J. Indian Inst. Sci., Sep.–Oct. 2005, **85**, 265–277

A comparative study of fuzzy and neural network approaches to discriminant analysis with linguistic variables

CHANDAN CHAKRABORTY AND DEBJANI CHAKRABORTY*
Department of Mathematics, Indian Institute of Technology, Kharagpur 721 302, India.
email: debjani@maths.iitkgp.ernet.in

Received on June 15, 2005; Revised on August 22, 2005

Abstract

This paper proposes a fuzzy discriminant analysis to solve the two-group classification problem where the measured variables are linguistic in nature. Especially under imprecise framework, the linguistic variables capture more information although vagueness is inherent. In analogy to classical statistics, a fuzzy linear discriminant function is introduced here, which directly deals with continuous fuzzy numbers as the representative of linguistic values to obtain fuzzy scores for classification. To make a comparative study, the backpropagation neural network approach has also been studied in this paper. Finally admission to management programme is considered as an example of the application on two-level classification problem of the proposed method.

Keywords: Linguistic variable, fuzzy number, linear fuzzy discriminant analysis, neural network.

1. Introduction

Discriminant analysis [1] for hard classification has been widely applied for effective decision-making in many real-world problems for the last few decades. It becomes more effective for the following advantages: firstly, separation of classes as much as possible based on the measured variables and secondly, classification of a new entity into a labeled class. As a multivariate data analytic technique, the discriminant analysis is usually strongly recommended for classification problems with precise data. But it becomes problematic if at least some of the variables are linguistic (i.e. qualitative) while designing the classification problems under imprecise environment. It must be logically accepted that the variables assessing the values as outcomes of human factors, especially experience, perception, thinking, reasoning and attitude, etc. become fuzzy. In practice, linguistic variables possess the linguistic or fuzzy values represented in terms of natural languages due to their flexibility and simplicity although vagueness and ambiguity are inherent. Therefore, there is a need of developing fuzzy discriminant technique for the classification problems on the basis of linguistic variables under the paradigm of 'fuzzy statistics'. The linguistic values are perfectly quantized by fuzzy sets and subsequently fuzzy numbers [2, 3]. In doing so this paper introduces a fuzzy approach to solve a two-group classification problem in discriminant analysis. Firstly, the fuzzy discriminant scores for each of the candidates computed and then defuzzified to compare with the defuzzified threshold value for classification. In fact,

*Author for correspondence.

we use here centroid method [3] for defuzzification. Among many statistical and nontraditional techniques of classification, neural network has been widely used for classification [4–6]. In many situations, neural network approximates continuous fuzzy data into discrete inputs. In this connection, we have studied a backpropagation neural network approach to make a comparative study where the linguistic data are initially defuzzified. Here, we have considered a two-level classification problem to study whether the students are eligible for admission or not in the management programme.

Some research developments on fuzzy discriminant analysis have been done recently. Lin *et al*. [7] have shown a fuzzy method for two-group discriminant analysis where the membership functions of the groups to be discriminated is obtained by minimizing the sum of squares of classification errors. A method for performing fuzzy multiple discriminant analysis [8] on groups of crisp data is proposed, which is able to detect membership function of each group by minimizing the classification error using genetic algorithm. A fuzzy mathematical programming approach to develop fuzzy linear discriminant function has been devised by Chiang *et al*. [9] for separable as well as nonseparable data sets, which is not at all based on fuzzy variables. Chen *et al*. [10] have proposed a discrimination technique for chemical data sets with a few overlapping data points that are considered equally important for all classes in ordinary discriminant analysis. Watada *et al*. [11] have proposed a fuzzy discrimination method only for fuzzy data in fuzzy groups. A Hopfiled neural network approach [6] for classification is introduced where the incomplete pattern is first translated into fuzzy terms, but these terms have been discretized. In fact, all the techniques for fuzzy classification are not subjected to the linguistic data at all. Rather the methods consider degrees of membership in discrete form that is nothing but oversimplification.

Section 2 describes the definitions and notations of fuzzy variable, *LR*-type fuzzy number in Sub-sections 2.1 and 2.2, respectively. In Sub-section 2.3, the variance-covariance for fuzzy variables is introduced. A fuzzy statistical approach to discriminant analysis with linguistic variables is described in Sub-section 3.1. Sub-section 3.2 highlights the backpropagation neural network framework for classification with linguistic input. Finally, a two-level classification problem whether the students are eligible for admission or not in the management programme is considered in Section 4. Conclusions are drawn in Section 5.

2. Preliminaries

2.1. *Fuzzy variable*

A fuzzy variable is a variable whose values are fuzzy. The concept of fuzzy variable [3] is very useful in situations where decision problems are too complex or too ill-defined to be described properly using conventional quantitative expressions. For example, the ability, performance ratings, etc. could be well expressed using fuzzy values such as very poor, poor, fair, good, very good, excellent, etc.

2.2. *LR-type fuzzy number*

Fuzzy set introduced by Zadeh [2] is a gradual transition from nonmembership to fullmembership. A fuzzy set $\tilde{X} = (m, a, b)_{LR}$ is said to be a *LR*-type fuzzy number [3] where '*L*' and '*R*' stand for left and right references if

(i) $m_{\tilde{X}}$ is bounded and upper semicontinuous;
(ii) the membership function $m_{\tilde{X}}$ is of the form

$$m_{\tilde{X}}(x) = \begin{cases} L\left(\frac{m-x}{a}\right) & \text{for } m-a \le x < m \\ R\left(\frac{x-m}{b}\right) & \text{for } m < x \le m+b \end{cases}$$

Now *LR*-type fuzzy number reduces to a triangular fuzzy number if $L(y) = R(y) = \max\{0, |1-y|\}$. Without loss of generality we consider *LR*-type fuzzy numbers as fuzzy realizations of the fuzzy variables throughout the paper.

2.3. *Fuzzy variance–covariance*

In classical sense, variance measures the dispersion around the central point of a set of observations, which is computed on the basis of crisp observations. But when we have fuzzy data, then computation of crisp variance–covariance is really oversimplification. Human intuition says that if the observations are vague (i.e. unstable) then variance–covariance will definitely be imprecise in nature. In view of this, a concept of computing fuzzy variance and fuzzy covariance for fuzzy variables based on fuzzy arithmetic [3] is introduced in this paper. Let us consider two fuzzy variables, say, $\tilde{X} = (m_x,\ a_x,\ b_x)_{LR}$ and $\tilde{Y} = (m_y,\ a_y,\ b_y)_{LR}$ defined on the universe say, U. A sample $\{(\tilde{x}_i, \tilde{y}_i)\}$ of size 'n' is drawn on $\tilde{X}$ and $\tilde{Y}$, respectively. Now we define fuzzy variance by $\tilde{S}_{\tilde{X}}$ and fuzzy covariance by $\tilde{S}_{\tilde{X},\tilde{Y}}$ as follows:

Fuzzy means:

$$\bar{\tilde{X}} = (\bar{m}_x, \bar{a}_x, \bar{b}_x)_{LR} \text{ where } \bar{m}_x = \frac{1}{n}\sum_{i=1}^{n} m_{x_i};\ \bar{a}_x = \frac{1}{n}\sum_{i=1}^{n} a_{x_i} \text{ and } \bar{b}_x = \frac{1}{n}\sum_{i=1}^{n} b_{x_i} \tag{1}$$

Fuzzy variance:

$$\tilde{S}_{\tilde{X}} = \frac{1}{n-1}\sum_{i=1}^{n}(\tilde{X}_i - \bar{\tilde{X}})^2 = \frac{1}{n-1}\sum_{i=1}^{n}[(m_{x_i}, a_{x_i}, b_{x_i}) - (\bar{m}_x, \bar{a}_x, \bar{b}_x)]^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}[(m_{x_i} - \bar{m}_x,\ \ a_{x_i} + \bar{a}_x,\ \ b_{x_i} + \bar{b}_x)]^2 = (s_x,\ \tilde{a}_x,\ \tilde{b}_x)_{LR} \tag{2}$$

where $s_x = \frac{1}{n-1}\sum_{i=1}^{n}(m_{x_i} - \bar{m}_x)^2;\ \tilde{a}_x = \frac{2}{n-1}\sum_{i=1}^{n}(m_{x_i} - \bar{m}_x)(a_{x_i} + \bar{a}_x);$

$$\tilde{b}_x = \frac{2}{n-1}\sum_{i=1}^{n}(m_{x_i} - \bar{m}_x)(b_{x_i} + \bar{b}_x)\ .$$

Therefore,

Fuzzy covariance: $\tilde{S}_{\tilde{X},\tilde{Y}} = \frac{1}{n-1}\sum_{i=1}^{n}(\tilde{X}_i - \bar{\tilde{X}})(\tilde{Y}_i - \bar{\tilde{Y}})$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left[\begin{pmatrix}(m_{x_i}-\bar{m}_x)(m_{y_i}-\bar{m}_y),\ (m_{x_i}-\bar{m}_x)(a_{y_i}+\bar{a}_y)+(m_{y_i}-\bar{m}_y)(a_{x_i}+\bar{a}_x),\\ (m_{x_i}-\bar{m}_x)(b_{y_i}+b_y)+(m_{y_i}-\bar{m}_y)(b_{x_i}+\bar{b}_x)\end{pmatrix}_{LR}\right]$$

$$= (s_{xy},\ \tilde{a}_{xy},\ \tilde{b}_{xy})_{LR} \tag{3}$$

where $s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(m_{x_i}-\bar{m}_x)(m_{y_i}-\bar{m}_y)$;

$$\tilde{a}_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}[(m_{x_i}-\bar{m}_x)(a_{y_i}+\bar{a}_y)+(m_{y_i}-\bar{m}_y)(a_{x_i}+\bar{a}_x)],\ \text{and}$$

$$\tilde{b}_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}[(m_{x_i}-\bar{m}_x)(b_{y_i}+\bar{b}_y)+(m_{y_i}-\bar{m}_y)(b_{x_i}+\bar{b}_x)]\,.$$

While making a suitable decision for certain purpose a decision-maker should emphasize on precise and concrete decision irrespective of fuzzy or crisp environment. Here we have used the well-known centroid method to defuzzify the computed fuzzy variance–covariance for computing its inverse as follows:

$$S_{\tilde{X}} = \frac{\int_U x m_{\tilde{X}}(x)dx}{\int_U m_{\tilde{X}}(x)dx} = \frac{\int_{s_x-\tilde{a}_x}^{s_x} x\,L\left(\frac{s_x-x}{\tilde{a}_x}\right)dx + \int_{s_x}^{s_x+\tilde{b}_x} x\,R\left(\frac{x-s_x}{\tilde{b}_x}\right)dx}{\int_{s_x-\tilde{a}_x}^{s_x} L\left(\frac{s_x-x}{\tilde{a}_x}\right)dx + \int_{s_x}^{s_x+\tilde{b}_x} R\left(\frac{x-s_x}{\tilde{b}_x}\right)dx} \tag{4}$$

and

$$S_{\tilde{X}\tilde{Y}} == \frac{\int_{s_{xy}-\tilde{a}_{xy}}^{s_{xy}} t\,L\left(\frac{s_{xy}-t}{\tilde{a}_{xy}}\right)dt \ + \int_{s_{xy}}^{s_{xy}+\tilde{b}_{yx}} t\,R\left(\frac{t-s_{xy}}{\tilde{b}_{xy}}\right)dt}{\int_{s_{xy}-\tilde{a}_{xy}}^{s_{xy}} L\left(\frac{s_{xy}-t}{\tilde{a}_{xy}}\right)dt \ + \int_{s_{xy}}^{s_{xy}+\tilde{b}_{yx}} R\left(\frac{t-s_{xy}}{\tilde{b}_{xy}}\right)dt} \tag{5}$$

3. Methodological development

3.1. *A fuzzy statistical approach to fuzzy linear discriminant analysis* (*FLDA*)

As a soft computing tool, fuzzy set theory [2] has been well established to deal with vagueness, which is especially reflected in all the natural languages and artificial intelligencerelated problems. In such situations, while human cognitive aspects like experience, knowledge, and reasoning drive the classification problems, the data become imprecise or fuzzy in nature. Practically, there are many real-world problems where some of the decision variables are purely fuzzy in nature. In those cases, considering the crisp values, instead of fuzzy values, may be crude oversimplification. Fuzzy logic has the power of considering the whole content of fuzzy values, which are represented by fuzzy numbers. Here a paradigm of classification problem consisting of all fuzzy variables is considered. Firstly, a number of groups is considered and hence collect the sample fuzzy data of different sizes for each of the groups. Now, the classical discriminant analysis is extended to fuzzy discriminant analysis for fuzzy data. Fuzzy discriminant analysis is a fuzzy statistical approach to model with fuzzy variables. Let us design the problem interface as follows: (a) a number of fuzzy variables is considered; and (b) a new entity is to be classified into one of the labeled groups. For the sake of simplicity we assume here only two classes, say C_1 and C_2. All the variables, say $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_m$ are considered here as fuzzy where a crisp value can be treated as fuzzy number with zero vagueness. The fuzzy data of fuzzy variables can be semantically represented as follows:

Let us consider a fuzzy vector as $\underline{\tilde{X}} = [\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_m]^T$ such that $\tilde{X}_j^{(k)} = (m^{(k)}_{X_j}, a^{(k)}_{X_j}, b^{(k)}_{X_j})_{LR}$ and $\tilde{x}_{ij}^{(k)} = (m^{(k)}_{x_{ij}}, a^{(k)}_{x_{ij}}, b^{(k)}_{x_{ij}})_{LR}$, where $\tilde{x}_{ij}^{(k)}$ denote the fuzzy opinion of the i^{th} (i = 1, 2, ..., n) person for the j^{th} (j = 1, 2, …, m) criterion in the k^{th} (here k = 1, 2) group. Two groups are to be discriminated based on 'm' triangular fuzzy variables. The key idea here is to transform the multifuzzy variables ($\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_m$) into univariate fuzzy variable $\tilde{Z}$ such that $\tilde{Z}$'s derived from two groups are separated as much as possible. In doing so, the fuzzy means using eqn (1) are computed as follows:

$$\tilde{\bar{X}}_j = \left(m^{(k)}_{\bar{X}_j},\ a^{(k)}_{\bar{X}_j},\ b^{(k)}_{\bar{X}_j}\right)_{LR} \text{ and } \tilde{\bar{X}}_t = \left(m^{(k)}_{\bar{X}_t},\ a^{(k)}_{\bar{X}_t},\ b^{(k)}_{\bar{X}_t}\right)_{LR} \tag{6}$$

Table I
Fuzzy values of fuzzy variables

Sample	Class	$\tilde{X}_1$	$\tilde{X}_2$	$\cdots$	$\tilde{X}_m$
1		$\tilde{x}_{11}^{(1)}$	$\tilde{x}_{12}^{(1)}$	$\cdots$	$\tilde{x}_{1m}^{(1)}$
2		$\tilde{x}_{21}^{(1)}$	$\tilde{x}_{22}^{(1)}$	$\cdots$	$\tilde{x}_{2m}^{(1)}$
$\vdots$	C_1	…………			
$\vdots$		$\tilde{x}_{i1}^{(1)}$	$\tilde{x}_{i2}^{(1)}$	$\cdots$	$\tilde{x}_{im}^{(1)}$
i		…………			
$\vdots$					
n_1		$\tilde{x}_{n_1 1}^{(1)}$	$\tilde{x}_{n_1 2}^{(1)}$	$\cdots$	$\tilde{x}_{n_1 m}^{(1)}$
1		$\tilde{x}_{11}^{(2)}$	$\tilde{x}_{12}^{(2)}$	$\cdots$	$\tilde{x}_{1m}^{(2)}$
2		$\tilde{x}_{21}^{(2)}$	$\tilde{x}_{22}^{(2)}$	$\cdots$	$\tilde{x}_{2m}^{(2)}$
$\vdots$	C_2	…………			
$\vdots$		$\tilde{x}_{i1}^{(2)}$	$\tilde{x}_{i2}^{(2)}$	$\cdots$	$\tilde{x}_{im}^{(2)}$
i		…………			
$\vdots$					
n_2		$\tilde{x}_{n_2 1}^{(2)}$	$\tilde{x}_{n_2 2}^{(2)}$	$\cdots$	$\tilde{x}_{n_2 m}^{(2)}$

where $m_{\bar{X}_j}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} m_{x_{ij}}^{(k)},\ a_{\bar{X}_j}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} a_{x_{ij}}^{(k)},\ b_{\bar{X}_j}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} b_{x_{ij}}^{(k)}$

$$m_{\bar{X}_t}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} m_{x_{it}}^{(k)},\ a_{\bar{X}_t}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} a_{x_{it}}^{(k)},\ b_{\bar{X}_t}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} b_{x_{it}}^{(k)}\ . \tag{7}$$

Therefore, the fuzzy mean vector and covariance matrix for the k^{th} group based on observed fuzzy observations can be theoretically calculated using equations (1)–(5) as follows:

$$\bar{\tilde{X}}^{(k)} = \begin{bmatrix} \bar{\tilde{X}}_1^{(k)} \\ \bar{\tilde{X}}_2^{(k)} \\ \vdots \\ \bar{\tilde{X}}_m^{(k)} \end{bmatrix} = \begin{bmatrix} \left(m_{\bar{X}_1}^{(k)},\ a_{\bar{X}_1}^{(k)},\ b_{\bar{X}_1}^{(k)}\right)_{LR} \\ \left(m_{\bar{X}_2}^{(k)},\ a_{\bar{X}_2}^{(k)},\ b_{\bar{X}_2}^{(k)}\right)_{LR} \\ \vdots \\ \left(m_{\bar{X}_m}^{(k)},\ a_{\bar{X}_m}^{(k)},\ b_{\bar{X}_m}^{(k)}\right)_{LR} \end{bmatrix} \text{ and } S^{(k)} = ((s_{jt}^{(k)}))_{m\times m} = \begin{bmatrix} s_{11}^{(k)} & s_{12}^{(k)} & \cdots & s_{1m}^{(k)} \\ s_{21}^{(k)} & s_{22}^{(k)} & \cdots & s_{2m}^{(k)} \\ \cdots & \cdots & \cdots & \cdots \\ s_{m1}^{(k)} & s_{m2}^{(k)} & \cdots & s_{mm}^{(k)} \end{bmatrix}$$

Hence, the objective is to select the linear combination of fuzzy variables to achieve maximum separation of fuzzy sample means $\bar{\tilde{Z}}_1$ and $\bar{\tilde{Z}}_2$ for C_1 and C_2, respectively. Let us consider the fuzzy discriminant function as a linear combination of $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_m$ i.e.

$$\tilde{Z} = \tilde{b}_1 \otimes \tilde{X}_1 \oplus \tilde{b}_2 \oplus \tilde{X}_2 + \cdots \oplus \tilde{b}_m \otimes \tilde{X}_m = \tilde{b}' \otimes \tilde{X}\ , \tag{8}$$

where '$\oplus$' and '$\otimes$' denote the extended sum and multiplication operator, respectively.

Suppose n_1 and n_2 numbers of fuzzy responses are sampled from the two groups say, $(\tilde{Z}_{11}, \tilde{Z}_{12}, \cdots, \tilde{Z}_{1n_1})$ and $(\tilde{Z}_{21}, \tilde{Z}_{22}, \cdots, \tilde{Z}_{2n_2})$. Now the fuzzy discriminant axes for two groups are calculated as

$$\bar{\tilde{Z}}_1 = \tilde{b}' \otimes \bar{\tilde{X}}^{(1)};\ \bar{\tilde{Z}}_2 = \tilde{b}' \otimes \bar{\tilde{X}}^{(2)} \text{ and } \mathrm{Var}\,(\tilde{Z}) = \frac{\sum_{j=1}^{n_1} (\tilde{Z}_{1j} - \bar{\tilde{Z}}_1)^2 + \sum_{j=1}^{n_2} (\tilde{Z}_{2j} - \bar{\tilde{Z}}_2)^2}{n_1 + n_2 - 2}\ .$$

The pooled covariance matrix for all the groups is computed by

$$S_p = \frac{\sum_k (n_k - 1) S^{(k)}}{\sum (n_k - 1)}\ .$$

The separation between two groups is defined by

$$\left[\frac{\text{squared distance between sample means of } \tilde{Z}}{\text{sample variance of } \tilde{Z}}\right]$$

$$= \left[\frac{(\tilde{Z}_1 - \tilde{Z}_2)^2}{\text{Var}\,(\tilde{Z})} \right] = \left[\frac{(\tilde{b}' \otimes \tilde{X}^{(1)} - \tilde{b}' \otimes \tilde{X}^{(2)})^2}{\tilde{b}' S_p \tilde{b}} \right].$$

The maximum is achieved for the choice of $\tilde{b}$ where $\hat{\tilde{b}} = \left(\tilde{X}^{(1)} - \tilde{X}^{(2)} \right)' S_p^{-1}$ from the linear programming:

$$\underset{\tilde{b}}{\text{Max}} \left[\frac{(\tilde{b}' \otimes \tilde{X}^{(1)} - \tilde{b}' \otimes \tilde{X}^{(2)})^2}{\tilde{b}' S_p \tilde{b}} \right].$$

The deviations between two fuzzy vectors are calculated as follows:

$$\tilde{X}^{(1)} - \tilde{X}^{(2)} = \begin{bmatrix} (m_{\bar{X}_1}^{(1)} - m_{\bar{X}_1}^{(2)},\ a_{\bar{X}_1}^{(1)} + a_{\bar{X}_1}^{(2)},\ b_{\bar{X}_1}^{(1)} + b_{\bar{X}_1}^{(2)})_{LR} \\ (m_{\bar{X}_2}^{(1)} - m_{\bar{X}_2}^{(2)},\ a_{\bar{X}_2}^{(1)} + a_{\bar{X}_2}^{(2)},\ b_{\bar{X}_2}^{(1)} + b_{\bar{X}_2}^{(2)})_{LR} \\ \vdots \\ (m_{\bar{X}_m}^{(1)} - m_{\bar{X}_m}^{(2)},\ a_{\bar{X}_m}^{(1)} + a_{\bar{X}_m}^{(2)},\ b_{\bar{X}_m}^{(1)} + b_{\bar{X}_m}^{(2)})_{LR} \end{bmatrix}.$$

Therefore, the estimated fuzzy discriminant function from eqn (8) is formulated as follows:

$$\tilde{Z} = (\tilde{X}^{(1)} - \tilde{X}^{(2)})^T S_{\text{p}}^{-1}\ \underline{\tilde{X}}. \tag{9}$$

And the fuzzy threshold on the basis of which a fuzzy classification rule is set up is computed

$$\tilde{m} = \frac{1}{2}(\tilde{Z}_1 + \tilde{Z}_2),\ \text{where}\ \tilde{Z}_1 = (\tilde{X}^{(1)} - \tilde{X}^{(2)})^T S_{\text{pooled}}^{-1}\ \tilde{X}^{(1)}\ \text{and}\ \tilde{Z}_2 = (\tilde{X}^{(1)} - \tilde{X}^{(2)})^T S_{\text{pooled}}^{-1}\ \tilde{X}^{(2)}.$$

Therefore, the fuzzy discriminant scores for two groups are obtained based on the fuzzy discriminant axes and the entities are categorized based on the defuzzified value of $\tilde{m}$, denoted by d($\tilde{m}$). Here, we actually have applied the centroid method for defuzzification. The obtained fuzzy scores are also defuzzifed to make a better comparison with d($\tilde{m}$). Hence, a fuzzy classification rule can be formulated as follows:

IF ($\tilde{X}_1$ *is* $\tilde{x}_1$) & ($\tilde{X}_2$ *is* $\tilde{x}_2$) & … & ($\tilde{X}_m$ *is* $\tilde{x}_m$) THEN *decision class is*

$$\begin{cases} G_1 & \text{if } \mathrm{d}(\tilde{Z}) \geq \mathrm{d}(\tilde{m}) \\ G_2 & \text{if } \mathrm{d}(\tilde{Z}) < \mathrm{d}(\tilde{m}) \end{cases}.$$

3.2. *Backpropagation neural network framework*

The most commonly used ANN is the feedforward network trained using the backpropagation algorithm [12], which is adopted in the present study. The neural network model is designed with the linguistic variables (first layer) where the linguistic values are defined in the second layer, called fuzzification layer. Then the third layer inputs defuzzified values that constitute a representative of fuzzy rule.

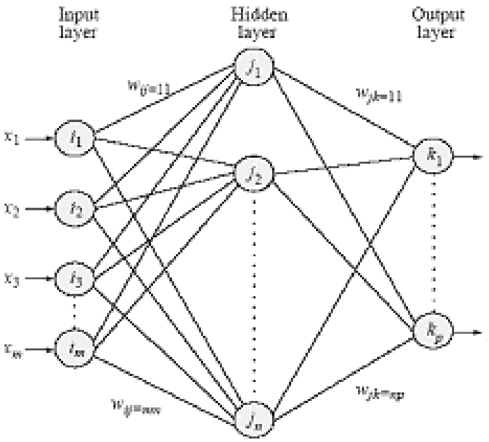

FIG. 1. Backpropagation neural network topology.

Now the backpropagation algorithm can be described in three equations for classification. First, weight connections are changed in each learning step (k) with

$$\Delta w_{ij(k)}^{[s]} = h d_{pj}^{[s]} x_j^{[s-1]} + m \Delta w_{ij(k-1)}^{[s]}. \tag{10}$$

Second, for output nodes it holds that

$$d_{pj}^{[o]} = (d_j - o_j) f_j^{/}(I_j^{[s]}) \tag{11}$$

and third, for the remaining nodes it hold that

$$d_{pj}^{[s]} = f_j^{/}(I_j^{[s]}) \sum_k d_{pj}^{[s+1]} w_{jk}^{[s+1]} \tag{12}$$

where $x_j^{[s]}$ is the actual output of node j in layer s; $w_{ij}^{[s]}$, the weight of the connection between node i at layer (s – 1) and node j at layer (s); $d_{pj}^{[s]}$, the measure for the actual error of node j; $I_j^{[s]}$, the weighted sum of the inputs of node j in layer s; h, the time-dependent learning rate; $f(\ )$, the transfer function; m, the momentum factor (between 0 and 1); and d_j, o_j are the desired and actual activity of node j (for output nodes only). Parameter values (i.e. the learning rate h, momentum factor m, and the number of hidden nodes h_j) are selected experimentally. The input and output nodes are selected according to the linguistic variables and class of the objects to be classified.

4. Example: Admission to management programme

A practical example on admission to management programme in a business school has been considered here to employ the proposed methodology. There are several factors on the basis of which students are to be evaluated by the evaluators. Especially in management programme, it becomes important to focus on some cognitive factors [13]. Perception, attitude reasoning and thinking, etc. are the cognitive factors. Here we are considering only three major factors: JMAT score, Work experience and Overall performance (communication

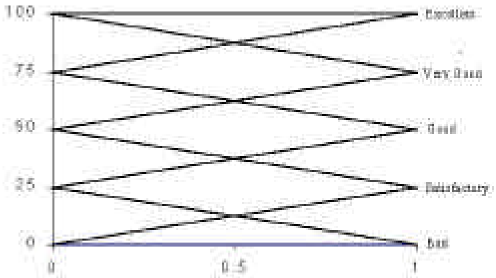

FIG. 2. Linguistic scale for $\tilde{X}_1$.

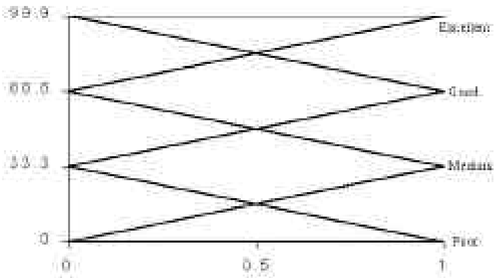

FIG. 3. Linguistic scale for $\tilde{X}_2$ and $\tilde{X}_3$.

skill and other abilities) of a candidate. In connection with experience, knowledge, perception, attitude, reasoning and thinking, the factors in fact lead to three fuzzy variables. The variables consider the experts' fuzzy responses for evaluation of a candidate whether s/he will be either admitted or not admitted into the programme. Therefore, we have two groups: 'Admitted' (G_1) and 'Not-admitted' (G_2) and three fuzzy variables: $\tilde{X}_1$: JMAT score; $\tilde{X}_2$: Work experience and $\tilde{X}_3$: Overall performance. Also the fuzzy scales for the variables are defined on the domain [0, 100] according to expert.

Here, a data set of size 60 (see Appendix A1) has been surveyed during the admission of the management programme where the first 45 data are set as training data and the last 15 are to be tested using fuzzy linear discriminant analysis and neural network.

4.1. *Results*: *fuzzy linear discriminant analysis (FLDA)*

Here the fuzzy means and variance–covariance are computed using eqns (1)–(5) for two classes.

(a) Class (C_1): Admitted (A)

$$\begin{aligned}\tilde{X}_1 &= (79.2,\ 25.1,\ 12.9)_{LR}\\ \tilde{X}_2 &= (73.6,\ 33.3,\ 22.2)_{LR}\\ \tilde{X}_3 &= (61.1,\ 33.3,\ 27.8)_{LR}\end{aligned} \quad \text{and} \quad S_A = \begin{pmatrix} 244.42 & -81.01 & 40.53 \\ -81.01 & 305.88 & -28.61 \\ 40.53 & -28.61 & 415.21 \end{pmatrix}$$

(b) Class (C_2): Not-Admitted (NA)

$$\begin{aligned}\tilde{X}_1 &= (41.7,\ 23.9,\ 25)_{LR}\\ \tilde{X}_2 &= (57.1,\ 30.2,\ 25.4)_{LR}\\ \tilde{X}_3 &= (41.3,\ 32.2,\ 33.3)_{LR}\end{aligned} \quad \text{and} \quad S_{NA} = \begin{pmatrix} 364.42 & -245.41 & 19.4 \\ -245.41 & 691.64 & -43.2 \\ 19.4 & -43.2 & 353.39 \end{pmatrix}.$$

Therefore, the pooled estimated covariance matrix is computed by $S_p = \frac{1}{2}\left(S_A + S_{NA}\right)$. Hence, the linear fuzzy discriminant function using eqn (9) is obtained as

$$\tilde{Z} = (0.17,\ 0.26,\ 0.20)_{LR} \otimes \tilde{X}_1 \oplus (0.53,\ 0.8,\ 0.61)_{LR} \otimes \tilde{X}_2 \oplus (0.04,\ 0.17,\ 0.15)_{LR} \otimes \tilde{X}_3 \quad (10)$$

Table II
Classification using FLDA for 15 testing samples (Appendix)

Sample		Class
46	$(28.69, 107.48, 63.49)_{LR}$	Admitted
	(23.08, 121.89, 35.827)	Admitted
48	$_{LR}$	
49	(20.17, 23.42, 53.49)	Not Admitted
	(18.83, 17.74, 48.476)	Not Admitted
	(24.42, 100.98, 58.49)	Admitted
52	$_{LR}$	
53	$(21.75, 61.54, 47.20)_{LR}$	Admitted
	(23.08, 68.53, 53.53)	Admitted
**55	$_{LR}$	-Admitted
**56	$4.42, 127.57, 40.84)_{LR}$	Admitted
	(28.67, 134.07, 45.84)	Admitted
58	$_{LR}$	-Admitted
59	$_{LR}$	-Admitted
60	$(21.75, 17.25, 47.14)_{LR}$	Not-Admitted

Therefore, the fuzzy threshold is automatically computed by the methodology as $\tilde{m} = (47, 39.6, 25.7)_{LR}$. After building the fuzzy discriminant function based on the training fuzzy data set, the training samples are tested in the following. In doing so, the fuzzy discriminant scores and threshold value are defuzzified using MATLAB 7.0 to classify the objects either in 'Admitted' class or in 'Not-Admitted' class. Now the tested results are given in Table II:

It can be observed from the above table that only three samples (marked by '*') have been misclassified.

4.2. *Results: neural network approach*

The network is initially trained with 45 samples where the error history is depicted in Fig. 4. Here, we have considered four nodes in the hidden layer with h = 0.90 and m = 0.75.

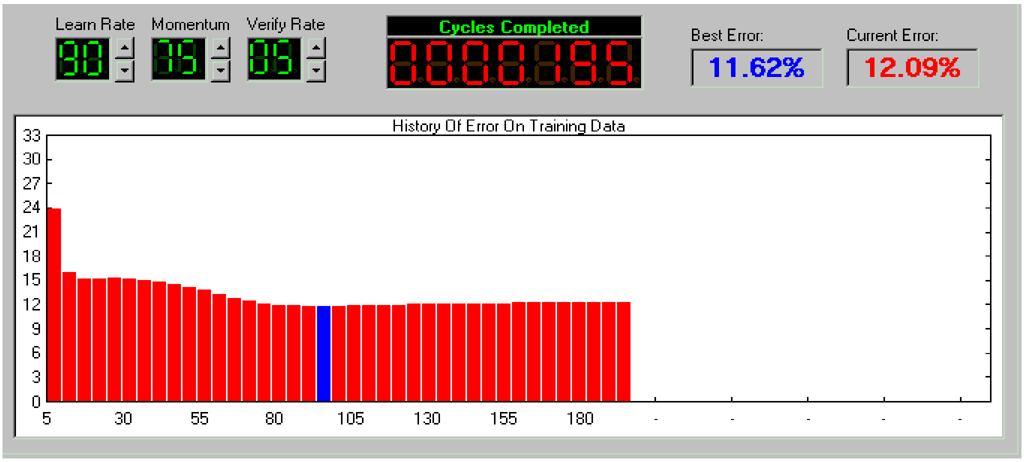

FIG. 4. History of error on training data.

Table III
Classification by backpropagation neural network for the testing data

*	ncoded_Clas	Predicted	Difference	JMAT score	Work-Exp	Overall Performance	ID	class
	1	1.02	0.02	75	64.97	94.97	46	Admitted
	1	0.63	0.37	50	94.97	34.97	47	Admitted
	1	0.96	0.04	75	64.97	64.97	48	Admitted
	0	-0.22	0.22	25	11.03	64.97	49	Not-Admitted
	0	-0.15	0.15	50	11.03	34.97	50	Not-Admitted
	1	0.96	0.04	75	64.97	64.97	51	Admitted
	1	0.63	0.37	50	94.97	34.97	52	Admitted
√	1	-0.08	1.08	50	34.97	11.03	53	Admitted
√	1	-0.04	1.04	50	34.97	34.97	54	Admitted
√	1	-0.28	1.28	25	11.03	11.03	55	Admitted
√	0	0.99	0.99	50	94.97	94.97	56	Not-Admitted
√	0	1.03	1.03	75	94.97	64.97	57	Not-Admitted
	0	-0.25	0.25	25	11.03	34.97	58	Not-Admitted
	0	-0.29	0.29	8	34.97	11.03	59	Not-Admitted
	0	-0.19	0.19	50	11.03	11.03	60	Not-Admitted

The neural network approach to fuzzy variables (Table III) leads to five samples (from 53 to 57) to be misclassified (marked by √) by means of comparing the predicted values with the actual values.

5. Conclusions

The results in Tables II and III obtained by the proposed FLDA and neural network depict almost the same classification except sample nos 53 and 54. Though sample 53 is correctly classified by neural network, sample no. 54 is not. The reverse case can be investigated in the case of FLDA. But it is recommendable here that both the methods are able to detect the misclassified entities. In FLDA, the fuzzy threshold is automatically computed on the basis of training linguistic data. The computational complexities of neural network are relatively higher than FLDA while using the continuous fuzzy numbers. Rather the proposed tool is a simplistic approach to obtain similar classification results in comparison with backpropagation neural network. This method directly deals with the fuzzy numbers and enhances the fuzzy scores, those of which can also be compared with fuzzy threshold by approximate reasoning [13]. This method can also be applicable in designing a fuzzy decision-making interface for understanding group membership based on fuzzy perception.

Acknowledgement

The authors are grateful to the anonymous referees for their valuable comments and suggestions. The authors also acknowledge the support provided by the Council of Scientific and Industrial Research, Govt of India (scheme no. 9/81(522)/05-EMR-II), and the Department of Science and Technology, Government of India (DST/MS/157/01).

References

1. J. F. Hair, R. E. Anderson, R. Tatham, and W. C. Black, *Multivariate data analysis*, fifth edition, Anand Sons (2003).
2. L. A. Zadeh, Fuzzy Sets, *Info. Control*, **8**, 338–353 (1965).
3. D. Dubois, and H. Prade. *Fuzzy sets and systems* (*Theory and Applications*), Academic Press (1980).

4. K. Y. Tam and N. Y. Kiang, Managerial applications of neural networks: the case of bank failure predictions, *Mgmt Sci.*, **38**, 926–947 (1992).
5. B. Baesens, R. Setiono, C. Mues, and J. Vanthienen, Using neural network rule extraction and decision tables for credit-risk evaluation, *Mgmt Sci.*, **49**, 312–329 (2003).
6. Shouhong Wang, Classification with incomplete survey data: A Hopfield neural network approach, *Computers Oper., Res.*, **32**, 2583–2594 (2005).
7. C. C. Lin, and A. P. Chen, A method for two group fuzzy discriminant analysis, *Int. J. Fuzzy Systems*, **3**, 341–345 (2001).
8. C. C. Lin, and A. P. Chen, Fuzzy discriminant analysis with outlier detection by genetic algorithm, *Computers Oper. Res.*, **31**, 877–888 (2004).
9. C. Chiang, Q. Liu, and Y. Liu, Fuzzy linear programming and discriminant analysis, *Int. J. Fuzzy Systems*, **3**, 462–465 (2001).
10. Z. P. Chen, J. H. Jiang, Y. Li, Y. Z. Liang, and R. Q. Yu, Fuzzy linear discriminant analysis for chemical data sets, *Chemometrics Intelligent Laboratory Systems*, **45**, 295–302 (1999).
11. J. Watada, H. Tanaka, and K. Asai, Fuzzy discriminant analysis in fuzzy groups, *Fuzzy Sets Systems*, **19**, 261–271 (1986).
12. D. E. Rumelhart, and J. McClelland, *Parallel distributed processing*, Vol. 1, MIT Press (1986).
13. Chandan Chakraborty, and Debjani Chakraborty, A decision scheme based on OWA operator for an evaluation programme: An approximate reasoning approach, *Int. J. Appl. Soft Computing*, **5**, 45–53 (2004).

Appendix A1

Fuzzy data set of size 60 (Training data: 1–15 and Test data: 46–60)

ID	Class	JMAT score	Work experience	Overall performance	ID	Class	JMAT score	Work experience	Overall performance
1	Admitted	G	G	E	31	Admitted	E	E	E
2	Admitted	VG	E	M	32	Admitted	VG	G	M
3	Admitted	G	E	G	33	Admitted	E	G	E
4	Admitted	E	G	M	34	Admitted	VG	M	G
5	Admitted	V	G	G	35	Admitted	VG	E	E
6	Admitted	E	G	G	36	Not-Admitted	S	G	M
7	Admitted	VG	G	M	37	Not-Admitted	G	VG	M
8	Not-Admitted	G	M	G	38	Not-Admitted	S	M	M
9	Not-Admitted	S	E	G	39	Not-Admitted	B	G	M
10	Not-Admitted	VG	G	M	40	Not-Admitted	VG	M	M
11	Not-Admitted	VG	M	G	41	Not-Admitted	G	P	G
12	Not-Admitted	S	E	G	42	Not-Admitted	S	M	M
13	Not-Admitted	G	E	M	43	Admitted	E	G	G
14	Not-Admitted	S	E	M	44	Admitted	E	G	E
15	Not-Admitted	G	G	M	45	Admitted	VG	E	G
16	Admitted	G	E	M	46	Admitted	VG	G	E
17	Not-Admitted	G	M	M	47	Admitted	G	E	M
18	Admitted	VG	G	G	48	Admitted	VG	G	G
19	Admitted	E	M	G	49	Not-Admitted	S	P	G
20	Admitted	E	G	M	50	Not-Admitted	G	P	M
21	Not-Admitted	S	E	G	51	Admitted	VG	G	G
22	Not-Admitted	S	G	M	52	Admitted	G	E	M
23	Admitted	VG	G	M	53	Admitted	G	M	P
24	Admitted	VG	G	G	54	Admitted	G	M	M
25	Admitted	VG	E	M	55	Admitted	S	P	P
26	Not-Admitted	G	M	P	56	Not-Admitted	G	E	E
27	Not-Admitted	G	P	G	57	Not-Admitted	VG	E	G
28	Admitted	G	E	G	58	Not-Admitted	S	P	M
29	Not-Admitted	S	G	P	59	Not-Admitted	B	M	P
30	Admitted	VG	M	G	60	Not-Admitted	G	P	P

‘E’ = Excellent; ‘VG’ = Very Good; ‘G’ = Good; ‘S’ = Satisfactory; ‘M’ = Medium; ‘P’ = Poor; ‘B’ = Bad.

Appendix A2
Defuzzification using Matlab 7.0

ID	Class	JMAT score	Work experience	Overall performance	ID	Class	JMAT score	Work experience	Overall performance
1	Admitted	50	64.97	94.97	31	Admitted	100	94.97	94.97
2	Admitted	75	94.97	34.97	32	Admitted	75	64.97	34.97
3	Admitted	50	94.97	64.97	33	Admitted	100	64.97	94.97
4	Admitted	100	64.97	34.97	34	Admitted	75	34.97	64.97
5	Admitted	75	64.97	64.97	35	Admitted	75	94.97	94.97
6	Admitted	100	64.97	64.97	36	Not-Admitted	25	64.97	34.97
7	Admitted	75	64.97	34.97	37	Not-Admitted	50	64.97	34.97
8	Not-Admitted	50	34.97	64.97	38	Not-Admitted	25	34.97	34.97
9	Not-Admitted	25	94.97	64.97	39	Not-Admitted	8	64.97	34.97
10	Not-Admitted	75	64.97	34.97	40	Not-Admitted	75	34.97	34.97
11	Not-Admitted	75	34.97	64.97	41	Not-Admitted	50	11.03	64.97
12	Not-Admitted	25	94.97	64.97	42	Not-Admitted	25	34.97	34.97
13	Not-Admitted	50	94.97	34.97	43	Admitted	100	64.97	64.97
14	Not-Admitted	25	94.97	34.97	44	Admitted	100	64.97	94.97
15	Not-Admitted	50	64.97	34.97	45	Admitted	75	94.97	64.97
16	Admitted	50	94.97	34.97	46	Admitted	75	64.97	94.97
17	Not-Admitted	50	34.97	34.97	47	Admitted	50	94.97	34.97
18	Admitted	75	64.97	64.97	48	Admitted	75	64.97	64.97
19	Admitted	100	34.97	64.97	49	Not-Admitted	25	11.03	64.97
20	Admitted	100	64.97	34.97	50	Not-Admitted	50	11.03	34.97
21	Not-Admitted	25	94.97	64.97	51	Admitted	75	64.97	64.97
22	Not-Admitted	25	64.97	34.97	52	Admitted	50	94.97	34.97
23	Admitted	75	64.97	34.97	53	Admitted	50	34.97	11.03
24	Admitted	75	64.97	64.97	54	Admitted	50	34.97	34.97
25	Admitted	75	94.97	34.97	55	Admitted	25	11.03	11.03
26	Not-Admitted	50	34.97	11.03	56	Not-Admitted	50	94.97	94.97
27	Not-Admitted	50	11.03	64.97	57	Not-Admitted	75	94.97	64.97
28	Admitted	50	94.97	64.97	58	Not-Admitted	25	11.03	34.97
29	Not-Admitted	25	64.97	11.03	59	Not-Admitted	8	34.97	11.03
30	Admitted	75	34.97	64.97	60	Not-Admitted	50	11.03	11.03