Knowledge-Based Systems 85 (2015) 62–70

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A cooperative belief rule based decision support system for lymph node metastasis diagnosis in gastric cancer

Zhi-Guo Zhou [a,b], Fang Liu [a,b,*], Ling-Ling Li [b], Li-Cheng Jiao [b], Zhi-Jie Zhou [c], Jian-Bo Yang [d], Zhi-Long Wang [e]

[a] School of Computer Science and Technology, Xidian University, Xi'an 710071, PR China
[b] Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an 710071, PR China
[c] High-Tech Institute of Xi'an, Xi'an, Shaanxi 710025, PR China
[d] Manchester Business School, The University of Manchester, Manchester M15 6PB, UK
[e] Key Laboratory of Carcinogenesis and Translational Research of Ministry of Education of China, Department of Radiology, Peking University School of Oncology, Beijing Cancer Hospital & Institute, Beijing 100142, PR China

ARTICLE INFO

Article history:
Received 5 April 2014
Received in revised form 21 April 2015
Accepted 22 April 2015
Available online 12 May 2015

Keywords:
Clinical decision support system
Gastric cancer
Lymph node metastasis
Belief rule base
Cooperative coevolutionary algorithm

ABSTRACT

Lymph Node Metastasis (LNM) has become one of the most important prognostic factors regarding long-term survival in gastric cancer. As it is difficult for doctors to integrate multiple factors for a comprehensive analysis, Clinical Decision Support System (CDSS) is used to help the analysis. In this paper, a new Cooperative Belief Rule Based (CBRB) prototype CDSS is proposed. CBRB consists of two independent Belief Rule Base (BRB) systems and the final output is combined by the Evidential Reasoning (ER) approach. A corresponding new Cooperative CoEvolutionary Algorithm (CCEA) is proposed to train the proposed CBRB model that is nonlinear. A case study demonstrates that the proposed CDSS prototype can obtain the better performance than other CDSS.

1. Introduction

Gastric cancer has become one of the leading causes for cancer-related death around the world [1]. Lymph Node Metastasis (LNM) is one of most important prognostic factors regarding long-term survival [2–5]. As such, diagnosing LNM accurately is very important. Currently, doctors diagnose LNM mainly according to the size of lymph node. However, large lymph nodes may be caused by inflammation, while small ones may be metastatic [6]. So a single lymph node size is not a strong predictor and more factors should be considered by applying Clinical Decision Support System (CDSS).

In previous research, several CDSSs have been proposed for LNM diagnosis, such as Artificial Neural Network (ANN) based CDSS [7], and Support Vector Machine (SVM) based CDSS [8]. However, both of them have some limitations. First of all, since they are black-box modeling methods, the reasoning process cannot be seen and doctors do not know which factors are more important for diagnosis. Moreover, as doctors play a critical role in diagnosis, it is important to diagnose LNM by using both clinical data and doctors' knowledge [9].

According to the above analysis, a knowledge-based CDSS which can capture human judgments is more suitable for LNM diagnosis. In [9], a Bi-level Belief Rule Based (BBRB) CDSS was proposed by the authors. That CDSS is constructed based on a recently developed belief Rule-based Inference Methodology using the Evidential Reasoning approach (RIMER) [10]. The CDSS consists of two parts: (1) a Clinical domain knowledge base modeled by Belief Rule Base (BRB) and (2) a reasoning process supported by the Evidential Reasoning (ER) approach [11–15]. Compared to other knowledge based CDSSs, BBRB have some advantages as described in [9].

In BBRB, however, only the number and size of lymph nodes are utilized for LNM diagnosis. In fact, LNM is not only related with lymph node factors, but also with tumor factors [16–20]. To fully utilize the above factors, a Cooperative Belief Rule Based (CBRB) prototype CDSS is proposed in this paper. CBRB consists of two independent BRBs. One is used for modeling lymph node factors, while the other is used for modeling tumor factors. The final result is obtained by integrating the output of two BRBs, which is

* Corresponding author at: School of Computer Science and Technology, Xidian University, Xi'an 710071, PR China. Tel.: +86 029 88204310; fax: +86 029 88201023.
E-mail address: f63liu@163.com (F. Liu).

http://dx.doi.org/10.1016/j.knosys.2015.04.019
0950-7051/

implemented by the ER approach. Meanwhile, as manually constructed belief rule may not be accurate, it is necessary to train CBRB. In this paper, a corresponding new Cooperative CoEvolutionary Algorithm (CCEA) is designed, which can be used to optimize the two BRBs and weight coefficients simultaneously.

The rest of the paper is organized as follows. The problem formulation is shown in Section 2. In Section 3, the CBRB CDSS prototype is presented. A new CCEA based method for optimizing CBRB is developed in Section 4. In Section 5, the proposed CDSS prototype for diagnosing LNM is presented. The validation of BBRB is discussed in Section 6. This paper is concluded in Section 7.

2. Problem formulation

Suppose that $T = [T_1, T_2, \ldots, T_M]$ is the set of diagnostic factors which are extracted from tumor, while $L = [L_1, L_2, \ldots, L_N]$ is the set of factors from lymph nodes, where M and N are the number of attributes for the two type of factors respectively. Suppose that P^T and P^L are the corresponding parameter vectors for the two BRBs, $\omega = [\omega_1, \omega_2]$ is the set of weight coefficients that represent the relative importance for each BRB. In other words, the problem is in essence to construct a causal relationship between the medical attributes and the output. As there are two types of factors, two independent BRBs are constructed for them at first, which are represented as follows:

$$D_T = f_T(T, P_T) \tag{1}$$

$$D_L = f_T(L, P_L) \tag{2}$$

where f_T is a diagnostic function for tumor factors, and f_L is the function for lymph node factors. D_T and D_L are the outputs of the two functions. Since there are four stages in LNM diagnosis, D_T and D_L can also be represented by the following belief distributions:

$$D_T = \{(D_T^0, \beta_T^0), (D_T^1, \beta_T^1), (D_T^2, \beta_T^2), (D_T^3, \beta_T^3)\} \tag{3}$$

$$D_L = \{(D_L^0, \beta_L^1), (D_L^1, \beta_L^1), (D_L^2, \beta_L^2), (D_L^3, \beta_L^3)\} \tag{4}$$

where $\beta_T^i (i = 0, \ldots, 3)$ and $\beta_L^j (j = 0, \ldots, 3)$ are the corresponding belief degrees for tumor and lymph node at each stage. Then the final output D can be obtained by combining the above two outputs, which is given as follows:

$$D = f(D_T, D_L, \omega) \tag{5}$$

where f is in general a function of D_T, D_L and ω. The final output can also be represented by the following belief distribution:

$$D = \{(D^0, \beta^0), (D^1, \beta^1), (D^2, \beta^2), (D^3, \beta^3)\} \tag{6}$$

where $D^i (i = 0, \ldots, 3)$ represents the final stage. $\beta^i (i = 0, \ldots, 3)$ is the belief degree, which should satisfy the following constraints:

$$\sum_{i=0}^{3} \beta^i = 1 \text{ and } 0 \leqslant \beta^i \leqslant 1, i = 0, \ldots, 3 \tag{7}$$

As tumor factors and lymph node factors are utilized simultaneously, CBRB is presented in the following section. In addition, the parameters P^T, P^L are initialized by experts and may not be accurate. Therefore, a new CCEA based method is proposed, which will be described in Section 4.

3. CBRB prototype CDSS

As CBRB is constructed on the basis of BRB, BRB will be briefly described at first. Then the proposed CBRB model will be presented.

3.1. A brief description of BRB and ER

BRB consists of belief rules for domain knowledge representation. A typical belief rule is defined as follows [10]:

$$R_k^i: \text{ If } x_1 \text{ is } A_1^k \wedge x_2 \text{ is } A_2^k \cdots \wedge x_M \text{ is } A_M^k, \text{ Then } \{(D_1, \beta_{1,k}), \ldots, (D_N, \beta_{N,k})\} \text{ With a rule weight } \theta_k \text{ and attribute weight } \delta_{1,k}, \delta_{2,k}, \ldots, \delta_{M,k} \tag{8}$$

where $x_1, x_2, \ldots, x_M$ represents the antecedent attributes in the kth rule. The referential value of the jth antecedent attribute in the kth rule is denoted by $A_j^k (j = 1, \ldots, M,\ k = 1, \ldots, L)$ and "$\wedge$" represents the "AND" relationship, where M is the number of medical antecedent attributes, and L represents the number of rules. $D_j (j = 1, \ldots, M)$ represents the stage of LNM, and $\beta_{j,k} (j = 1, \ldots, N, k = 1, \ldots, L)$ is the corresponding belief degree. Note that the kth rule is complete if $\sum_{j=1}^{N} \beta_{j,k} = 1$; otherwise, it is incomplete. $\theta_k (k = 1, \ldots, L)$ represents the relative weight of the kth rule, while $\delta_{j,k} (j = 1, \ldots, M)$ is the relative weight of the jth antecedent attributes in the kth rules, which demonstrates the relative importance of each medical attribute.

The inference of BRB and the final integration is implemented by the ER approach, which was originally proposed to deal with Multiple Attribute Decision Analysis (MADA) problems [11,12]. The kernel of the ER approach is the ER algorithm which is developed based on the Dempster–Shafer (D–S) theory of evidence [20,21] and the decision theory. It is utilized to aggregate nonlinear information under uncertainty, which consists of the following two steps.

(1) Calculating the activation weight.
The activation weight of the kth rule ω_k is calculated as:

$$\omega_k = \theta_k \prod_{i=1}^{M} (\alpha_i^k)^{\bar{\delta}_{i,k}} \Big/ \sum_{l=1}^{L} \theta_l \prod_{i=1}^{M} (\alpha_i^l)^{\bar{\delta}_{i,k}} \text{ and}$$
$$\bar{\delta}_{i,k} = \delta_{i,k} \Big/ \max_{i=1,\ldots,M} \{\delta_{i,k}\} \tag{9}$$

where $\bar{\delta}_{i,k}\ (i = 1, \ldots, M)$ represent the normalized attribute weights. $\alpha_i^k\ (i = 1, \ldots, M)$, which is named as the individual matching degree, is the degree of belief to which the input for the ith antecedent attribute belongs to its jth referential value A_i^k in the kth rule. It can be generated using various ways [22]. α_k is the normalized combined matching degree, that is:

$$\alpha_k = \prod_{i=1}^{M} (\alpha_i^k)^{\bar{\delta}_{ki}} \tag{10}$$

(2) Combining rules by the ER approach.
The final conclusion D which is generated through aggregating all rules using the ER analytical algorithms [23] is represented as:

$$D = f(x) = \{(D_j, \hat{\beta}_j),\ j = 1, \ldots, N\} \tag{11}$$

where $\hat{\beta}_j$ denotes the belief degree in D_j, and is calculated as follows,

$$\hat{\beta}_j = \mu \times \left[\prod_{k=1}^{L} \left(\omega_k \beta_{j,k} + 1 - \omega_k \sum_{i=1}^{N} \beta_{i,k} \right) - \prod_{k=1}^{L} \left(1 - \omega_k \sum_{i=1}^{N} \beta_{i,k} \right) \right] \Big/ 1 - \mu \times \left[\prod_{k=1}^{L} (1 - \omega_k) \right] \tag{12}$$

$$\mu = \left[\sum_{j=1}^{N}\prod_{k=1}^{L}\left(\omega_k\beta_{j,k} + 1 - \omega_k\sum_{i=1}^{N}\beta_{i,k}\right) - \left(N-1\right)\prod_{k=1}^{L}\left(1-\omega_k\sum_{i=1}^{N}\beta_{i,k}\right)\right]^{-1} \quad (13)$$

3.2. Proposed CBRB model

The proposed CDSS prototype is shown in Fig. 1. Tumor factors $T_1,\ldots,T_M$ and lymph node factors $L_1,\ldots,L_N$ are the inputs of the BRB models. BRB_T represents BRB for tumor factor, while BRB_L represents BRB for lymph node. Then the corresponding outputs D_T and D_L can be obtained as the outputs of the two models. The final result D can be obtained by combining the outputs of the two BRBs using the ER approach.

Suppose that the evaluating result for one patient in BRB_T is $D_T = \{(D_T^0, 0.00), (D_T^1, 0.70), (D_T^2, 0.30), (D_T^3, 0.00)\}$, while the output for BRB_LN is $\{(D_T^0, 0.00), (D_T^1, 0.60), (D_T^2, 0.40), (D_T^3, 0.00)\}$. The weight for the two BRBs is $\{\omega_1 = 0.4, \omega_2 = 0.6\}$. Then the ER approach which will be shown in Section 3.2 is employed for combining the two outputs. After calculation, the final output is given by $D = \{(D^0, 0.00), (D^1, 0.6601), (D^2, 0.3399), (D^3, 0.00)\}$. The result shows that the prediction is mainly stage 1. Moreover, it shows the changeable tendency of the illness because the degree of stage 2 is 0.3399, which means that the condition may become worse in the future. As such, doctors should consider both stage 1 and stage 2 when creating a treatment plan for the patient that $P_i, i = 1,2$ represents a vector composed of the above parameters in the two BRBs. P_i includes rule weight, attribute weights and consequent belief degrees as follows:

$$P_i = [\theta_1^i,\ldots,\theta_L^i,\delta_1^i,\ldots,\delta_M^i,\beta_{1,1}^i,\ldots,\beta_{N,L}^i]^T, \quad i = 1,2 \quad (14)$$

They should satisfy the following constraints:

$$0 \leqslant \theta_k^i \leqslant 1, \quad 0 \leqslant \delta_m^i \leqslant 1, \quad 0 \leqslant \beta_{j,k}^i \leqslant 1, \quad \sum_{j=1}^{N}\beta_{j,k}^i = 1, k = 1,\ldots,L, \quad m = 1,\ldots,M, \quad j = 1,\ldots,N \quad (15)$$

In addition to the above parameters, there is a weight vector $\omega = [\omega_1, \omega_2]$, which represents the relative importance of the two BRB models under the following constraints:

$$0 \leqslant \omega_i \leqslant 1, \quad \sum_{i=1}^{2}\omega_i = 1, \quad i = 1,2 \quad (16)$$

The initial parameters $P_i, i = 1,2$ and $\omega = [\omega_1, \omega_2]$ in CBRB are given using clinical domain knowledge and may not be accurate. Therefore, it is necessary to adjust them. In the following section, a new CCEA based optimization method is proposed to fine-tune these parameters for obtaining the optimal performance.

As is well-known, cooperative system is defined to be a system of multiple dynamic entities which share information or tasks to achieve a common object [36]. In CBRB, the objective of the two BRBs is to diagnose LNM. To improve the final diagnostic performance in CBRB, when constructing the two models, they share the trained parameters between each other, which can be seen in the following section.

There are other inference methods, such as Bayesian inference and fuzzy inference. In Bayesian inference, complete knowledge on all parameters and prior distribution are required, which can be time consuming and over-ambitious in the sense that it can be difficult to get appropriate prior distribution. In fuzzy inference, it always needs a fuzzification and de-fuzzification steps, which causes controversial in most fuzzy logic based inference system. However, the ER approach can overcome the above drawback and is regarded to be more suitable for LNM diagnosis as investigated in this paper.

4. New optimization model and CCEA for training CBRB

In this section, a new optimization model for training CBRB and a corresponding CCEA based optimization algorithm is proposed.

4.1. Optimization model for training CBRB

The objective is to find a set of parameters which can lead to better diagnostic performance. Fig. 2 shows the corresponding model. T, L are the given input, $\hat{O}$ is the observed output, and O is the simulated output. ζ is the objective function, which represents the diagnosing error rate, that is:

$$\min_{P_i,\omega}\{\zeta(P_i,\omega)\}, \quad i = 1,2 \quad (17)$$

As $P_i, i = 1,2$ and ω are two types of parameters which affect each other, they cannot be optimized simultaneously. So a new CCEA is proposed, which will be described in the following subsection.

4.2. The proposed CCEA

Several methods for training a single BRB have been proposed. The typical method for training a single BRB was proposed by Yang et al. [22]. Then a recursive online optimal method for updating BRB was proposed [24–26]. The above strategies are sensitive to initialization and may achieve local optima. Therefore, a Clonal Selection Algorithm (CSA) based method was proposed by the authors [9], which can obtain better performance. However, in CBRB, multiple BRBs and the corresponding weight coefficients should be optimized cooperatively, while the parameters in each BRB model and the weight coefficients of the two BRB model will affect each other. So, the current available methods cannot be used to train CBRB., and new methods which can optimize multiple BRBs and their weight coefficients need to be developed.

CCEA is an extended method which was proposed for improving the performance of Evolutionary Algorithms (EAs) for problems with multiple subcomponents [27]. It is suitable for optimization problems having certain structures which can be described by its nature of decomposition [28]. In CBRB, there are two BRBs which

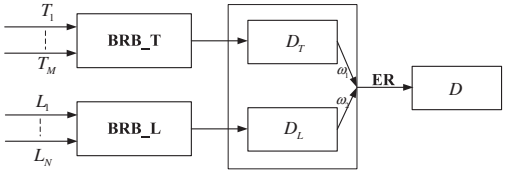

Fig. 1. The CBRB CDSS prototype.

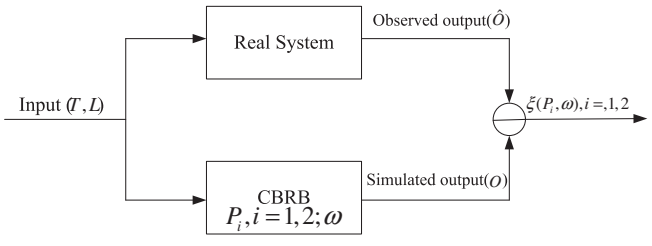

Fig. 2. Optimization model for training CBRB.

should be optimized, separately. Nevertheless, the objective function for each BRB is the same. As CCEA can solve the problems by partitioning potential solutions into smaller components that can be solved separately, it is suitable for optimizing CBRB. According to the nature of the CBRB, a new CCEA is proposed in this paper.

It is known that CCEA should initialize the population at first. In CBRB, two BRBs and the corresponding weight coefficients need to be optimized. Therefore, three initial populations should be generated. Currently, most existing CCEAs initialize population randomly. However, the clinical expert knowledge can be embedded into CBRB, which can improve the diagnostic performance. And the initial populations for two BRBs are suitable places to embed clinical domain knowledge.

In this paper, initial rules in two BRBs are extracted through expert knowledge. To obtain enough individuals, other individuals are generated based on the individuals provided by experts. Assume that the generated population for BRB_T is:

$$Pop_T(j) = \{I_{T,j}^1, I_{T,j}^2, \ldots, I_{T,j}^{Num}\}, \quad j = 0 \tag{18}$$

The generated population for BRB_L is:

$$Pop_L(j) = \{I_{L,j}^1, I_{L,j}^2, \ldots, I_{L,j}^{Num}\}, \quad j = 0 \tag{19}$$

where Num is the population number. Besides these, we need to initialize the weight coefficient. As the number of parameters is small, it is initialized randomly. The generated population for weight vector is represented as:

$$Pop_\omega(j) = \{I_{\omega,j}^1, I_{\omega,j}^2, \ldots, I_{\omega,j}^{Num}\}, \quad j = 0 \tag{20}$$

Assume that the number of iteration is I, and G is the maximum number of evolutionary generation in each iteration.

In the following steps, the parameters in weight coefficients, BRB_T and BRB_LN will be trained step by step, which means that a corresponding training algorithm in each step is needed, especially for training BRB. In [22], a method for training BRB was proposed. However, this method is sensitive to initialization and is only capable of searching for local optimal solutions. To overcome this problem, a new Clonal selection Algorithm (CSA) which is also one of the evolutionary computation methods was proposed in [9], which can improve the performance greatly. In this paper, to keep a better performance, CSA is also utilized for training weight coefficients and model parameter in BRB. The whole process is shown as follows.

Step 1: The weight coefficients are trained. When training the weight coefficients, we need the optimal parameters from two BRBs. In the first iteration, the best individuals in the two BRBs are obtained by calculating the following objective function:

$$\xi(P) = \frac{C}{M} \tag{21}$$

where M is the number in training set, and C is the number of error prediction samples. P is the parameters in each BRB, respectively. It is shown that the optimal individuals are obtained by calculating the single BRB. Assume that the best individuals from the two BRBs are denoted by B_T, B_{LN}. Then weight coefficients are optimized.

Assume that output of BRB_T is obtained from the following equation,

$$\xi_T = f_T(B_T) \tag{22}$$

where ξ_T is the objective-function. Similarity, the output of BRB_LN is:

$$\xi_{LN} = f_{LN}(B_{LN}) \tag{23}$$

where ξ_{LN} is the fitness value. Therefore, the final objective-function is:

$$\xi = \omega_1 * f_T(B_T) + \omega_2 * f_{LN}(B_{LN}) \tag{24}$$

where ξ is the final object-function value. If Eq. (24) is represented by f, the objective is to minimize the following Eq., that is:

$$\min_{\omega} f(B_T, B_{LN}, \omega) \tag{25}$$

At first, the fitness is calculated for each individual in population $Pop_\omega(j)$ according to Eq. (24). Then clonal operator is executed and the clonal scale is denoted by CS_w. So each individual in $Pop_\omega(j)$ is cloned CS_w times. In the third step, Mutation operator is operated on the cloned population. Several individuals are selected for mutation with probability $1/CS_w$. In the following step, the selected individuals are replaced by the randomly generated new individuals, which also satisfy the constraint in Eq. (16). Then the population is updated and Num best individuals are selected. A cloned individual is replaced only when its fitness is smaller than the original individual. The above three steps are executed iteratively until $j = G$ and the individual with minimal objective-function value is considered as output which is denoted by B_ω. Meanwhile, B_ω is used for training two BRBs. The corresponding algorithm is shown in Table 1.

Step 2: The parameters in BRB_T are trained. In this step, another CSA based method is proposed, which is similar to the training method in [9]. The objective-function is Eq. (24). However, as the training parameters are different, the optimization function is also different, which is shown as follows:

$$\min_{B_T} f(B_T, B_{LN}, \omega) \tag{26}$$

The clonal operator is the same as step 1 and the Group based Mutation Operator (GMO) [9] is employed in the second step. Then the new population is also selected. The three steps are executed repeatedly until $j = G$. The best individual is kept for training the other BRB and weight coefficient.

Step 3: The parameters in BRB_LN are trained. The only difference between this step and step 2 is the objective-function, which is shown as:

$$\min_{B_{LN}} f(B_T, B_{LN}, \omega) \tag{27}$$

Step 4: Step 1–3 are executed iteratively until the number of iteration equals I. The corresponding algorithm is given in Table 2. Based on CBRB and new CCEA, the process for LNM diagnosis is described in section 5.

Table 1
Training weight coefficient in CBRB.

Input: Initialize population $Pop_\omega(j)$ and calculate ζ^j using Eq. (24), Set $j = 0$
while $j < G$
  Each individual in $Pop_\omega(j)$ is cloned CS times
  Perform mutation operator on the cloned population
  Compute the objective-function values for each individual in the cloned population
  Select individuals and form next $Pop_\omega(j), j = j + 1$
End *while*
Output: the ω with minimal objective-function value and $Pop_\omega(j)$

Table 2
CCEA for training CBRB.

Input: Initial population $Pop_{\omega}, Pop_T, Pop_{LN}$, Set $i = 0, j = 0$
while $i < I$
  Training ω using CSA based method until $j = G$. The best individual is used for training two BRBs
  Training B_T using another CSA based method until $j = G$. The best individual is used for training ω and BRB_LN
  Training B_T using the above CSA based method until $j = G$. The best individual is used for training ω and BRB_T
End *while*
Output: the B_w, B_T, B_{LN} with minimal objective-function value

In step 2, when training the parameters in BRB_T, the current best parameters in BRB_LN are needed, which means the two BRBs share the information. Step 3 is also the same as the last step. These two BRBs are trained in a cooperative way and the performance of CBRB can improve greatly.

5. CBRB prototype CDSS for diagnosing LNM

In this section, the proposed CBRB CDSS prototype is utilized for LNM diagnosis. It consists of two parts: (1) constructing each BRB model in CBRB system and (2) the training process for optimizing CBRB.

5.1. Generating the initial belief rules in CBRB

5.1.1. Referential points for the antecedents and consequence

In CBRB, both tumor factors and lymph node factors are used for LNM diagnosis. According to the clinical domain knowledge, the Tumor Longest Diameter (named as *TLD*) and Tumor Thickness (named as *TT*) are the two important diagnostic factors in tumor, while the size of Lymph Node (LN) (named as *LNsize*) and the number of LN (named as *LNnumber*) are the two important diagnostic factors in lymph node. Therefore, they are chosen as the medical antecedent attributes, which are described as follows:

(1) Tumor Longest diameter: The longest diameter of tumor is measured at the axial CT images.
(2) Tumor thickness: The maximal thickness of tumor is measured at the axial CT images [29].
(3) The number of LN: The number of all visible gastric regional lymph nodes in CT images by groups is counted.
(4) The size of LN: The short axis of the largest lymph node detected in CT images is measured.

Two BRBs in CBRB are constructed for tumor and LN factors, respectively. To represent the clinical expert knowledge in each BRB, the reference points should be provided. The set of the number and value of the referential points are based on the clinical domain knowledge [30,31]. Since there are two BRBs in CBRB, two groups of the referential points are given, which is described as follows.

For *TLD* in BRB_T, eleven referential points are used, including zero (Z), super small (SS), very small (VS), small (S), a little small (LS), middle (M), a little large (LL), large (L), very large (VL), greatly large (GL), and super larger (SL), which is:

$$T_1 \in \{Z_T, SS_T, VS_T, S_T, LS_T, M_T, LL_T, L_T, VL_T, GL_L, SL_T\} \tag{28}$$

For *TT* in BRB_T, nine referential points are used, and they are zero (Z), very small (VS), small (S), a little small (LS), middle (M), a little larger (LL), large (L), very large (VL), and greatly large (GL), that is:

$$T_2 \in \{Z_T, VS_T, S_T, LS_T, M_T, LL_T, L_T, VL_T, GL_L\} \tag{29}$$

There are four referential points in BRB_T, which is shown as follows:

Table 3
The referential points of *TLD* in BRB_T.

Linguistic terms	Z	SS	VS	S	LS	M	LL	L	VL	GL	SL
Numerical values	0	26	30	36	40	45	50	60	65	80	130

Table 4
The referential points of *TT* in BRB_T.

Linguistic terms	Z	VS	S	LS	M	LL	L	VL	GL
Numerical values	0	9	11	12	13	14	16	20	50

$$D_T = (D_T^0, D_T^1, D_T^2, D_T^3) \tag{30}$$

Meanwhile, all the referential points need to be quantified. Tables 3 and 4 shows the quantified value of *TLD* and *TT*, respectively.

Then the referential points for LN in BRB_LN are presented. The referential points of *LNsize* are the same as *TLD*, while the referential points of *LNnumber* are the same as *TT*. They are represented as:

$$LN_1 \in \{Z_{LN}, SS_{LN}, VS_{LN}, S_{LN}, LS_{LN}, M_{LN}, LL_{LN}, L_{LN}, VL_{LN}, GL_{LN}, SL_{LN}\} \tag{31}$$

$$LN_2 \in \{Z_{LN}, VS_{LN}, S_{LN}, LS_{LN}, M_{LN}, LL_{LN}, L_{LN}, VL_{LN}, GL_{LN}\} \tag{32}$$

There are also four referential points in output for BRB_LN, that is:

$$D_{LN} = (D_{LN}^0, D_{LN}^1, D_{LN}^2, D_{LN}^3) \tag{33}$$

The corresponding quantified values are shown in Tables 4 and 5.

The final outputs are obtained by integrating the output of BRB_T and BRB_LN, which also have four referential points, that is:

$$D = (D^0, D^1, D^2, D^3) \tag{34}$$

5.1.2. Constructing the rule base

Based on the above analysis, two BRBs in CBRB are generated. A belief rule in BRB_T is described as:

$$R_k^T: \text{ IF } TLD \text{ is } T_1^k \text{ AND } TT \text{ is } T_2^k$$
$$\text{THEN stage is } \{(D_T^0, \beta_{T,k}^0), (D_T^1, \beta_{T,k}^1), (D_T^2, \beta_{T,k}^2), (D_T^3, \beta_{T,k}^3)\}, \left(\sum_{i=0}^{3} \beta_{Tk}^i \leqslant 1\right)$$
$$\text{with a rule weight } \theta_k^T (k = 1, 2, \ldots, 99) \text{ and attribute weight } \delta_1^T, \delta_2^T \tag{35}$$

Table 5
The referential points of *LNsize* in BRB_LN.

Linguistic terms	Z	SS	VS	S	LS	M	LL	L	VL	GL	SL
Numerical values	0	5	6	7	8	9	10	12	17	21	50

where T_k^1 and T_k^2 are the referential values in BRB_T. As there are eleven terms in *TLD* and nine terms in *TT*, 99 rules are generated totally in BRB_T after combining the two factors. Similarly, the belief rule in BRB_LN is represented as:

$$R_k^{LN}: \text{ IF } LNsize \text{ is } LN_1^k \text{ AND } LNnumber \text{ is } LN_2^k$$
$$\text{THEN stage is } \{(D_{LN}^0, \beta_{LN,k}^0), (D_{LN}^1, \beta_{LN,k}^1), (D_{LN}^2, \beta_{LN,k}^2), (D_{LN}^3, \beta_{LN,k}^3)\},$$
$$\left(\sum_{i=0}^{3}\beta_{LN,k}^i \leqslant 1\right)$$
with a rule weight $\theta_k^{LN}(k = 1, 2, \ldots, 99)$ and attribute weight
$$\delta_1^{LN}, \delta_2^{LN} \tag{36}$$

where LN_k^1 and LN_k^2 are the referential values in BRB_LN. The number of generated rules is also 99. In CBRB, it is assumed that all the belief rules have equal rule weight and all the antecedent attributes have equal weight. Meanwhile, the initial belief degrees which are assigned to each stage are based on the expert knowledge.

After obtaining the output of two BRBs, the final output D can be obtained by using the ER approach. This is more informative as ER can provide a distributed clinical conclusion, which can show an overall view of the prediction result.

5.2. CCEA for training CBRB

The parameters in two BRBs and weight coefficient are independent, while the objective function for them is the same. So they have to be optimized iteratively. The parameters in the two BRBs are needed to be optimized are shown in Eq. (15), and the weight coefficient ω shown in Eq. (16) should be trained.

The inputs for BRB_T are *TLD* and *TT*, that is

$$T = (TLD, TT) \tag{37}$$

and the inputs for BRB_LN are

$$L = (LNsize, LNnumber) \tag{38}$$

The training process is outlined as follows.

Step 1: Set initial parameters.

Initial populations are generated for two BRBs and weight coefficient. To fully utilize the clinical domain knowledge, the initial belief degrees in BRB_T and BRB_LN are generated according to the expert knowledge. Meanwhile, to simplify the calculation process, the *fminmax* function which is from the optimization toolbox in MATLAB is used to generate the individuals. Note that different generated individuals are considered as the individuals in a population. As the number of parameters in ω is small, the initial population is generated randomly. In each iteration, the maximum number of evolutionary generation is considered as 20. The number of iteration is 5, which means the total number of evolutionary generation is 100.

Step 2: Transforming the input data.

Since the input data is quantitative, a rule-based transformation technique is utilized [32]. For BRB_T, the input T is transformed and represented in terms of the referential values which are defined in Tables 3 and 4. Meanwhile, the input L in BRB_LN is also transformed and represented according to Tables 5 and 6.

Step 3: Combining the outputs of the two single BRBs.

Table 6
The referential points of *LNnumber* in BRB_LN.

Linguistic terms	Z	VS	S	LS	M	LL	L	VL	GL
Numerical values	0	5	7	9	12	16	20	30	50

After getting the output of each BRB and corresponding weight coefficient, the final output is obtained by using the ER analytical algorithm [23].

Step 4: Calculating the predicting error rate.

After obtaining the belief degrees for each stage, the predicting error rate is calculated using Eq. (21).

Step 5: Updating all the parameters.

The proposed CCEA is used for updating the parameters in two BRBs and weight coefficient. Step 2–Step 5 are executed iteratively until the maximum number of iteration is achieved. To show the performance of CBRB and the proposed CCEA, a case study is given in the following section.

6. An experimental case study

In this section, CBRB is validated using a set of real patient data. Just as in [9], the 2-cross-validation approach is used. To better analyze the performance, Confusion Matrix (CM) is utilized to show the results. In this matrix, each column represents the instances in a predicted stage, and each row represents the instances in an actual stage. In CM, two evaluated indexes are used, that is: (1) User Accuracy (UA); (2) Procedure Accuracy (PA). UA represents the ratio of the number of corrected predicted stage to the number of test sample in this stage, while PA is the ratio of the number of the correct predicted stage to the number of a part of test data whose result is this stage.

In this section, the real patient data set is described at first. Then CBRB is validated. The new CCEA based optimization method is validated in the following subsection. Finally, a comparative study is shown.

Table 7
Patient characteristics.

Clinic pathological features	Value
Number of patients	255
Average age(y)	57(29–85)
Ratio of men to women	185:69
Lymph node metastasis	
Stage 0	46(18.1%)
Stage 1	49(19.2%)
Stage 2	88(34.5%)
Stage 3	72(28.2%)

Table 8
Feature description of data.

Patient data	Stage 0	Stage 1	Stage 2	Stage 3
Tumor longest diameter	35.54 ± 24.46	48.7 ± 81.3	46.72 ± 74.38	56.36 ± 53.14
Tumor thickness	6.98 ± 5.08	13.15 ± 7.85	15.34 ± 11.66	21.72 ± 23.28
The size of LN (mm)	6.3 ± 18.4	9.6 ± 43	9.7 ± 23	10.7 ± 32
The number of LN	6.4 ± 16.0	8.5 ± 32	11.9 ± 37	13.9 ± 41

Table 9
Confusion matrix of CBRB.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	44	5	0	1	50	88.00
Stage 1	3	48	6	3	60	80.00
Stage 2	2	4	63	7	76	82.89
Stage 3	0	0	5	64	69	92.75
Total	49	57	74	75	255	
PA (%)	89.80	84.21	85.14	85.33	85.88	

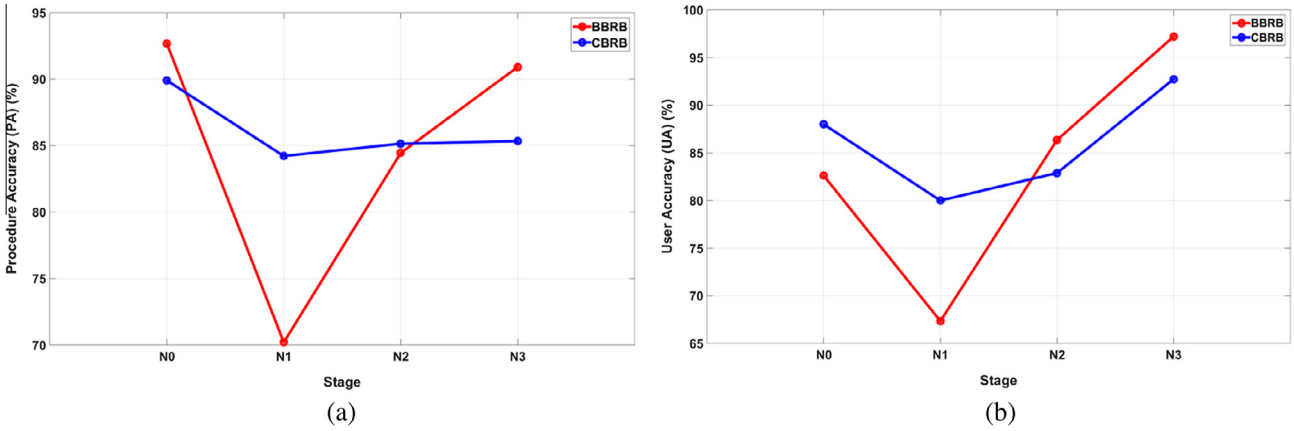

Fig. 3. PA and UA for CBRB and BBRB. (a) is the PA and (b) is the UA.

6.1. Data set description

255 real patient data are utilized, which are provided by Beijing Cancer Hospital (Beijing, PR China). Note that the informed consent from all selected patients was obtained prior to the routine clinical course. Table 7 shows the patient characteristics, and Table 8 shows the details of the four medical antecedent attributes, which shows the average value and changing range.

6.2. Validation of CBRB

In this experiment, the proposed CCEA is used for optimizing CBRB. Meanwhile, the experimental result of CBRB which was proposed in [9] is also shown. Table 9 shows the corresponding CM. It can be seen that the overall correct rate is 85.88%, while the BBRB in [9] is lower. Fig. 3 shows the PA and UA of CBRB and BBRB. The results show that CBRB is more effective than CBRB for LNM diagnosis.

Table 10
Confusion matrix of CBRB-pre.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	35	12	2	1	50	70.00
Stage 1	4	35	20	1	60	58.33
Stage 2	0	4	61	11	76	80.826
Stage 3	0	0	4	65	69	94.20
Total	39	51	87	78	255	
PA (%)	89.74	68.63	70.11	83.33	76.86	

6.3. Validation of CCEA

To show the effectiveness of the proposed CCEA based training algorithm (represented as CBRB-CCEA), we compare it with the diagnostic performance in test set before training (represented as CBRB-pre). The initial belief degree and other initial parameters are the same.

Table 11
Confusion matrix of CANN.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	44	4	2	0	50	88.00
Stage 1	6	26	25	3	60	43.33
Stage 2	0	5	65	6	76	85.53
Stage 3	0	0	7	62	69	89.86
Total	50	35	99	71	255	
PA (%)	88.00	74.29	65.66	87.32	77.25	

Table 12
Confusion matrix of CSVM.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	47	2	1	0	50	94.00
Stage 1	11	39	10	0	60	65.00
Stage 2	2	12	61	1	76	80.26
Stage 3	0	0	21	48	69	69.57
Total	60	53	93	49	255	
PA (%)	78.33	73.58	67.03	97.96	76.47	

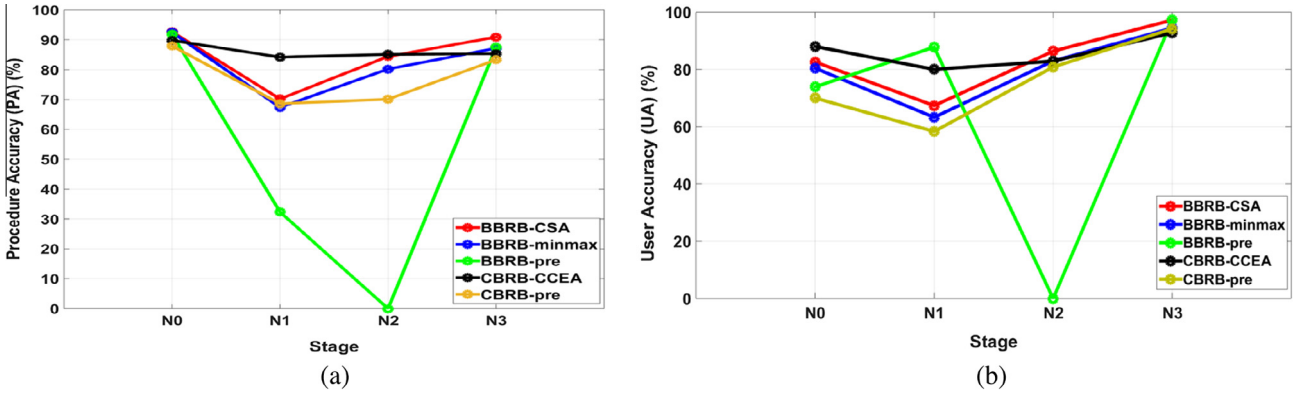

Fig. 4. PA and UA for training algorithms. (a) is the PA and (b) is the UA.

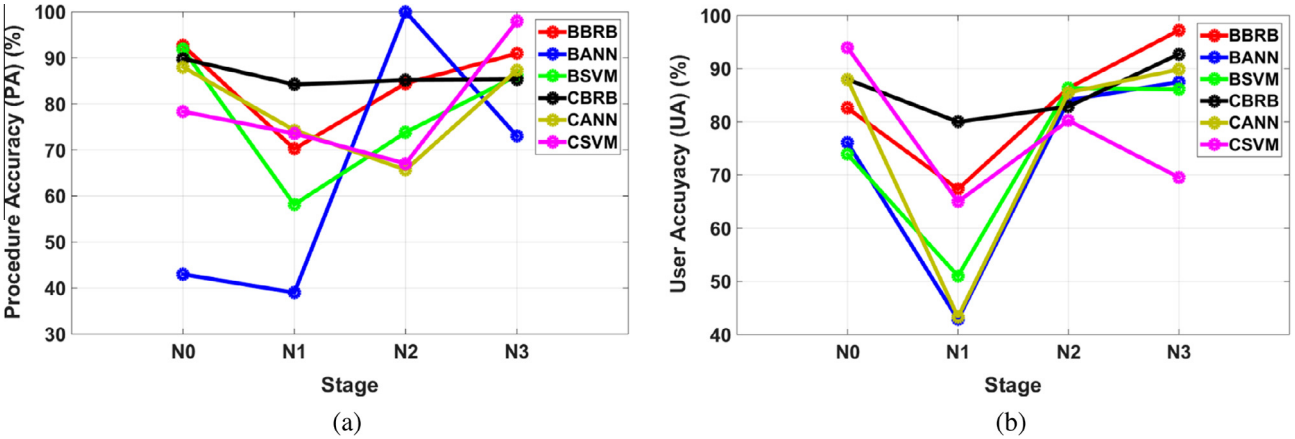

Fig. 5. PA and UA of the six models. (a) is the PA and (b) is the UA.

Table 10 shows the CM of CBRB-pre. The overall correct rate is 76.86%, which is greatly lower than CBRB-CCEA. Fig. 4 shows the PA and UA. Meanwhile, three conditions for BBRB in [9] are also compared, that is: (1) CSA based training algorithm (represented as BBRB-CSA). (2) Minmax based training algorithm (represented as BBRB-minmax). (3) The diagnostic performance in test set before training (represented as BBRB-pre). According to Fig. 4, the performance after CCEA based training is better in most cases. It demonstrates that CCEA is effective for improving the prediction results. It is also shown that it is very necessary to train CBRB.

6.4. Comparative study

Since SVM and ANN have been applied to diagnose LNM, a comparative study is presented in the subsection. To make fair comparisons, the same cooperative systems for SVM (represented as CSVM) and ANN (represented as CANN) are developed. The final output is also combined using the ER approach. For CSVM, the type of SVM is LibSVM2.91 [33] and the kernel function is Radial Basis Function (RBF) which the regularization and kernel parameters are set to $\{2^{-3}, 2^{-2}, \ldots, 2^{10}\}$. The lowest error rate is considered as the output. The feed forward neural network [34,35] in MATLAB tool box is used for CANN. Each ANN has a single hidden layer and the number of node is 5. In addition, the best result in 10 different runs is taken as the predicting result. Training and testing data in 2-fold cross validation is same. Meanwhile, BBRB, Bi-level ANN (BANN) and Bi-level SVM (BSVM) in [9] are also presented to make a comparison.

Tables 11 and 12 shows the CM of CANN and CSVM. The overall correct rate is 77.25%, 76.47%, respectively. These results are greatly lower than CBRB. Fig. 5 shows the PA and UA of the six models. In most cases, CBRB performs the best. This may be due to that the expert knowledge is embedded into the CBRB and more factors are introduced.

Compared with CANN can CSVM, CBRB is an open box system whose internal structure and parameters are transparent to the users. However, CBRB can have more parameters to be trained. This could be a cause of concern in dynamic training for real time diagnosis, although an iterative training scheme can be designed for such purposes [25]. Once CBRB is trained, the updating and prediction is easy to execute.

7. Conclusions

In this paper, a Cooperative Belief Rule Based (CBRB) CDSS prototype was proposed for LNM diagnosis in gastric cancer. CBRB consists of two independent BRBs and the final output is obtained by using the ER approach. In addition, a corresponding CCEA based method was proposed for training CBRB. By utilizing both the tumor and lymph node features, CBRB can obtain better diagnostic performance. A comparative case study on current CDSSs showed that CBRB performed better than other methods, needless to say that a CDSS is an open box system which is easy for expert validation and scrutiny.

Besides tumor and lymph node features, there are other features. It is hoped that more useful features can be introduced into a CBRB model, and more accurate diagnostic results can be obtained. In addition, the proposed CCEA was used for training two BRBs in this paper. In fact, the new training algorithm is a general method which can optimize multiple BRBs, which is also considered in the future work.

Acknowledgments

The authors would like to thank their collaborators of Beijing Cancer Hospital for their suggestions and impressive work and thank Professor Dong-Ling Xu from Manchester Business School, The University of Manchester for her suggestive comments. This research was supported in part the National Basic Research Program (973 Program) of China (No. 2013CB329402), in part by the National Natural Science Foundation of China (Nos. 61173090, 61173092, 61303119, 61271302, 61272282), in part by the National Research Foundation for the Doctoral Program of Higher Education of China (No. 20110203110006), in part by the Fund for Foreign Scholars in University Research and Teaching Programs (the 111 Project) (Grant No. B07048), and in part by the Program for Cheung Kong Scholars and Innovative Research Team in University (No. IRT1170).

References

[1] D.M. Parkin et al., Global cancer statistics, Can. J. Clinic. 55 (2005) 74–108.
[2] D. Cunningham, Y.J. Chua, East meets west in the treatment of gastric cancer, New Eng. J. Med. 357 (2007) 1863–1865.
[3] K. Ohkuma, K. Hiramatsu, Computed tomography for stage of gastric cancer, Jap. J. Can. Clin. 32 (1986) 1176–1179.
[4] D.D. Smith et al., Impact of total lymph node count on stage and survival after gastrectomy for gastric cancer: data from a large US-population database, J. Clin. Oncol. 23 (2005) 7114–7124.
[5] C.W. Wu et al., Nodal dissection for patients with gastric cancer: a randomized controlled trial, Lancet Uncoil. 7 (2006) 309–315.
[6] Zhi-Guo Zhou, Fang Liu, Li-Cheng Jiao, Zhi-Long Wang, Xiao-Peng Zhang, Xiao-Dong Wang, Xiao-Peng Zhang, An evidential reasoning based model for diagnosis of lymph node metastasis in gastric cancer, BMC Med. Inform. Dec. Making 13 (2013) 123.
[7] H. Elfriede et al., Artificial neural network for prediction of lymph node metastases in gastric cancer: a phase II diagnostic study, Ann. Surg. Oncol. 11 (2004) 506–511.

[8] X.P. Zhang et al., Support vector machine model for diagnosis of lymph node metastasis in gastric cancer with multi-detector computed tomography: a preliminary study, BMC Can. 11 (2011) 1471–2407.
[9] Zhi-Guo Zhou, Fang Liu, Li-Cheng Jiao, Zhi-Jie Zhou, Jian-Bo Yang, Mao-Guo Gong, Xiao-Peng Zhang, A bi-level belief rule based decision support system for diagnosis of lymph node metastasis in gastric cancer, Knowl.-Based Syst. 54 (2013) 128–136.
[10] J.B. Yang et al., Belief rule-base inference methodology using the evidential reasoning approach–RIMER, IEEE Trans. Syst., Man, Cybernet.-Part A: Syst. Hum. 36 (2006) 266–285.
[11] J.B. Yang, P. Sen, A general multi-level evaluation process for hybrid MADM with uncertainty, IEEE Trans. Syst., Man, Cybernet.-Part A: Syst. Hum. 24 (1994) 1458–1473.
[12] J.B. Yang, M.G. Singh, An evidential reasoning approach for multiple attribute decision making with uncertainty, IEEE Trans. Syst., Man, Cybernet. – Part A: Syst. Hum. 24 (1994) 1–18.
[13] J.B. Yang, D.L. Xu, On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty, IEEE Trans. Syst., Man, Cybernet. – Part A: Syst. Hum. 32 (2002) 289–304.
[14] J.B. Yang et al., Belief rule-base inference methodology using the evidential reasoning approach–RIMER, IEEE Trans. Syst., Man, Cybernet. – Part A: Syst. Hum. 36 (2006) 266–285.
[15] J.B. Yang et al., The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties, Euro. J. Operat. Res. 171 (2006) 309–343.
[16] R.M. Kwee, T.C. Kwee, Imaging in assessing lymph node status in gastric cancer, Gastric Cancer 12 (1) (2009) 6–22.
[17] Y. Fang, D.B. Zhao, J.G. Zhou, J.Q. Cai, Multivariate analysis of risk factors of lymph node metastasis in early gastric cancer, Zhonghua Wei Chang Wai Ke Za Zhi 12 (March) (2009) 130–132.
[18] L. Shen, Y. Huang, M. Sun, H. Xu, W. Wei, W. Wu, Clinicopathological features associated with lymph node metastasis in early gastric cancer: analysis of a single-institution experience in China, Can. J. Gastroenterol. 23 (2009) 353–356.
[19] C.Y. Wu, J.T. Chen, G.H. Chen, H.Z. Yeh, Lymph node metastasis in early gastric cancer: a clinicopathological analysis, Hepatogastroenterology 49 (2002) 1465–1468.
[20] A.P. Dempster, A generalization of Bayesian inference, J. Roy. Statist. Soc., Ser. B 30 (1968) 205–247.
[21] G. Shafer, A Mathematical Theory of Evidence, Princeton Univ. Press, Princeton, NJ, 1976.
[22] J.B. Yang et al., Optimal learning method for training belief rule based systems, IEEE Trans. Syst., Man, Cybernet. – Part A: Syst. Hum. 37 (2007) 569–585.
[23] Y.M. Wang et al., Environmental impact assessment using the evidential reasoning approach, Euro. J. Operat. Res. 174 (2006) 1885–1913.
[24] Z.J. Zhou et al., A sequential learning algorithm for online constructing belief-rule-based systems, Exp. Syst. Appl. 37 (2010) 1790–1799.
[25] Z.J. Zhou et al., A model for real-time failure prognosis based on hidden Markov model and belief rule base, Euro. J. Operat. Res. 207 (2010) 269–283.
[26] Z.J. Zhou et al., New model for system behavior prediction based on belief rule based systems, Inf. Sci. 180 (2010) 4843–4846.
[27] Tapabrata Ray, Xin Yao, A cooperative coevolutionary algorithm with correlation based adaptive variable partitioning, in: IEEE Congress on Evolutionary Computation, 2009, pp. 983–989.
[28] Thomas Jansen, R. Paul Wiegand, The cooperative coevolutionary (1 + 1) EA, Evolution. Comput. 12 (4) (2004) 405–434.
[29] Zhi-Guo Zhou, Fang Liu, Li-Cheng Jiao, Zhi-Long Wang, Xiao-Peng Zhang, Xiao-Dong Wang, Xiao-Peng Zhang, An evidential reasoning based model for diagnosis of lymph node metastasis in gastric cancer 13 (2013) 123.
[30] S.P. Monig et al., Staging of gastric cancer: correlation of lymph node size and metastatic infiltration, Am. J. Roentgenol. 173 (1999) 365–367.
[31] C.W. Wu et al., Relation of number of positive lymph nodes to the prognosis of patients with primary gastric adenocarcinoma, Int. J. Gastroenterol. Hepatol. 38 (1996) 526.
[32] J.B. Yang, Rule and utility based evidential reasoning approach for multi-attribute decision analysis under uncertainties, Euro. J. Operat. Res. 131 (2001) 31–61.
[33] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines (2011).
[34] C. Bishop, Pattern Recognition and Machine Learning, Springer Press, 2006.
[35] M.T. Hagan, M. Menhaj, Training feed-forward networks with the Marquardt algorithm, IEEE Trans. Neural Netw. 5 (1999) 989–993.
[36] S. Butenko, R. Murphey, P.M. Pardalos, Recent Developments in Cooperative Control and Optimization, Springer Science & Business Media, 2003.