason, the high dimensional membership function of the cluster is represented as the form of conjunction of these linguistic labels in the premise of the corresponding classification rule. The conclusion part of this rule is the class to which the cluster is assigned. Such kind of representation of the premise of classification rule is formulated as the Cartesian product of the corresponding one-dimensional fuzzy set as description of a first part of the relative class. The Cartesian product is described bellow.

If to take into account that the object is p-dimensional real vector $x = (\chi_1,\chi_2,\dots\chi_j,\dots,\chi_p)$ then classification rule $R \in \Re$ ($\Re$ be a finite set of possible rules) may be written as

$$\text{R: If } \chi_1 \text{ is } \mu_R^{(1)} \text{ and } \dots \chi_j \text{ is } \mu_R^{(j)} \text{ and } \dots \chi_p \text{ is } \mu_R^{(p)} \text{ then class is } C_R \tag{3.2}$$

where $C_R \in C$ is one of the finite set C classes.

Fuzzy sets $\mu_R^{(1)}, \ldots, \mu_R^{(j)}, \ldots, \mu_R^{(p)}$ are defined in the universe of course (domains) $X_1, X_2, \ldots X_p$, respectively i.e. $\mu_R^{(j)} : X_j \to [0,1]$. The Cartesian product of these fuzzy sets is also fuzzy set such as

$$\mu_R(\chi_1 \ldots \chi_j \ldots \chi_p) = \mu_R^{(1)} \times \mu_R^{(2)} \times \ldots \times \mu_R^{(j)} \times \mu_R^{(p)}$$

defined in the product space $X_1 \times X_2 \times \ldots \times X_j \times \ldots \times X_p$ where $X = X_1 \times X_2 \times \ldots \times X_j \times \ldots \times X_p$.

During the classification process according to corresponding fuzzy rules is accomplished a partial mapping

$$\text{Class: } X_1 \times X_2 \times \ldots \times X_j \times \ldots \times X_p \to C$$

that assigns classes to some vectors $\{x = (\chi_1,\chi_2 \ldots \chi_j \ldots \chi_p) \in X_1 \times X_2 \times \ldots \times X_j \times \ldots \times X_p\}$

In order to gain the linguistic classification rules from (3.2) fuzzy sets $\mu_R^{(j)}$, have to be replacing by suitable linguistic values, as mentioned above.

It should be noted that fuzzy rules induced by means of the projection method, however, in general does not yield the same results as the original rule with multidimensional membership function, since it is an approximation of the latter.

In fuzzy rule R (see (3,2)) by replacing the fuzzy sets $\mu_R^{(1)}, \mu_R^{(2)}, \ldots, \mu_R^{(j)}, \ldots, \mu_R^{(p)}$ with corresponding linguistic labels assigned to them, is provided linguistic classification rule. Such form of rules is very useful because of its interpretability and transparency. Therefore, fuzzy classification will be accomplished on base such kind of rules. As mentioned above, these fuzzy sets are trapezoidal or triangular types of sets (after approximation) and as membership function associated with correspondent linguistic labels.

4. FUZZY CLASSIFICATION BY USING FUZZY QUERYING

4.1 Fuzzy Inference for Classification

For many classification problems, a unique assignment of an object to class is required. The assignment, in general, is based on the mapping the class to the relative object presented by vector $x = (\chi_1,\chi_2 \ldots \chi_j \ldots \chi_p)$. As a result of matching the object with regard to classification rules one will be assigned to corresponding class if that object possesses highest degree of membership to them. It is the unique assignment that corresponds to a defuzzification process that simply chooses the mentioned class. Hence, to achieve this aim we have to apply compositional operators (rules), which have mainly four types. Due to simplicity in computation max-min compositional rule is used, i.e. the conjunction in the rules is evaluated by the Mamdani mini rule (intersection operation AND) and the result of the rules is aggregated by the maximum (union operation OR) [2,11]

Therefore, the conjunction with respect to rule R according (3.2) will be defined as follows:

$$\mu_R(x_1,...x_j...x_p) = \min_{j\in\{1,...,p\}} \{\mu_R^{(j)}(x_j)\} \quad (4.1)$$

Thus, by using equation (4.1) the membership degree is determined to which the premise of rule R or belongingness of the object to class C_R is satisfied. The membership degree to which the vector $x = (x_1,x_2...x_j...x_p)$ is assigned to class $C \in \mathcal{C}$ can be determined by the following equation

$$\mu_C^{(R)}(x_1,...,x_p) = \max \{\mu_C(x_1,..x_j...x_p) \mid C_R = C\} \quad (4.2)$$

The defuzzification operation, that is, the final assignment of a unique class to given vector $x = (x_1,x_2...x_j...x_p)$ is carried out by the mapping over rule base $\Re$

$$\Re(x_1, ...x_p) = \begin{cases} C & \text{if } \mu_C^{(R)}(x_1,...,x_p) > \mu_D^{(R)}(x_1,...,x_p) \\ & \text{for all } D \in \mathcal{C}, D \neq C \\ \text{unknown} \notin \mathcal{C} & \text{otherwise} \end{cases} \quad (4.3)$$

Finally, by applying all the data to rules $\Re$ for defuzzification is determined the subset of objects that are assigned to class C. This may be expressed by equation:

$$\Re^{-1}(C) = \{(x_1,..., xp) \mid \Re(x_1,..., xp) = C\}$$
(4.4)

4.2 Fuzzy Classification Using Fuzzy Querying

In previous section we dealt with derivation of fuzzy linguistic rules based on fuzzy clustering analysis of the numerical data that contain set of the objects $X = \{x_1,x_2,...,x_n\} \in R^p$ each of which is presented by vectors $x = (x_1,x_2...x_j...x_p) \in R^p$. Each of these vectors being feature or attribute vector comprises crisp values of all attributes $\{ x_j \}$. Relevant numerical data in many applications is huge and unstructured and therefore it is accumulated usually in the form of relational database. Obviously, the multidimensional data in database may be treated to be matrix type data or 2-way data. This 2-way data consists of object and attributes represented as 0 x A (object x attributes) or matrix $\{ x_{ji} \}$.

In ordinary IF-THEN rule-based classifier for assignment of the data to relative classes, as is easily understood after section 4.1, each first feature vector $x = (x_1,...x_p)$ is matched with each of linguistic rules successively and afterwards results of fuzzy matchings are estimated according to (4.2) and (4.3) for execution of fuzzy inference.

In this paper, in order to construct the fast and efficient fuzzy classifier, it is suggested to accomplish classification of the numerical data using fuzzy query processing. Reasons for this suggestion are to support the form of relational database representation of the data to be classified and the existence of standard many SQL (Structured Query Language) tools. These SQLs supports quickly searching and retrieving with respect to crisp data in according to crisp query. Each query here contains search criterion that involves p-numbers attributes presented through their numerical values. However, as seen, crisp query and therefore its search criterion cannot be applied to fuzzy classification. Fuzzy query can be contained in only search criterion that consist in fuzzy predicates in each of which the attribute may be presented through linguistic values like good, weight, long, etc. Therefore, received fuzzy search criterion will correspond to premise of the linguistic classification rules. For instance, the form of the fuzzy query will be as following

Select*
From fuzzy
Where Age is Young and Height is Tall and Weight is Leigh

which can be interpreted as following: from database "Fuzzy" to select the individuals on three attributes with relevant linguistic values. Obviously, current SQLs do not support above considered imprecise query with respect to crisp data because its grammar does not provide the use of the fuzzy (imprecise) predicates. Few different extension to SQL such as QUEL, SQL^f, etc. which may tackle with fuzzy query processing was proposed [10].

As proposed by us new extension to current SQL [8] have been adopted to assign the crisp data to relevant classes according to derivated fuzzy linguistic rules. The principal idea of fuzzy classification based on fuzzy query processing is the following. Firstly, all the feature vectors $X = \{x_1 x_2 ..., x_n\} \in R^p$ are successively matched with the premise of all the linguistic classification rules and accordingly to (4.1) degree of meeting for each search criterion, i.e. for premise of relevant rule are calculated. Secondly, with respect to all object (records) the results of matching (fire strengths) along the same class $C \in \mathcal{C}$ are estimated after (4.2). Afterwards the results of previous estimation with respect to the classes are processed according to (4.3) and the objects are uniquely assigned to relative class $C \in \mathcal{C}$. Finally, the subsets of the objects associated with relevant classes are composed.

As presented by us extended SQL by means of interface and application procedures are added to them, the four steps of fuzzy inference dealing with the assignment of numerical data to relevant classes. As it is noted above, linguistic labels in the premise of the classification rules are associated with relative fuzzy sets. Hence, during fuzzy inference the objects (feature vectors) are matched with the fuzzy sets that take place in classification rule or search criterion to one of which corresponds relative the membership function.

Utilizing SQL's manipulation procedures fuzzy classifier may supply the assignment of subset of the objects to class ranking them within the same class. That happens in according to degree of belongingness for satisfaction of relevant class.

4.3 Implementation of Fuzzy Classifier

This subsection deals with implementation of fuzzy classifier, for instance, regarding classification of population (individuals) in sense their performance capacity with respect to candidates for the basketball team.

Each individual in the training data set is represented as real valued vector x = $(x_1, x_2, x_3) \in R^3$ including three variables such as Age, Height and Weight attributes, respectively. This data set contains 110 individuals (objects) or feature vectors.

First by c-means fuzzy clustering algorithm we find the optimal number of clusters to be 3 in product space. Since we are concerned with the unsupervised clustering after these are assigned class labels fire "High Performance", "Middle Performance" and "Low Performance". As stated from the results, to second and third are associated by 2 clusters in the input (projection) space. By projecting the obtained clusters on the three input domains, we yield non-convex and after approximation the convex fuzzy sets over respective projection the convex fuzzy sets over respective projection spaces x_1, x_2, x_3 with respect to each fuzzy clusters. Obviously, these fuzzy sets are continuous membership functions are presented according to (3.2) in the respective fuzzy rules. The number of fuzzy rules to be equal to the number of cluster in the projection space is five (see Fig. 4.1). As shown in Fig. 4.1, the membership functions have triangular and trapezoidal types. These membership functions after the functionally definition took place in application program for fuzzy matching.

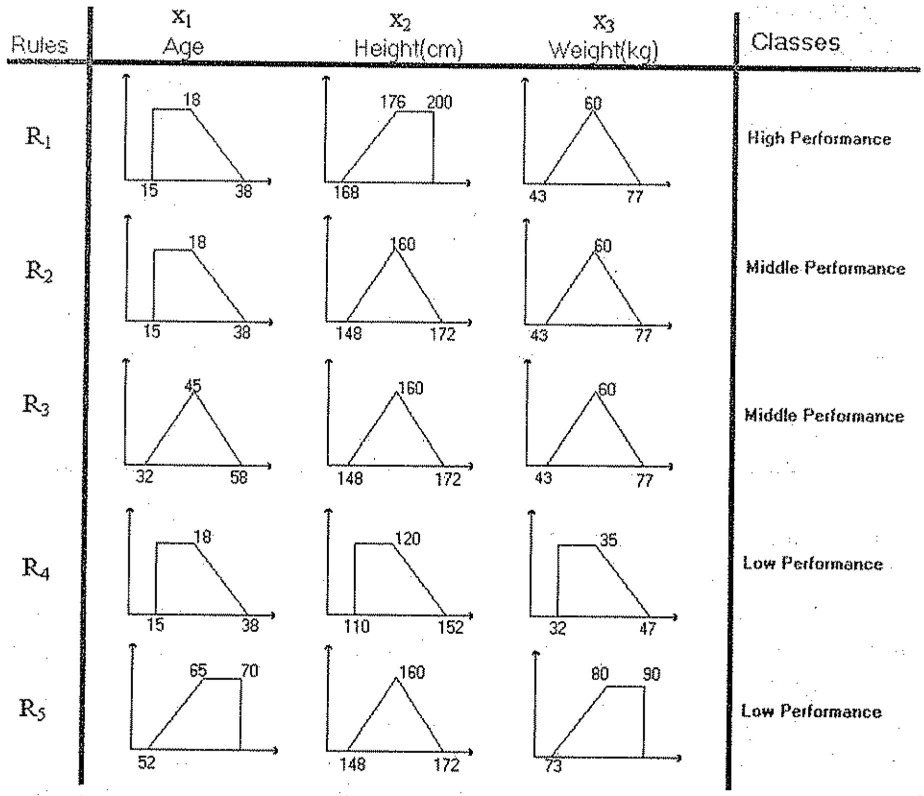


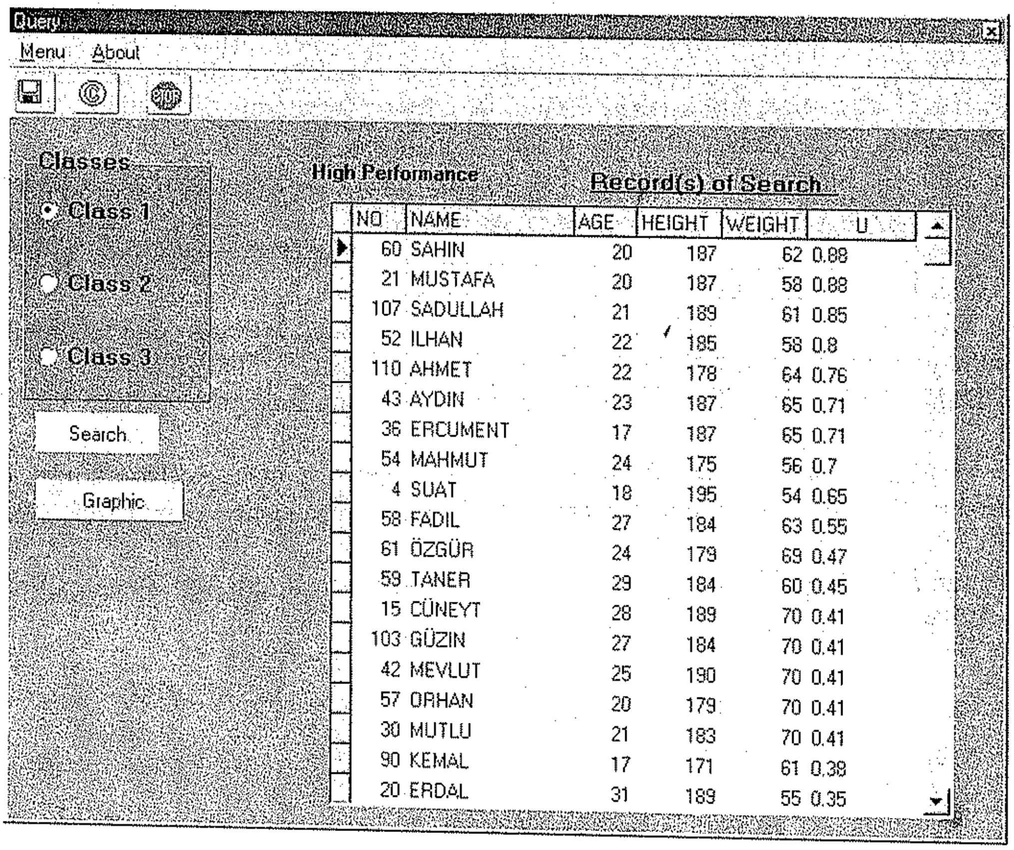
Fig. 4.1. Obtained Fuzzy Classification Rules

After assigning the respective labels to mentioned fuzzy membership functions, has been obtained linguistic fuzzy classification rules presented such as

Rule 1: IF x_1 is Young and x_2 is Tall and x_3 is Middle Weight
THEN individual belongs to High Performance

Rule 2: IF x_1 is Young and x_2 is Middle Height and x_3 is Middle Weight
THEN individual belongs to Middle Performance
Rule 3: IF x_1 is Middle Age and x_2 is Middle Height and x_3 is Middle Weight
THEN individual belongs to Middle Performance
Rule 4: IF x_1 is Young and x_2 is Short and x_3 is Light
THEN individual belongs to Low Performance
Rule 5: IF x_1 is Old and x_2 is Middle Height and x_3 is Heave
THEN individual belongs to Low Performance

By setting five search criterion successively in according to premise part of the derivated above linguistic rules fuzzy classifier supplies the assignment of the individual to be classified to one of three classes. As clearly seen in Fig. 4.2, the subset of the individuals assigned to class "High Performance" is reported as table after searching. The list of the assigned individuals is given through them ranking according to degree of satisfaction for relevant in a descending direction.



NO	NAME	AGE	HEIGHT	WEIGHT	U
60	SAHIN	20	187	62	0.88
21	MUSTAFA	20	187	58	0.88
107	SADULLAH	21	189	61	0.85
52	ILHAN	22	185	58	0.8
110	AHMET	22	178	64	0.76
43	AYDIN	23	187	65	0.71
36	ERCUMENT	17	187	65	0.71
54	MAHMUT	24	175	56	0.7
4	SUAT	18	195	54	0.65
58	FADIL	27	184	63	0.55
61	ÖZGÜR	24	179	69	0.47
59	TANER	29	184	60	0.45
15	CÜNEYT	28	189	70	0.41
103	GÜZIN	27	184	70	0.41
42	MEVLUT	25	190	70	0.41
57	ORHAN	20	179	70	0.41
30	MUTLU	21	183	70	0.41
90	KEMAL	17	171	61	0.38
20	ERDAL	31	189	55	0.35

Fig. 4.2. Computer Report for Classification to "High Performance"

Proposed fuzzy classification has been implemented successfully in Dbase for Windows Database Management System. Interface to standard SQL and other application programs have been developed in environment of Delphi 4.0.

5. CONCLUSION

By applying fuzzy clustering to the numerical data to be classified the classifications is obtained that have the information inherent in membership degrees to which it is able to judge the object to classes. The use of fuzzy querying for classification provides the fast and efficient assignment of the data that is best suitable especially for data mining dealing with huge data sets. On the other hand, it supplies the rank of a datum in the same class according to degree of belongingness to them.

REFERENCES

1. F.Klawonn and R. Kruse, Derivation of Fuzzy Classification Rules from Multidimensional Data, The International Institute for Advanced Studies in System Research and Cybernetics, Windsor, Ontario, 90-94, 1995
2. F. Hoppner, F. Klawonn, R. Kruse and T. Runkler, Fuzzy Cluster Analysis, Chichester, New York, John Wiley, 1999
3. J. C. Bezdek, Fuzzy Mathematics in Pattern Classification, Ph. D. Thesis, Applied Math. Center, Cornell University, Ithaca, 1973
4. J. C. Bezdek, J. Keller, R. Krisnapuram and N. R. Pal, Fuzzy Models and Algorithms for Pattern Recognition and Image Processing, Kluwer Academic Publishers, 1999
5. L.A. Zadeh, Fuzzy Sets, Information and Control, vol. 8, 338-353, 1965
6. L.A. Zadeh, Fuzzy Set. Theoretic Interpretation of Linguistic Hedges, Cybernetics, vol.2, 4-34. 1972
7. M. Sugeno and T. Yasucawa, A. Fuzzy-Logic-Based Approach to Qualitative Modeling, IEEE Transactions on Fuzzy Systems, vol.1, no.1, 7-31,1993
8. M. Eminov, Querying a Database by Fuzzification of Attribute Values, 5.National Econometrics and Statistics Symposium, September, 19-22, Adana, 2001
9. R. Duda and H. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1973
10. W.H. Mansfield and R.M. Fleischman, A. High Performance, Ad Hoc, Fuzzy Query Processing System, Intelligent Information Systems, vol.2, no. November, 397-419, 1993
11. Y. Jan and M. Ryan, Using Fuzzy Logic: Forwards in Intelligent Systems, Prentice Hall, 1994