Health Care Management Science
https://doi.org/10.1007/s10729-019-09477-1

Fuzzy DEA-based classifier and its applications in healthcare management

Ai-bing Ji[1] · Yanhua Qiao[1] · Chang Liu[1]

Received: 28 June 2018 / Accepted: 11 February 2019

Abstract

Nonlinear fuzzy classification models have better classification performance than linear fuzzy classifiers. In many nonlinear fuzzy classification problems, piecewise-linear fuzzy discriminant functions can approximate nonlinear fuzzy discriminant functions. In this paper, we first build fuzzy classifier based on data envelopment analysis (DEA) for incremental separable fuzzy training data, which can be widely applied in the healthcare management with fuzzy attributes, then we apply the proposed fuzzy DEA-based classifier in the diagnosis of Coronary with fuzzy symptoms and the classification of breast cancer dataset with fuzzy disturbance. Numerical experiments show the proposed fuzzy DEA-based classifier is accurate and robust.

Keywords Fuzzy data envelopment analysis · DEA classification machine · Piecewise-linear discriminant analysis · Fuzzy training set · Healthcare management

1 Introduction

Data mining is an important method of knowledge management and health care management [1], the classical data mining algorithm assumes the data to be exact, but in some applications, the data which are to be processed are rarely exactly determined. The indeterminism is more or less hidden in their structure with different types: imprecision or approximation, randomness, and also vagueness. The last type of indeterminism is very frequently connected with the properties of natural language used during the process of data acquisition. In the communication among people, certain degree of vagueness is not only acceptable but even useful. In disease diagnosis, some symptom of an illness are also vague. The world appears rather more uncertain than the results aiming to state. Consequently, the development of adequate mathematical tools for processing vagueness became quite urgent.

Since its introduction by Lotfi Zadeh, fuzzy set theory has been fruitfully used in many domains. Data mining is one of the important fields of its application. Early after the appearance of data mining, several works have proposed the use of fuzzy set theory to this domain.

The contributions of fuzzy sets in data mining are various: increasing the interpretability, enhancing the robustness of the process and managing unclear information. Both are provided by the introduction of fuzzy set theory to build up model of fuzzy data mining, offering to that process the capacity to mine complex information difficult to treat in a classical environment. Robustness of the process enables it to produce similar results when facing data with small changes (for instance, in the presence of noise).

Data mining aims to set up a model from a set of data providing some background knowledge. The model can be viewed as a new knowledge that is produced from the learning, it can be of various forms: e.g. mathematical function, neural network, rule base, patterns, association rules.

The data mining techniques provide several benefits in healthcare management [2, 3] like detection diseases cause, medical treatment methods identification, constructing drug recommendation systems, individuals health profile development, healthcare research policies, even detection of fraud in health insurance. The data mining technologies can provide benefits to healthcare by grouping the patients having similar types of diseases so that the healthcare organizations provide them better treatments.

Fuzzy data mining is an extension of data mining in which fuzzy set modeling is introduced. Many classical data mining algorithms have been extended to fuzzy cases [4–7].

✉ Yanhua Qiao
jab@hbu.edu.cn

[1] College of Public Health, Affiliated Hospital of Hebei University, Hebei University, Baoding 071000, Hebei, People's Republic of China

Published online: 08 March 2019

Extending a classical algorithm to build up a fuzzy learning algorithm is an interesting task. Many papers have been published in the case of fuzzy decision trees [8–10], in fuzzy rule base construction [11–13] and fuzzy kernel method [14–18]. The challenge in such a case is to propose an algorithm that can both handle fuzzy input and still satisfy the main properties of the classical algorithm.

Classification, as an important data mining method, is an organizing system based on similarities and differences, it is an important forecasting (decision making) method and is widely used in healthcare management [2, 19] and disease surveillance and early warning [20]. A classification aims to judge whether a data belongs to a particular group by evaluating a set of attribute values, the techniques and methods for the conventional classification machine can be found in the works of Han and Kamber [21]. In practice, a health care decision making system determines if a patient has an illness according to some given symptom indexes and physical conditions, in essence, diseases diagnosis is a classification problem.

Data envelopment analysis (DEA) is used to evaluate the relative efficiency among a given number of decision making units (DMUs) with multiple inputs and multiple outputs, by solving linear programming problems for each DMU; The well-known DEA models include the CCR model by Charnes et al. [22] and BCC model by Banker et al. [23].

DEA are mainly used in the performance evaluation [24–26]. Besides performance evaluation, DEA can also be applied in classification, The pioneering work in the application of DEA to classification problems was performed by Troutt et al. [27]. In their works, acceptance domains were constructed, and a sample-based decision system was presented, which makes decisions on whether to accept or reject a credit risk, based on samples predetermined by experts. Based on DEA model, two new mathematical programming approaches are developed for the minimization of the sum of the deviations and for the relative efficiency concept of DEA in solving two group classification problems [28]. Yan and Wei [29] established an equivalence relation between the DEA classification machine and DEA model, and then created a DEA-based classification machine, in which the classification of data is equivalent to testing whether a particular DMU is in the production possibility set. In all above mentioned DEA-based classifiers, they are all based on exact data, Wei [30] developed the classifiers for dealing with binary classification problems with interval data, although it has some uncertainties, also unambiguous, so it is very necessary to build the classifier based on the fuzzy DEA for the fuzzy training data.

In this paper, a novel DEA-based fuzzy classifier is constructed. we first introduce the fuzzy data envelopment analysis classification machine for conditional monotonicity fuzzy training data. Then experiments are carried out on the actual fuzzy data set and disturbed UCI data set to prove the accuracy and efficiency of the proposed model.

2 Background

In this section, we first review the related concepts and models of DEA, fuzzy DEA and DEA classification machine, then we give some basic definitions on possibility measure and fuzzy chance constrained programming, some important theorems are proved.

2.1 Data envelopment analysis

Suppose that there are n DMUs, each producing m outputs from s inputs. DMU_0 is the DMU to be evaluated.DMU_k uses the input bundle $x_k=(x_{1k},x_{2k},\dots,x_{sk})$ to produce the output bundle $y_k=(y_{1k},y_{2k},\dots,y_{mk})$. A standard DEA model for assessing DMU_0, known as the CCR model [22], is formulated in model:

$$\begin{cases} \max \sum\limits_{i=1}^{m} \mu_i y_{i0} \\ \sum\limits_{i=1}^{m} \mu_i y_{ik} - \sum\limits_{j=1}^{s} \omega_j x_{jk} \le 0, (k=1,2,\dots,n); \\ \sum\limits_{j=1}^{s} \omega_j x_{j0} = 1; \\ \mu_i \ge 0, \omega_j \ge 0, (i=1,2,\dots m; j=1,2,\dots,s); \end{cases} \tag{1}$$

The efficiency evaluation models above are called CCR models which are treated output oriented forms.

The dual form of this model (1) is

$$\begin{cases} \min\theta \\ \sum\limits_{k=1}^{n} \lambda_k x_{jk} \le \theta x_{j0}, (j=1,2,\dots,s); \\ \sum\limits_{k=1}^{n} \lambda_k y_{ik} \ge y_{i0}, (i=1,2,\dots,m); \\ \lambda_k \ge 0, (k=1,2,\dots n). \end{cases} \tag{2}$$

The essence of CCR models is to find the optimal weight vector to maximize its weights for inputs and outputs which maximize the efficiency score of the DMU under evaluation.

2.2 Fuzzy data envelopment analysis

In order to evaluate the efficiency with fuzzy data, fuzzy set theory has been proposed as a way to quantify imprecise and vague data in DEA models. Fuzzy DEA models adopt the form of fuzzy linear programming models. CCR model with fuzzy data is given as following:

$$\begin{cases} \max \sum_{i=1}^{m} u_i \tilde{y}_{i0} \\ \sum_{i=1}^{m} u_i \tilde{y}_{ik} - \sum_{i=1}^{s} v_i \tilde{x}_{ik} \tilde{<} \tilde{0} \\ \sum_{i=1}^{s} v_i \tilde{x}_{i0} \tilde{=} 1 \\ u_i \geq 0, v_j \geq 0 (i = 1, 2, ..., m; j = 1, 2, ..., s); \end{cases} \quad (3)$$

Where the input $\tilde{x}_i (i = 1, 2, ..., m)$ and output $\tilde{y}_i$ $(i = 1, 2, ..., s)$ are all fuzzy numbers.

Several fuzzy methods have been proposed to solve the fuzzy DEA model [31], the fuzzy DEA methods can be classified into the following five groups: (1) The tolerance approach [32]; (2) The $\alpha -$level based approach [33–35]; (3) The fuzzy ranking approach [36, 37]; (4) The possibility approach [38, 39]; (5) Other developments [40–42]. In order to improve the discriminative power of the fuzzy DEA model [43], cross-efficiency fuzzy DEA [44] was introduced.

2.3 DEA classification machine (DCM)

Hong Yan, Quanling Wei [29] extended DEA method to the classification of large data. They treat each datum as an evaluated DMU, with the attribute values as input and single output of value 1. Toloo [45, 46] extended the DEA approach to find the most efficient unit of a data set without explicit inputs (outputs).

Consider a sample training data set $\hat{T}$ (with single class)

$$\hat{T} = \{x_k | k = 1, 2, ..., n\}$$

where $x_k \in E^m$, $x_k > 0$, $k = 1, 2, ..., n$. Using the terms in the DEA research, x_k is DMUs with certain characteristics described by $x_k = (x_{1k}, x_{2k}, ..., x_{sk})^T$, where x_{ik} is the ith characteristic value of DMU_k.

This problem can be described by CCR model [22] with the DMU_k of the input given by x_k and output $y_k = 1 (k = 1, 2, ..., n)$. That is, the sample training data set in the DEA model is given by $\{(x_k, 1) | k = 1, 2, ..., n\}$.

Then CCR model with input–output value $(x_k, 1)$ for DMU_k (k = 1,2,…,n) is given by

$$(P) \begin{cases} \max \mu_0 \\ \omega x_k - \mu_0 \geq 0, k = 1, 2, ..., n; \\ \omega x_0 = 1; \\ \omega \geq 0, \mu_0 \geq 0. \end{cases} \quad (4)$$

Where DMU_{j_0} is the evaluated unit and $x_0 = x_{j_0}, 1 \leq j_0 \leq n$.

Definition 2.1 [28] If the optimal values of (P) are 1, then DMU_0 is called weakly DEA efficient.

For weakly DEA efficient DMU_0, x_0 satisfies (note that ω^0 $x_0 = 1, \mu_0^0 = 1$)

$$L : \omega^0 x - 1 = 0$$

L is a supporting plane of the acceptance domain and is called the classification hyperplane, x_0 is on this plane.

The acceptance domain T can be given in its intersection form

$$T = \left\{ x | \omega^k x - 1 \geq 0, k = 1, 2, .., n \right\}$$

Where ω^k, $\mu_0^k = 1$ (k = 1,2,…,n) are the optimal solutions of (4). Then, for $x \in \hat{T}$, the classification function is given by

$$d(x) = sign\left(\min_{1 \leq k \leq n} \left(\omega^k x - \mu_0^k \right) \right)$$

If $d(x) = 1$, then $x \in T$, $d(x) = -1$, then $x \notin T$.

2.4 Possibility measure and fuzzy chance constrained programming

Definition 2.2 Let X be a nonempty set, P(X) be the class of all subsets of X, a mapping.

Pos: P(X)→[0, 1] is called a possibility measure if it satisfies:

(1) $Pos(\varphi) = 0$; (2) $Pos(X) = 1$; (3) $Pos(\underset{t \in T}{\cup} A_t$ $) = Sup_{t \in T} \{Pos(A_t)\}$.

Definition 2.3 Let $\tilde{a}$ be a fuzzy number, and its membership function is

$$\mu_{\tilde{a}}(x) = \begin{cases} \frac{x - r_1}{r_2 - r_1}, & r_1 \leq x < r_2 \\ 1 & x = r_2 \\ \frac{x - r_3}{r_2 - r_3}, & r_2 < x \leq r_3 \end{cases}$$

where

$r_1 \leq r_2 \leq r_3$, *and* r_1, r_2, r_3	are real numbers
$\tilde{a}$	is called a triangular fuzzy number, denoted by (r_1, r_2, r_3).

The classes of all triangular fuzzy number are denoted by $T(R)$, If $\tilde{x}_i (i = 1, 2, \cdots, n)$ are all fuzzy numbers, then $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ is called a fuzzy number vector, the classes of all fuzzy number vectors are denoted by $F^n(R)$, especially when $\tilde{x}_i (i = 1, 2, \cdots, n)$ are all triangular fuzzy numbers, then $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \cdots \tilde{x}_n)$ is called a triangular fuzzy number vector, the class of all triangular fuzzy number vectors is denoted by $T^n(R)$.

Definition 2.4 Let $\tilde{a}, \tilde{b}$ be two fuzzy numbers, then $\tilde{a} \vee \tilde{b}$ represents the fuzzy number having the following membership function $\mu_{\tilde{a} \vee \tilde{b}}(x) = \sup_{s \vee t = x} \left\{ \mu_{\tilde{a}}(s) \wedge \mu_{\tilde{b}}(t) \right\}$.

Definition 2.5 Let $\tilde{a}, \tilde{b}$ be two fuzzy number, then $\tilde{a} \tilde{>} \tilde{b} \Leftrightarrow \tilde{a} \vee \tilde{b} = \tilde{a}$..

Lemma 1 [43] Let $\tilde{a}, \tilde{b}$ be two fuzzy number, then $\tilde{a} \tilde{>} \tilde{b}$ if and only if, for $\forall h \in [0, 1]$, the two statement below hold:

$$\inf\left\{s : \mu_{\tilde{a}}(s) \geq h\right\} \geq \inf\left\{t : \mu_{\tilde{b}}(t) \geq h\right\},$$

$$\sup\left\{s : \mu_{\tilde{a}}(s) \geq h\right\} \geq \sup\left\{t : \mu_{\tilde{b}}(t) \geq h\right\}.$$

By lemma 1, we can easily prove the following lemma.

Lemma 2 Let $\tilde{a}, \tilde{b}$ be two fuzzy number, (1) For $\lambda > 0, \lambda\tilde{a} \tilde{>} \lambda\tilde{b}$; (2) For every fuzzy number $\tilde{c}, \tilde{a} + \tilde{c} > \tilde{b} + \tilde{c}$.

Definition 2.6 Let $\tilde{a}$ be a fuzzy number, b is a real number, then the possibility measure of fuzzy event $\tilde{a} < b$ is defined by $Pos(\tilde{a} \leq b) = \sup\{\mu_{\tilde{a}}(x) | x \in R, x \leq b\}$.

Similarly, $Pos(\tilde{a} < b) = \sup\{\mu_{\tilde{a}}(x) | x \in R, x < b\}$, $Pos(\tilde{a} \geq b) = \sup\{\mu_{\tilde{a}}(x) | x \in R, x \geq b\}$ $Pos(\tilde{a} = b) = \mu_{\tilde{a}}(b)$.

Theorem1 Let $\tilde{a}, \tilde{b}$ be two fuzzy numbers, c be a real number, for a possibility level $\alpha \in [0, 1]$ if $\tilde{a} \tilde{>} \tilde{b}$ and $Pos(\tilde{b} \geq c) \geq \alpha$, then $Pos(\tilde{a} \geq c) \geq \alpha$.

Proof Because $\tilde{a} \tilde{>} \tilde{b}$, for a given x, t = x, by Lemma 1, there are s: $s \leq t$ such that $\mu_{\tilde{a}}(s) \leq \mu_{\tilde{b}}(x)$, thus $\mu_{\tilde{a} \vee \tilde{b}}(x) = \sup_{s \vee t = x} \{\mu_{\tilde{a}}(s) \wedge \mu_{\tilde{b}}(t)\} \geq \mu_{\tilde{a}}(s) \wedge \mu_{\tilde{b}}(x) \geq \mu_{\tilde{b}}(x)$, therefore $Pos(\tilde{a} \geq c) = \sup\{\mu_{\tilde{a} \vee \tilde{b}}(x) | x \geq c\} \geq \sup\{\mu_{\tilde{b}}(x) | x \geq c\} \geq \alpha$.

If $\tilde{x}_i (i = 1, 2, \cdots, n)$ are all fuzzy numbers, then $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ is called a fuzzy number vector, the class of all fuzzy number vectors is denoted by $F^n(R)$, especially when $\tilde{x}_i (i = 1, 2, \cdots, n)$ are all triangular fuzzy numbers, then $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \cdots \tilde{x}_n)$ is called a triangular fuzzy number vector, all the triangular fuzzy number vectors is denoted by $T^n(R)$.

Following from the Zadeh extension principle, then for function $f: R^n \to R$, $\tilde{y} = f(\tilde{x}_1, \tilde{x}_2, \cdots \tilde{x}_n)$ is a fuzzy number, and its membership function is:

$$\mu_{\tilde{y}}(v) = \sup_{u_1, u_2, \ldots, u_n} \left\{\min_{1 \leq i \leq n} \mu_{\tilde{x}_i}(u_i) | v = f(u_1, u_2, \ldots, u_n)\right\}.$$

Theorem 2 Let $\tilde{x}_i (i = 1, 2, \cdots, n)$ are fuzzy numbers, $\omega_i (i = 1, 2, \ldots, n)$ are positive real coefficient, let $()^L_\alpha$ and $()^U_\alpha$ denote the lower and upper bounds of the α-level set of a fuzzy number, then for any given possibility levels α_1, α_2 and α_3 with $0 \leq \alpha_1, \alpha_2, \alpha_3 \leq 1$, $Pos\{\omega_1\tilde{x}_1 + \omega_2\tilde{x}_2 + \ldots + \omega_n\tilde{x}_n \leq b\} \geq \alpha_1$ if $\omega_1(\tilde{x}_1)^L_{\alpha_1} + \omega_2(\tilde{x}_2)^L_{\alpha_1} + \ldots + \omega_n(\tilde{x}_n)^L_{\alpha_1} \leq b$; $Pos\{\omega_1\tilde{x}_1 + \omega_2\tilde{x}_2 + \ldots + \omega_n\tilde{x}_n \geq b\} \geq \alpha_2$ if $\omega_1(\tilde{x}_1)^U_{\alpha_2} + \omega_2(\tilde{x}_2)^U_{\alpha_2} + \ldots + \omega_n(\tilde{x}_n)^U_{\alpha_2} \geq b$; $Pos\{\omega_1\tilde{x}_1 + \omega_2\tilde{x}_2 + \ldots + \omega_n\tilde{x}_n = b\} \geq \alpha_3$ if $\omega_1(\tilde{x}_1)^L_{\alpha_3} + \omega_2(\tilde{x}_2)^L_{\alpha_3} + \ldots + \omega_n(\tilde{x}_n)^L_{\alpha_3} \leq b$ and $\omega_1(\tilde{x}_1)^U_{\alpha_3} + \omega_2(\tilde{x}_2)^U_{\alpha_3} + \ldots + \omega_n(\tilde{x}_n)^U_{\alpha_3} \geq b$.

Proof Only the proof for the first case will be given. The other cases can be proved using similar arguments.

$Pos\{\omega_1\tilde{x}_1 + \omega_2\tilde{x}_2 + \ldots + \omega_n\tilde{x}_n \leq b\} \geq \alpha_1$, suppose that $(s_1^*, \ldots, s_n^*) =^\Delta \arg \sup_{s_1, \ldots, s_n \in R} \{\min[\mu_{\tilde{x}_1}(s_1), \ldots, \mu_{\tilde{x}_n}(s_n)] | \omega_1 s_1 + \omega_2 s_2 +, \ldots, + \omega_n s_n \leq b\}$.

Then $\min\{\mu_{\tilde{x}_1}(s_1), \ldots, \mu_{\tilde{x}_n}(s_n)\} \geq \alpha_1$ and $\omega_1 s_1^* + \omega_2 s_2^* + \ldots + \omega_n s_n^* \leq b$.

$\min\{\mu_{\tilde{x}_1}(s_1), \ldots, \mu_{\tilde{x}_n}(s_n)\} \geq \alpha_1$ implies that $\mu_{\tilde{x}_1}(s_1^*) \geq \alpha_1, \ldots, \mu_{\tilde{x}_n}(s_n^*) \geq \alpha_1$, therefore $s_1^* \in \left[(\tilde{x}_1)^L_{\alpha_1}, (\tilde{x}_1)^U_{\alpha_1}\right], s_2^* \in \left[(\tilde{x}_2)^L_{\alpha_1}, (\tilde{x}_2)^U_{\alpha_1}\right], \ldots, s_n^* \in \left[(\tilde{x}_n)^L_{\alpha_1}, (\tilde{x}_n)^U_{\alpha_1}\right]$, together with $\omega_1 s_1^* + \omega_2 s_2^* + \ldots + \omega_n s_n^* \leq b$ and the non-negativity of $\omega_i (i = 1, 2, \cdots, n)$, thus $\omega_1(\tilde{x}_1)^L_{\alpha_1} + \omega_2(\tilde{x}_2)^L_{\alpha_1} + \ldots + \omega_n(\tilde{x}_n)^L_{\alpha_1} \leq b$.

Conversely, if $\omega_1(\tilde{x}_1)^L_{\alpha_1} + \omega_2(\tilde{x}_2)^L_{\alpha_1} + \ldots + \omega_n(\tilde{x}_n)^L_{\alpha_1} \leq b$, then there exists α_1' with $\alpha_1 \leq a_1' \leq 1$ such that $\omega_1(\tilde{x}_1)^L_{\alpha_1'} + \omega_2(\tilde{x}_2)^L_{\alpha_1'} + \ldots + \omega_n(\tilde{x}_n)^L_{\alpha_1'} \leq b$, obviously $\mu_{\tilde{x}_1}\left((\tilde{x}_1)^L_{\alpha_1'}\right) \geq \alpha_1, \ldots, \mu_{\tilde{x}_n}\left((\tilde{x}_n)^L_{\alpha_1'}\right) \geq \alpha_1$, this is equivalent to $\min\left\{\mu_{\tilde{x}_1}(\tilde{x}_1)^L_{\alpha_1'}, \ldots, \mu_{\tilde{x}_n}(\tilde{x}_n)^L_{\alpha_1'}\right\} \geq \alpha_1$, thus $\min\left\{\mu_{\tilde{x}_1}\left((\tilde{x}_1)^L_{\alpha_1'}\right), \ldots, \mu_{\tilde{x}_n}\left((\tilde{x}_n)^L_{\alpha_1'}\right) | \omega_1 (\tilde{x}_1)^L_{\alpha_1'} + \omega_2(\tilde{x}_2)^L_{\alpha_1'} + \ldots + \omega_n(\tilde{x}_n)^L_{\alpha_1'} \leq b\right\} \geq \alpha_1$, therefore $Pos\{\omega_1\tilde{x}_1 + \omega_2\tilde{x}_2 + \ldots + \omega_n\tilde{x}_n \leq b\} = \sup_{s_1, \ldots, s_n \in R} \min\{\mu_{\tilde{x}_1}(s_1), \ldots, \mu_{\tilde{x}_n}(s_n) | s_1 + \ldots + s_n \leq b\} \geq \alpha_1$.

3 Methodology

The data classification is to judge whether the data belong to a specified group according to some observed characteristics. In other words, the data classification is to construct a discriminant function using a preselected set of data, called sample training set, and then to test the class label for a new sample.

Fuzzy DEA-based classifier is to use the fuzzy DEA model to construct such a discriminant function.

3.1 Incremental fuzzy DEA-based classifier

In business, the given set of features is an incremental fuzzy data, that is, it satisfies the rule that "bigger is better" is applied. This implies that larger feature values have greater object preference.

Consider the fuzzy training sample set $S = \left\{(\tilde{X}_1, y_1), (\tilde{X}_2, y_2), \ldots, (\tilde{X}_l, y_l)\right\}$, where $\tilde{X}_j \in T^n(R), y_j \in \{-1, 1\}, j = 1, 2, \cdots l$. If $y_i = 1$, then $(\tilde{X}_i, y_i)$ is called a positive class; If $y_i = -1$, then $(\tilde{X}_i, y_i)$ is called a negative class. Without loss of generality, the positive fuzzy training sample set is denoted by $S^+ = \left\{(\tilde{X}_1, y_1), (\tilde{X}_2, y_2), \ldots, (\tilde{X}_{l_1}, y_{l_1})\right\}, S^-$

$= \left\{ \left(\tilde{X}_{l_1+1}, y_{l_1+1}\right), \left(\tilde{X}_{l_1+2}, y_{l_1+2}\right), ..., \left(\tilde{X}_l, y_l\right) \right\}$ is the negative fuzzy training sample set.

The classification based on the fuzzy training set $S = \left\{ \left(\tilde{X}_1, y_1\right), \left(\tilde{X}_2, y_2\right), ..., \left(\tilde{X}_l, y_l\right) \right\}$ is to find a decision function $g\left(\tilde{X}\right)$, such that the positive class and the negative class can separated with the low classification error and good generalization performance .

For simplicity, we suppose that the fuzzy training data $\tilde{X}_k = (\tilde{x}_{1k}, \tilde{x}_{2k}, ..., \tilde{x}_{nk})$ $(k = 1, 2, ..., n)$ are vectors of the triangular fuzzy numbers, $\tilde{x}_{ik} = (l_{ik}, m_{ik}, r_{ik}), i = 1, 2, ..., n; k = 1, 2, ..., l$.

Definition 3.1 For the fuzzy training set $S = S^+ \cup S^-$ and a given possibility level $\alpha(0 \le \alpha \le 1)$, if there are positive vector $\omega \ge 0$ and positive number μ such that $Pos\{\omega \cdot \tilde{x}_i - \mu > 0\} \ge \alpha$ for $\tilde{x}_i \in S^+$, then the fuzzy training set is called incremental separable with respect to the possibility level α; The fuzzy hyper-plane $\omega \cdot \tilde{x}_i - \mu \cong \tilde{0}$ is called a fuzzy classification hyper-plane at level α.

For incremental separable fuzzy training set $S = S^+ \cup S^-$ the acceptance domain can be represented as followings:

$$\Theta = \left\{ \tilde{x} \middle| \sum_{i=1}^{l_1} \lambda_i \tilde{x}_i \tilde{<} \tilde{x}, \tilde{x}_i \in S^+, \sum_{i=1}^{l_1} \lambda_i \ge 1, \lambda_i \ge 0 \right\}$$

The fuzzy CCR model with input-output $\left(\tilde{X}_i, 1\right)$ for DMU$_j$, $j = 1, 2, ..., l_1$, to evaluate the DMU_{j_0} is given by

$$\begin{cases} \max \mu_0 \\ \omega \tilde{x}_i - \mu_0 \tilde{>} 0, i = 1, 2, ..., l_1; \\ \omega \tilde{x}_0 \cong 1; \\ \omega \ge 0, \mu_0 \ge 0. \end{cases} \tag{5}$$

To solve the above fuzzy model (5), we apply the possibility approach [30, 31]. For given possibility level α, $\beta(0 \le \alpha, \beta \le 1)$, the possibility model of (5) is as follows:

$$\begin{cases} \max \mu_0 \\ pos\left\{\omega \tilde{x}_i - \mu_0 \ge 0\right\} \ge \alpha, i = 1, 2, ..., l_1; \\ pos\left\{\omega \tilde{x}_0 = 1\right\} \ge \beta; \\ \omega \ge 0, \mu_0 \ge 0. \end{cases} \tag{6}$$

For simplicity, $\tilde{x}_{j_0} = \tilde{x}_0$ and suppose that $\alpha = \beta$. From Lemma 1, the optimization problem is equivalent to the following model

$$\begin{cases} \max \mu_0 \\ \omega (\tilde{x}_i)_\alpha^U - \mu_0 \ge 0, i = 1, 2, ..., l_1; \\ \omega (\tilde{x}_0)_\alpha^L \le 1, \omega \left(\tilde{x}_0\right)_\alpha^U \ge 1; \\ \omega \ge 0, \mu_0 \ge 0. \end{cases} \tag{7}$$

Definition 3.2 For the evaluated DMU_{j_0} : $\left(\tilde{x}_{j_0}, 1\right) \in S^+$, if the optimal values of (6) is not less than 1, then $\left(\tilde{x}_{j_0}, 1\right)$ is called a efficient classification point of S^+.

For given possibility level α, $\beta(0 \le \alpha, \beta \le 1)$and an efficient classification point of S^+, by the model (7) (or (6) for $\alpha = \beta$),we can obtain a normal vector of the efficient frontier surface of the fuzzy acceptance domain Θ.

If training sample set $S = \{(\tilde{X}_1, y_1), (\tilde{X}_2, y_2), ..., (\tilde{X}_l, y_l)\}$ is a classical training sample set, that is to say $\tilde{X}_j \in R^n, j = 1, 2, ..., l$, then fuzzy DEA-based classifier is degenerated into data envelopment classification machine [28].

Theorem 3 If $\left(\tilde{x}_{j_0}, 1\right)$ is a efficient classification point of S^+, $\omega = \omega_0$, μ_0 are the optimal solution of (6), then for every $\tilde{x} \in \Theta, Pos\{\omega_0 \tilde{x} - \mu_0 \ge 0\} \ge \alpha$.

Proof Because $\omega_0 > 0$, μ_0 are the optimal solution of (6), thus $pos\{\omega_0 \tilde{x}_i - \mu_0 \ge 0\} \ge \alpha, i = 1, 2, ..., l_1, pos\{\omega \tilde{x}_0 = 1\} \ge \alpha$; By the lemma 2, for every $\tilde{x} \in \Theta, \omega_0 \tilde{x} - \mu_0 \tilde{>} \omega_0 \sum_{i=1}^{l_1} \lambda_i \tilde{x}_i - \mu_0 = \sum_{i=1}^{l_1} \lambda_i (\omega_0 \tilde{x}_i - \mu_0) + \mu_0 \left(\sum_{i=1}^{l_1} \lambda_i - 1 \right) \tilde{>} \sum_{i=1}^{l_1} \lambda_i (\omega_0 \tilde{x}_i - \mu_0)$ and by theorem 2, $Pos\{\omega_0 \tilde{x}_i - \mu_0 \ge 0\} \ge \alpha$ if and only if $\omega_0 (\tilde{x}_i)_\alpha^U - \mu_0 \ge 0$ for $i = 1, 2, ..., l_1$.Thus $\sum_{i=1}^{l_1} \lambda_i \left[\omega_0 (\tilde{x}_i)_\alpha^U - \mu_0 \right] \ge 0$, that is, $\omega_0 \sum_{i=1}^{l_1} \lambda_i (\tilde{x}_i)_\alpha^U - \mu_0 \sum_{i=1}^{l_1} \lambda_i \ge 0$, then for every $\tilde{x} \in \Theta$, $(\omega_0 \tilde{x} - \mu_0)_\alpha^U = \omega_0 (\tilde{x})_\alpha^U - \mu_0 \ge \omega_0 \sum_{i=1}^{l_1} \lambda_i (\tilde{x}_i)_\alpha^U - \mu_0 \ge \omega_0 \sum_{i=1}^{l_1} \lambda_i (\tilde{x}_i)_\alpha^U - \mu_0 \sum_{i=1}^{l_1} \lambda_i = \sum_{i=1}^{l_1} \lambda_i \left[\omega_0 (\tilde{x}_i)_\alpha^U - \mu_0 \right]$, also by theorem 2, $Pos\{\omega_0 \tilde{x} - \mu_0 \ge 0\} \ge \alpha$.

To take account of the noise of the training data or the fact that some data points may be misclassified, we introduce a vector of slack variables $\xi = \left(\xi_1, ..., \xi_{l_1}\right)^T$ that measure the amount of violation of the constraints, the corresponding fuzzy CCR model of (5) is as follows:

$$\begin{cases} \max \mu_0 - C \sum_{i=1}^{l_1} \xi_i \\ \omega \tilde{x}_i - \mu_0 \tilde{>} 0 + \xi_i, i = 1, 2, ..., l_1; \\ \omega \tilde{x}_0 \cong 1; \\ \xi_i \ge 0; \\ \omega \ge 0, \mu_0 \ge 0. \end{cases} \tag{8}$$

where $C > 0$ is specified beforehand. For given possibility level α, $\beta(0 \le \alpha, \beta \le 1)$, the possibility model of (8) is as follows:

$$\begin{cases} \max \mu_0 - C \sum_{i=1}^{l_1} \xi_i \\ pos\left\{\omega \tilde{x}_i - \mu_0 \ge 0 + \xi_i\right\} \ge \alpha, i = 1, 2, ..., l_1; \\ pos\left\{\omega \tilde{x}_0 = 1\right\} \ge \beta; \\ \xi_i \ge 0 \\ \omega \ge 0, \mu_0 \ge 0. \end{cases} \tag{9}$$

Also from Lemma 1, for β = α, the optimization problem (9) is equivalent to the following classical optimization model

$$\begin{cases} \max \mu_0 - C\sum_{i=1}^{l_1}\xi_i \\ \omega(\tilde{x}_i)_\alpha^U - \mu_0 \geq 0 + \xi_i, i = 1, 2, ..., l_1; \\ \omega(\tilde{x}_0)_\alpha^L \leq 1, \omega(\tilde{x}_0)_\alpha^U \geq 1; \\ \xi_i \geq 0; \\ \omega \geq 0, \mu_0 \geq 0. \end{cases} \tag{10}$$

If $\omega_0 > 0$, $\mu_0 = 1$ are the optimal solution of (7) (or (10)) and $\tilde{x}_0$ satisfies $pos\{\omega\tilde{x}_0 = 1\} \geq \alpha$.

Denote $L : \omega_0\tilde{x} - 1 \cong 0$.

The hyperplane L is a supporting plane of the acceptance domain, ω_0 is the normal vector of the hyperplane L and $\tilde{x}_0$ is on this plane.

Suppose all the supporting plane of the fuzzy training set $S = S^+ \cup S^-$ be denoted by $\omega_0^j\tilde{x} - 1 \cong 0, j = 1, 2, ..., l_0(l_0 < l_1)$, then the classification fuzzy function can be given by $d(\tilde{x}) = \min_{1 \leq i \leq l_0} Pos\left(\omega_0^i\tilde{x} - 1 \geq 0\right)$.

To judge if a given data x belong to the fuzzy acceptance domain Θ is equivalent to identify if x is above all the fuzzy hyper-plane(for a given confidence level).

For fuzzy example which unknown class $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$, the decision rule is: for a given confidence level $\alpha(0 < \alpha \leq 1)$, If $d(\tilde{x}) \geq \alpha$, then $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ is a positive example, otherwise $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ is a negative example.

The diagram of incremental fuzzy DEA-based classifier is shown in Fig. 1.

3.2 Applications in healthcare management

Our proposed fuzzy DEA-based classifier can be applied in healthcare management, like diseases diagnosis, diseases forecasting and individuals health surveillance or health warning.

In the following, the proposed fuzzy DEA-based classifier is applied in the diagnosis of Coronary with fuzzy symptoms and in prognosing the breast cancer with fuzzy disturbance data. The applications also prove the accuracy and efficiency of the proposed fuzzy DEA-based classifier.

Example 1 The diagnosis of Coronary [48]. 30 samples are randomly divided into two groups: 24 training samples and 6 test samples, the training data in Table 1 are the diastolic pressure ($\tilde{x}_{i1}$) and plasma cholesterol ($\tilde{x}_{i2}$) of twenty-four persons, where half of them are healthy ($y_i = 1$), the others are Coronary patients ($y_i = -1$), $\tilde{x}_{i1}$ and $\tilde{x}_{i2}$ are triangular fuzzy numbers.

We randomly divide the samples into training data (24 samples) and test data (6). Based on the fuzzy training data, for parameter C = 0.1, $\alpha = 0.95$, solving the programming (11), we can obtain four efficient classification points and their responding supporting planes

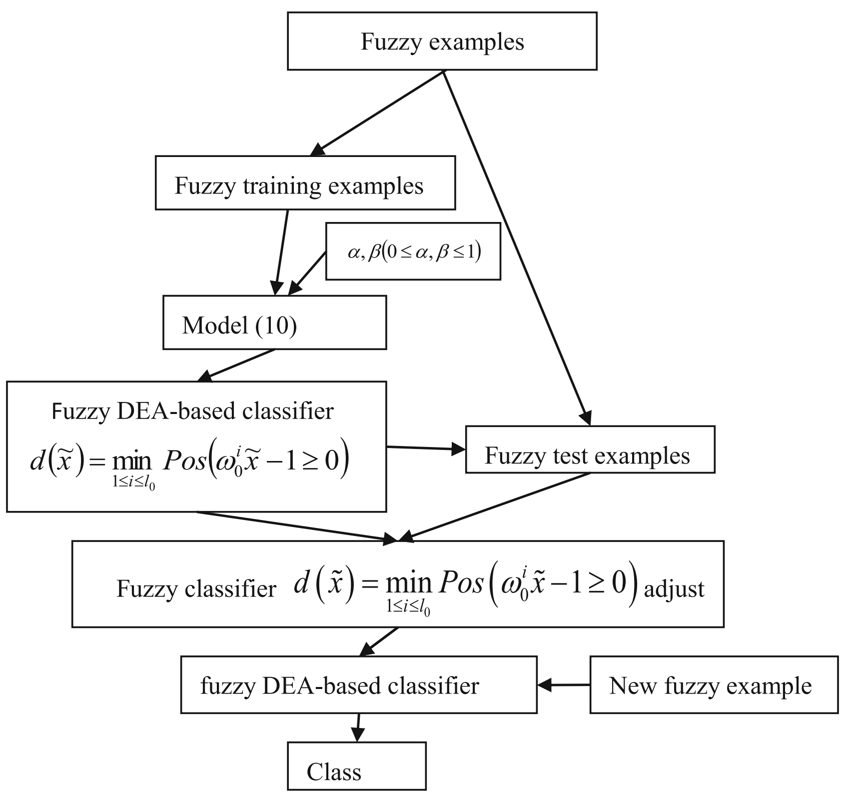

Fig. 1 The diagram of incremental fuzzy DEA-based classifier

Table 1 The data of diastolic pressure and plasma cholesterol of the patient of Coronary and healthy people

i	$\tilde{x}_{i1}$ (KPa)	$\tilde{x}_{i2}$ (mmol/L)	y_i	i	$\tilde{x}_{i1}$	$\tilde{x}_{i2}$	y_i
1	(9.84,9.86,9.88)	(5.17,5.18,5.19)	1	16	(10.62,10.66,10.70)	(2.06,2.07,2.08)	−1
2	(13.31,13.33,13.35)	(3.72,3.73,3.74)	1	17	(12.51,12.53,12.55)	(4.44,4.45,4.46)	−1
3	(14.63,14.66,14.69)	(3.87,3.89,3.91)	1	18	(13.30,13.33,13.36)	(3.04,3.06,3.08)	−1
4	(9.32,9.33,9.34)	(7.08,7.10,7.12)	1	19	(9.32,9.33,9.34)	(3.90,3.94,3.98)	−1
5	(12.87,12.80,12.83)	(5.47,5.49,5.51)	1	20	(10.64,10.66,10.68)	(4.43,4.45,4.47)	−1
6	(10.64,10.66,10.68)	(4.06,4.09,4.12)	1	21	(10.64,10.66,10.68)	(4.89,4.92,4.95)	−1
7	(10.65,10.66,10.67)	(4.43,4.45,4.47)	1	22	(9.31,9.33,9.35)	(3.66,3.68,3.70)	−1
8	(13.31,13.33,13.35)	(3.60,3.63,3.66)	1	23	(10.64,10.66,10.68)	(3.20,3.21,3.22)	−1
9	(13.32,13.33,13.34)	(5.68,5.70,5.72)	1	24	(10.37,10.40,10.43)	(3.92,3.94,3.96)	−1
10	(11.97,12.00,12.03)	(6.17,6.19,6.21)	1	25	(9.31,9.33,9.35)	(4.90,4.92,4.94)	−1
11	(14.64,14.66,14.68)	(4.00,4.01,4.02)	1	26	(11.19,11.20,11.21)	(3.40,3.42,3.44)	−1
12	(13.31,13.33,13.35)	(3.99,4.01,4.03)	1	27	(9.31,9.33,9.35)	(3.62,3.63,3.64)	−1
13	(12.72,12.80,12.88)	(5.93,5.96,5.99)	1	28	(10.64,10.66,10.68)	(4.43,4.45,4.47)	−1
14	(13.3,13.33,13.36)	(5.88,5.96,6.04)	1	29	(10.64,10.66,10.68)	(2.65,2.69,2.73)	−1
15	(10.63,10.66,10.69)	(5,5.02,5.04)	−1	30	(10.61,10.66,10.71)	(2.71,2.77,2.83)	−1

$$L_1 : f_1\left(\tilde{x}\right) = 0.0886786\tilde{x}_1 + 0.02450144\tilde{x}_2 - 1 \cong 0;$$
$$L_2 : f_2\left(\tilde{x}\right) = 0.08856158\tilde{x}_1 + 0.02496741\tilde{x}_2 - 1 \cong 0;$$
$$L_3 : f_3\left(\tilde{x}\right) = 0.02909458\tilde{x}_1 + 0.1688751\tilde{x}_2 - 1 \cong 0;$$
$$L_4 : f_4\left(\tilde{x}\right) = 0.275824\tilde{x}_2 - 1 \cong 0.$$

then the discriminant function is given as follows:

$$d(\tilde{x}) = \min_{1 \le i \le 4} Pos(f_i(\tilde{x}) - 1 \ge 0).$$

The decision rule is: for a given confidence level $\alpha = 0.95$, if $d(\tilde{x}) = \min_{1 \le i \le 4} Pos(f_i(\tilde{x}) - 1 \ge 0) \ge 0.95$, then $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)$ is a positive example; otherwise $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)$ is a negative example.

For the same fuzzy data set and confidence level, we compare with linear fuzzy support vector machine (LFSVM) [47], the experiment result is given in Table 2.

Compared with LFSVM, our proposed fuzzy DEA classification machine has better training accuracy and test training, but the computational time is longer; because the fuzzy data set isn't linear separable, the training accuracies and test accuracies of the LFSVM are worst.

Example 2 In this example, we select the dataset from University of California Irvine (UCI) repository: Wisconsin prognostic breast cancer dataset (Bennett and Mangasarian 1992) (http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Original%29).Wisconsin prognostic breast cancer dataset has 699 records and 9 variables (Clump Thickness; Uniformity of Cell Size; Uniformity of Cell Shape; Marginal Adhesion; Single Epithelial Cell Size; Bare Nuclei; Bland Chromatin; Normal Nucleoli; Mitoses). The dataset contains 16 samples that have attributes with missing values and 683 samples have complete data, these records belong to either benign or malignant class. In order to verify the robustness of our proposed fuzzy DEA-based classifier, we randomly add a fuzzy disturbance to every attribute values, then we can obtain the fuzzy samples set.

The malignant class is considered as the positive class, we randomly select 180 samples from the 239 positive samples as training samples and the rest of the positive samples and all the negative samples as the test samples. When parameter C = 0.1, $\alpha = 0.85$, solving the programming (11). We can obtain 37 efficient classification points and their responding supporting planes, then we have the discriminant function:

$$d\left(\tilde{x}\right) = \min_{1 \le i \le 37} Pos\left(\omega_0^i \tilde{x} - 1 \ge 0\right)$$

The proposed model, which is trained on the above fuzzy disturbance data set, is compared with support vector machine

Table 2 Comparison with LFSVM

Classifiers	parameter	Training accuracies	Test accuracies	training times(seconds)
FDEACM	C = 0.1, $\alpha = 0.95$	95.8%	100%	240
LFSVM	C = 0.1, $\alpha = 0.95$	91.7%	83.3%	27

Table 3 Comparison result of Robust FDEACM with SVM

Classifiers	parameter	Training accuracies	Test accuracies	training times (minute)
FDEACM	C = 0.1, α = 0.85	97.8%	96.8%	12
SVM	C = 0.1,	98.7%	96.3%	2.3

trained on original wisconsin prognostic breast cancer dataset (http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Original%29), the result is given in Table 3.

Experiments results show that although the proposed model has lower training accuracies and longer training times, it has better test accuracies than SVM. However, the results of the proposed model are obtained under the fuzzy disturbance to the data set. This numerical experiments show that the proposed fuzzy DEA classification machine is robust.

4 Conclusions and future research

In this paper, in line with the idea of Hong Yan et al. [29], we discussed and defined a DEA classification machine for fuzzy training data. We treat the fuzzy data under evaluation to be a decision making unit with the given fuzzy attribute values as the input and a single output of value 1. We then used a set of decision making units (DMUs) to form a fuzzy sample data set and construct the fuzzy acceptance domain based on the fuzzy sample set for classification. The proposed fuzzy DEA-based classifier use the piecewise-linear fuzzy discriminant functions to approximate nonlinear fuzzy discriminant functions. Experiments show that the proposed model has good performance. The theoretical work and the computational results show that the method developed has great potential in practice. But the proposed model has some drawbacks, (1) The proposed model is only applicable to the classification of incremental fuzzy data. (2) The training time of the proposed model is longer, which is not suitable for big data. In the future, we will generalize our classification model in the following directions. (1) The normal vectors of hyper-planes are positive, and it only uses the information from one class [29], the information of other class is lost. All of these may result in bad classification performance. We may construct a novel classifier based on generalized fuzzy DEA, in which the non-negative conditions can be loosed and the classification information cab be fully utilized; (2) The proposed model considers all features with equal importance, but in some cases, some features are more important than others. and we can discuss the fuzzy DEA classification machine with assurance region constraints; (3) The training data of the proposed model are fuzzy data, we can construct the stochastic data envelopment classification machine based on stochastic data envelopment analysis [49, 50]. Furthermore, build data envelopment classification machine with stochastic fuzzy training sample.

Acknowledgements This work was supported by a grant from National Social Science Fund (14BJY010), Hebei province Social Science Fund (HB18GL014).

References

1. Kim S-P, Gupta D, Ajay K et al (2015) Accept/decline decision module for the liver simulated allocation model. Health Care Manag Sci 18:35–57
2. Chen H, Cheng B-C, Liao G-T, Kuo T-C (2014) Hybrid classification engine for cardiac arrhythmia cloud service in elderly healthcare management. J Vis Lang Comput 25:745–753
3. Mahamune M, Ingle S, Deo P, Chowhan S (2015) Healthcare knowledge management using data mining techniques. Advances in Computational Research 1:274–278
4. Luukka P (2011) A New Nonlinear Fuzzy Robust PCA Algorithm and Similarity Classifier in Classification of Medical Data Sets. International Journal of Fuzzy Systems 3:153–163
5. Ishibushi H, Yamamoto T, Nakashima T (2001) Fuzzy data mining: effect of fuzzy discretization. In: Proc. of the IEEE International Conference on Data Mining, ICDM, San Jose, pp. 241–248
6. Guillaume S (2001) Designing fuzzy inference system from data: an interpretability-oriented review. IEEE Trans Fuzzy Syst 9(3): 426–443
7. Wu K, Yap K-H (2006) Fuzzy SVM for content-based image retrieval. IEEE Comput Intell Mag 1(2):10–16
8. Yuan Y, Shaw M (1995) Induction of fuzzy decision trees. Fuzzy Sets Syst 69:125–139
9. Boyen X, Wehenkel L (1999) Automatic induction of fuzzy decision tree and its application to power system security assessment. Fuzzy Sets Syst 102(1):3–19
10. Huang Y-P, Lai S-L, Sandnes FE, Liu S-I (2012) Improving Classifications of Medical Data Based on Fuzzy ART2 DecisionTrees. International Journal of Fuzzy Systems 3:444–454
11. Wang L-X, Mendel J (1992) Generating fuzzy rules by learning from examples. IEEE Trans Syst Man Cybern 22(6):1414–1427
12. Hong T-P, Chen J-B (2000) Processing individual fuzzy attributes for fuzzy rule induction. Fuzzy Sets Syst 112(1):127–140
13. Hühn J, Hüllermeier E (2009) FR3: a fuzzy rule learner for inducing reliable classifiers. IEEE Trans Fuzzy Syst 17(1):138–149
14. Wu X-H, Zhou J-J (2006) Fuzzy discriminant analysis with kernel methods. Pattern Recogn 39(11):2236–2239
15. Graves D, Pedrycz W (2010) Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study. Fuzzy Sets Syst 161(4):522–543
16. Ji A-b, Pang J-h, Qiu H-j (2010) Support vector machine for classification based on fuzzy training data. Expert Syst Appl 37(4): 3495–3498
17. Heo G, Gader P (2011) Robust kernel discriminant analysis using fuzzy memberships. Pattern Recogn 44(3):716–723
18. Baklouti R, Mansouri M, Nounou M, Nounou H, Hamida AB (2015) Iterated Robust kernel Fuzzy Principal Component Analysis and application to fault detection. J Comput Sci 2

19. Lorence DP, Spink A (2003) Assessment of preferences for classification detail in medical information: is uniformity better? Inf Process Manag 39:465–477
20. Kulldorff M, Fang Z, Walsh SJ A Tree-Based Scan Statistic for Database Disease Surveillance. Biometrics 59(2013):323–331
21. J. Han, M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufman Publishers, Inc. San Francisco, 2001
22. Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2:429–444
23. Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manag Sci 30(9):1078–1092
24. Narci HO, Ozcan YA et al (2015) An Examination of Competition and Efficiency for Hospital industry in Turkey. Health Care Management Science (4):407–418
25. Nayar P, Ozcan YA, Yu F, Nguyen AT (2013) Data Envelopment Analysis: A Benchmarking Tool for Performance in Urban Acute Care Hospitals. Health Care Manag Rev 2:137–145
26. Ozcan YA (2014) Health Care Benchmarking and Performance Evaluation: An Assessment using Data Envelopment Analysis (DEA) 2nd Edition. Springer, Newton
27. Troutt MD, Rai A, Zhang A (1996) The potential use of DEA for credit applicant acceptance systems. Comput Oper Res 23(4):405–408
28. Hasan Bal HHO (2007) Data envelopment analysis approach to two-group classification problem and experimental comparison with some classification models. Hacettepe Journal of Mathematics and Statistics 36(2):169–180
29. Yan H, Wei Q (2011) Data envelopment analysis classification machine. Inf Sci 181:5029–5041
30. Wei Q, Chang T-S, Han S (2014) Quantile–DEA classifiers with interval data. Ann Oper Res 217:535–563
31. Hatami-Marbini A, Emrouznejad A, Tavana M (2011) A taxonomy and review of the fuzzy Data Envelopment Analysis literature: Two decades in the making. Eur J Oper Res 214:457–472
32. Sengupta JK (1992) A fuzzy systems approach in data envelopment analysis. Computers and Mathematics with Applications 24(8–9): 259–266
33. Hatami-Marbini A, Saati S, Tavana M (2010) An ideal-seeking fuzzy data envelopment analysis framework. Appl Soft Comput 10(4):1062–1070
34. Kao C, Liu ST (2003) A mathematical programming approach to fuzzy efficiency ranking. Int J Prod Econ 86(2):145–154
35. Puri J, Yadav SP (2013) A concept of fuzzy input mix-efficiency in fuzzy DEA and its application in banking. Expert Syst Appl 40(5): 1437–1450
36. Guo P, Tanaka H (2001) Fuzzy DEA: a perceptual evaluation method. Fuzzy Sets Syst 119(1):149–160
37. León T, Liern V, Ruiz JL, Sirvent I (2003) A fuzzy mathematical programming approach to the assessment of efficiency with DEA models. Fuzzy Sets Syst 139(2):407–419
38. Guo P, Tanaka H, Inuiguchi M (2000) Self-organizing fuzzy aggregation models to rank the objects with multiple attributes. IEEE Transactions on Systems, Man and Cybernetics, Part A – Systems and Humans 30(5):573–580
39. Lertworasirikul S, Shu-Cherng F, Joines JA, Nuttle HLW (2003) Fuzzy data envelopment analysis (DEA): a possibility approach. Fuzzy Sets Syst 139(2):379–394
40. Muren ZM, Cui W (2014) Generalized fuzzy data envelopment analysis methods. Appl Soft Comput 19:215–225
41. Ghasemi M-R, Ignatius J, Lozano S, Emrouznejad A, Hatami-Marbini A (2015) A fuzzy expected value approach under generalized data envelopment analysis. Knowl-Based Syst 89:148–159
42. Tavana M, Shiraz RK, Hatami-Marbini A, Agrell PJ, Paryab K (2013) Chance-constrained DEA models with random fuzzy inputs and outputs. Knowl-Based Syst 52:32–52
43. Angulo Meza L, Pereira Estellita M (2002) Lins, "Review of methods for increasing iscrimination in Data Envelopment Analysis". Ann Oper Res 116(1–4):225–242
44. Dotoli M, Epicoco N, Falagario M, Sciancalepore F (2015) A cross-efficiency fuzzy Data Envelopment Analysis technique for performance evaluation of Decision Making Units under uncertainty. Comput Ind Eng 79:103–114
45. Toloo M, Kresta A (2014) Finding the best asset financing alternative: A DEA-WEO approach. Measurement 55:288–294
46. Toloo M (2013) The most efficient unit without explicit inputs: an extended MILP–DEA model. Measurement 46:3628–3634
47. Ramik J, Rimanek J (1985) Inequality relation between fuzzy numbers and its use in fuzzy optimization. Fuzzy Sets Syst 16:123–138
48. Ji A-b, Pang J-h, Qiu H-j (2010) Support vector machine for classification based on fuzzy training data. Expert Syst Appl 37:3495–3498
49. Olesen OB, Petersen NC (2016) Stochastic Data Envelopment Analysis- A review. Eur J Oper Res 251(1):2–21
50. Dotoli M, Epicoco N, Falagario M, Sciancalepore F (2016) A stochastic cross-efficiency Data Envelopment Analysis approach for supplier selection under uncertainty. Int Trans Oper Res 23(4):725–748