Applied Soft Computing 11 (2011) 632–644

Contents lists available at ScienceDirect

Applied Soft Computing

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

A hybrid model combining case-based reasoning and fuzzy decision tree for medical data classification

Chin-Yuan Fan [a], Pei-Chann Chang [b,*], Jyun-Jie Lin [b], J.C. Hsieh [b]

[a] Department of Business Innovation & Technology Management, MingDao University, Changhua, Taiwan, ROC
[b] Department of Information Management, Yuan Ze University, Taoyuan 32026, Taiwan, ROC

ARTICLE INFO

Article history:
Received 18 July 2008
Received in revised form 8 December 2009
Accepted 12 December 2009
Available online 22 December 2009

Keywords:
Hybrid intelligence
Fuzzy decision tree
Case-based reasoning
Genetic algorithm
Breast cancer diagnosis
Liver disorders classification

ABSTRACT

In this research, a hybrid model is developed by integrating a case-based data clustering method and a fuzzy decision tree for medical data classification. Two datasets from UCI Machine Learning Repository, i.e., liver disorders dataset and Breast Cancer Wisconsin (Diagnosis), are employed for benchmark test. Initially a case-based clustering method is applied to preprocess the dataset thus a more homogeneous data within each cluster will be attainted. A fuzzy decision tree is then applied to the data in each cluster and genetic algorithms (GAs) are further applied to construct a decision-making system based on the selected features and diseases identified. Finally, a set of fuzzy decision rules is generated for each cluster. As a result, the FDT model can accurately react to the test data by the inductions derived from the case-based fuzzy decision tree. The average forecasting accuracy for breast cancer of CBFDT model is 98.4% and for liver disorders is 81.6%. The accuracy of the hybrid model is the highest among those models compared. The hybrid model can produce accurate but also comprehensible decision rules that could potentially help medical doctors to extract effective conclusions in medical diagnosis.

1. Introduction

Recently, the incorporation of computational intelligence in medical diagnosis is a new tendency and with a large number of medical applications. Many of the medical diagnosis procedures can be categorized as intelligent data classification tasks. These classification procedures can be divided into two types, with regard to the number of categories that each time is classified. The first classification type separates the data between only two classes (known as binary classification or two-class task), and the second type classifies the data between more than two classes (multi-class task). For example, there are methods for intelligent classification that handle efficiently the two-class task such as the Ada Boost and the support vector machines. Any multi-class problem can be substituted by more than one or two-class problems. Such an approach is to build independent classification rules for each of the classes and then run these competitive rules simultaneously [5].

In medical area, many researchers have tried to use different methods to improve the accuracy of data classification. Methods with better classification accuracy will provide more sufficient information to identify the potential patients and to improve the diagnosis accuracy. Meta-heuristic algorithms (like genetic algorithms, Particle Swarm Optimizations, and Tabu Search) and data mining tools (Neural Network and decision tree) have been applied in this area and have derived significant results as described in Ref. [5]. Chang and Lilly [52] employ hybrid heuristics in breast cancer classification and achieve more than 90% accuracy rate. Other researchers as in Refs. [46,42] also apply hybrid Neural Network and Boolean rules to derive results which exhibiting good performance and reduced number of rules with relevant input variables. However, in liver disorders classification [8–10], the classification accuracy of current methods are still low and insignificant enough to be adopted in practical applications.

Aside from other traditional classification problems, medical dataset classification problems are further applied in future diagnosis. Therefore, patients or doctors not only need to know the answer (classification result), but also need to know the symptoms that derive this answer. Neural Networks [46,42] and linear programming models [26] have been reported and these models almost obtain high classification accuracy rate. However, their decision process is essentially a black box, with no explanation as to how they were attained. Hybrid heuristic methods like GA or Neural Networks combine with fuzzy rules will handle this problem caused by black box approaches, but they still cannot identify which input factors are more significant than the others.

In this research, our contribution is to develop a hybrid model for medical data classification in two medical domains: breast cancer diagnosis and the classification of a liver disorder diagnosis. This hybrid model combines the soft computing techniques including a

* Corresponding author. Tel.: +886 936101320; fax: +886 34638884.
E-mail address: iepchang@saturn.yzu.edu.tw (P.-C. Chang).

1568-4946/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2009.12.023

decision tree tool ID3 (Iterative Dichotomizer 3) [22], fuzzy theory and genetic algorithms. The proposed model is able to classify the breast cancer and liver disorder data more precisely and offer medical doctor a better information platform during the diagnosis of a cancer or liver disorder patient. The medical data classification is conducted by extracting and analyzing available data with an appropriate clustering procedure and fuzzy decision tree. The clustering procedure collects groups of similar data in terms of medical data profile (breast cancer diagnosis or the classification of a liver disorder diagnosis) and weight of each input variables. Therefore, the importance of each factor (input variable) will be identified through this procedure. In addition, the fuzzy decision tree tends to discover interpretable rules between these input variables with each cluster and the classification of breast cancer or liver disorder diagnosis.

This paper is organized as follows. A literature review for medical classification problems is introduced in Section 2. Section 3 detailing the method and algorithm of a case-based fuzzy decision tree model for medical classification problems, while Section 4 presents the experimental results. Section 5 provides the conclusions and future directions of researches.

2. Literature review

Medical database classification is a kind of complex optimization problem, and this problem not just trace the optimal solution, but need to provide diagnosis aid accurately. Many researchers have been applying different kinds of soft computing tools in solving this problem. This section reviews the literature covered in the areas of Neural Network; fuzzy classifier and data mining tools respectively.

2.1. Artificial neural network

Researches in the area of using ANNs for medical classification or diagnosis purposes have caught a lot of attentions before few years ago [6,11,8,39,34,55,37]. ANN research aims to provide a filter that distinguishes the cases which do not have cancer; therefore, reducing the cost of medication, and helping the doctors to focus on real patients. Previous research in this area has been undertaken by various researchers. Wu et al. [55] used an ANN to learn from 133 instances containing each 43 mammographic features rated between 0 and 10 by mammographer. The ANN was trained using the back-propagation (BPN) algorithm using 10 hidden nodes. A single output node was trained to produce 1 for malignancy and 0 for benign. The performance of the ANNs was found to be competitive to the domain expert and after a considerable amount of feature selection; the performance of the ANNs significantly outperformed the domain expert.

Another use of BPN was undertaken by Floyd et al. [6] who used eight input parameters; these are mass size and margin, asymmetric density, architectural distortion, calcification number, morphology, density and distribution. After extensive experiments with BPN over their limited dataset of 260 cases, they achieved a classification accuracy of 50%. In another study, Setiono used his rule extraction from ANNs algorithm [43] to extract useful rules that can predict breast cancer from the Wisconsin dataset. He needed to train an ANN first using BPN to achieve an accuracy level of approximately 94% on the test data. After applying his rule extraction technique, the accuracy of the extracted rule set did not change. Setiono [41] used feature selection before training the ANN. The New rule sets had an average accuracy of more than 96%. This is an improvement when compared to the initial results. It is also comparable to the results of Fogel et al. [11]

In summary, all previous methods depend on the conventional BPN algorithm can easily be trapped in a local minimum and require extensive computational time to find out the best number of hidden units. The approach of Fogel et al. [11] presents a successful attempt to use evolutionary computations for solving the problem. Although this approach achieved better predictive accuracy than the others, it suffered from high computational cost and dependence on either knowing the right number of hidden units in advance or experimenting with a number of different networks to find the best number of hidden units. Therefore, the results of previous studies indicated a need for an approach that is less expensive and can find the right number of hidden units without so much interference from the user. It should be as accurate as the algorithm presented by Fogel et al. [11].

2.2. Fuzzy classifier tools

There are many fuzzy classifier tools have been used in database classification problems in last decade, most of them employ fuzzy rule based tools [16,1,33,38] or connectionist fuzzy methodologies with other Artificial Intelligence models [53,27,7] to extract fuzzy rules directly from data. All these methods show very excellent results in classification problem. Gadaras and Mikhailov [16] presents a novel fuzzy classification framework for the automatic extraction of fuzzy rules from labeled numerical data and through three medical pattern classification problems test shows this methods can trace small number of rules and a simple, fast and robust training process to get excellent results. Fernández et al. [1] using a preprocessing step in their fuzzy rule classifier system in order to deal with the class imbalance. The aim of this research is analyze the behavior of fuzzy rule based classification systems in the framework of imbalanced datasets. Pulkkinen and Koivisto [33] suppose one kinds of hybrid classification fuzzy models (FMs). This model initialized by two algorithms. Modified Gath-Geva (MGG) is used for function estimation and C4.5 for classification problems and through three-step GA-based optimization, the proposed fitness function makes the favoring of simple FMs possible. Furthermore, the rule base is made more comprehensible by reducing the number of conditions in the rules.

Fuzzy rule based tools shows very significant results in database classification; however, those methods have a number of critical disadvantages. Fuzzy rule based always do not use class labels but usually a gain/loss objective function [16]. So, many researchers try to connect fuzzy methodologies with other Artificial Intelligence models. Song et al. [53] developed a novel fuzzy linear discriminant analysis algorithm using a relaxed normalized condition in the definition of fuzzy membership function. Thus, the classification limitation from the outlier samples is effectively alleviated. Gonçalves et al. [27] suppose a new neuro-fuzzy model (Inverted Hierarchical Neuro-Fuzzy BSP System), not like traditional neuro-fuzzy system, this model can extract classification rules and denominated inverted that has been specifically created for record classification and rule extraction in databases. Lee and Wang [7] present an adaptive fuzzy detector to detect the electrocardiogram (ECG) signal, and ontological fuzzy agent is presented to process the collection of ECG signals. The required knowledge is stored in the ontology, which comprises some personal ontologies and predefined by domain experts. The experimental results indicate that the proposed method can work effectively.

2.3. Data mining techniques

Many other data mining techniques exist for medical classification [45,3,23] and for time series identification [24]. Amongst these methods, a wide variety of statistical models, such as linear discriminant analysis or logistic regression perform well on a large number of applications [17]. However, the classification accu-

racy of these models is often limited when the relationships of the input/output dataset are complex and/or non-linear [14]. In such situation, which are frequently found in real world problems, machine learning methods are more suitable for building simple and interpretable pattern classification models [24]. The most common models are [50,29]: Bayesian networks [32], Neural Networks, rough sets [47], decision trees [28] genetic algorithm classifiers [12] and support vector machine [23].

In this research, a hybrid model is developed based on two basic methods, i.e., FDT to classify (ID3 algorithm) and CBR to cluster (case-based weighted cluster algorithm). These tools are effective not only to find and describe patterns in data in order to make prediction but also to build an explicit representation of the knowledge.

2.3.1. Fuzzy decision tree

The fuzzy decision tree is similar to the standard decision tree methods (e.g., CART [9,22]) based on a recursive binary partitioning algorithm. At each node during the construction process of a fuzzy decision tree, the most stable splitting region is selected and the boundary uncertainty is estimated based on an iterative resampling algorithm. The boundary uncertainty estimate is used within the region's fuzzy membership function to direct new samples to each resulting partition with a quantified confidence. The fuzzy membership function is used to recover those samples that lie within the uncertainty of the splitting regions.

2.3.2. Clustering methods

Unlike decision trees which assign a class to an instance (supervised method), clustering procedures are used when instances are divided into natural groups or clusters (unsupervised method). There are different ways to produce these clusters. The groups may be:

1. exclusive [19], i.e., any instance belongs to only one group;
2. probabilistic or fuzzy [20], i.e., an instance belongs to each group to a certain probability or degree (membership value);
3. hierarchical [48], i.e., there is a crude division of instances into groups at the top level and each of these groups are refined further up to individual instances.

The choice between these possibilities (exclusive clusters, membership degree, and hierarchical clusters) is dictated by the future exploitation of the results. Indeed, in data mining applications, clustering is often followed by a stage where a decision tree or a set of rules is inferred to allocate each instance to the cluster in which it belongs. The clustering method used in this paper is the classic and straightforward *k*-means algorithm [18] which has been used for several decades.

The *k*-means method is simplistic but reasonably effective to carry out the training for the decision trees. It allows dividing instances into disjoint clusters from numeric attributes. Different final clusters can be found because of the random initialization of the center of classes. Thus, as it is the case with other practical clustering techniques, each final cluster center do not represent a global optimum but a local one. To increase the chance of finding the global optimum, simply repeat the whole algorithm several times with different starting points and chose the best.

It has been a new tendency to combine the soft computing (SC) technologies of NNs, fuzzy logic (FL) and genetic algorithms (GAs) to significantly improve an analysis [1,38,39,40]. In general, NNs are used for learning and curve fitting, FL is used to deal with imprecision and uncertainty, and GAs are used for search and optimization. Zadeh [45] points out, merging these technologies results in a tolerance for imprecision, uncertainty, and partial truth to achieve tractability, robustness, and low solution cost.

Table 1
Basic indices for liver disorders.

Indices	Descriptions
mcv	Mean corpuscular volume
alkphos	Alkaline phosphatase
sgpt	Alamine aminotransferases
sgot	Aspartate aminotransferase
gammagt	Gamma-glutamyl transpeptidase
drinks	Number of half-pint equivalents of alcoholic beverages drunk per day
selector	Field used to split data into two sets

This research combines several soft computing techniques such as fuzzy decision tree, case-based weighted data clustering, and genetic algorithm to develop a classification model for breast cancer and liver disorder datasets. The weighted case-based clustering algorithm is applied to cluster the data before the fuzzy decision rules are generated. A set of fuzzy decision rules is generated for each cluster, which enables us to determine the fuzzy terms of each variable. Finally, a GA is applied as an evolving tool to further fine-tune the classified result from the FDT model.

3. A case-based fuzzy decision tree

Decision tree induction is free from parametric assumptions and it generates a reasonable tree by progressively selecting attributes to branch the tree. A decision tree is a flow-chart-like structure where each node represents a test on an attribute (such as mean corpuscular volume with liver disorders database), each branch represents an outcome of the test (such as mean corpuscular volume = high) and leaf nodes represent a classification of an instance (such as liver disorders). By combining all kinds of medical features on liver disorders and Breast Cancer Wisconsin database, this research will apply a fuzzy decision tree to develop a forecasting model for generating decision rules in disease classification.

A novel model is developed by clustering and evolving fuzzy decision tree for the medical data classification problems. This classification model integrates a data clustering technique, a fuzzy decision tree (FDT), and genetic algorithms (GAs) to construct a medical classification system based on medical database. The set of historical data is divided into *n* subclusters by adopting a case-based weighted algorithm. GA is then applied to evolve the number of fuzzy terms for each input index in FDT. The framework of CBFDT is shown in Fig. 1 and it can be divided into four major steps. They are (1) screening medical database from UCI dataset; (2) clustering case library into smaller cases; (3) establishing fuzzy decision tree; and finally (4) outputting the classification results.

3.1. The selection of medical databases

Two medical dataset including liver disorders and Breast Cancer Wisconsin are selected from UCI database. Liver disorder database was support by BUPA Medical Research Ltd. Breast Cancer Wisconsin was support by Dr. William H. Wolberg et al. All these data can be found in UCI machine library database.

The liver disorders database includes seven features, i.e., mcv, alkphos, sgpt, sgot, gammagt, drinks and selector. There is a total of 345 data and the first 5 features are all blood tests thought to be sensitive to liver disorders arise from excessive alcohol consumption. Each line in the data file constitutes the record of a single male individual. The basic indices for liver disorders are described in Table 1. This dataset donate by Richard S. Forsyth et al. in 1990-05-15.

Another dataset applied in this research is Wisconsin Diagnostic Breast Cancer (WDBC) (Diagnostic). Those dataset samples arrive periodically as Dr. Wolberg reports in his clinical cases. This dataset includes 32 features and 569 datasets. Basic Indices for Wisconsin Diagnostic Breast Cancer are described in Table 2.

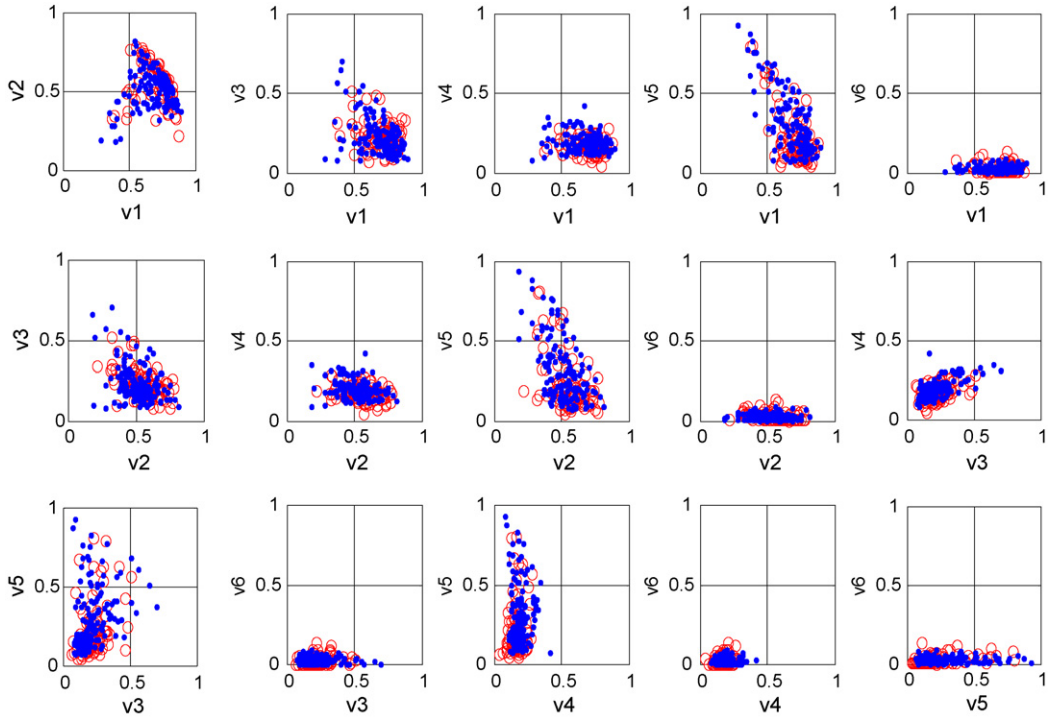

Fig. 1. A series of input data affecting the order of a liver function.

Table 2
Basic indices for Wisconsin Diagnostic Breast Cancer.

Indices		Descriptions
Radius	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	Mean of distances from center to points on the perimeter
Texture	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	Standard deviation of gray-scale values
Perimeter	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	
Area	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	
Smoothness	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	Local variation in radius lengths
Compactness	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	$\text{Perimeter}^2/\text{area} - 1.0$
Concavity	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	Severity of concave portions of the contour
Concave points	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	Number of concave portions of the contour
Symmetry	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	
Fractal dimension	1. Mean 2. Standard error 3. "Worst" or largest (mean of the three largest values)	"Coastline approximation" − 1
Diagnosis		M = malignant, B = benign

Table 3
Important input factors selected by stepwise regression analysis.

Medical database	Input factors		
Liver disorders	Original indices	Mcv Alkphos sgpt sgot Gammagt drinks	
	Through stepwise regression	Alkphos sgpt sgot gammagt	
Wisconsin Diagnostic Breast Cancer	Original indices	Radius Texture Perimeter Area Smoothness Compactness Concavity Concave points Symmetry Fractal dimension Radius: standard error	Mean, Standard error "worst"or largest (mean of the three largest values)
	Through stepwise regression choose important indices	Choose Important Indices Concavity: worst Concave points: worst Fractal dimension: mean	

The mean, standard error, and "worst" or largest (mean of the three largest values) of these features were computed for each image, resulting in 30 features. For instance, field 3 is mean radius, field 13 is radius SE, and field 23 is worst radius.

In this research, we will apply these two medical datasets to test the effectiveness and efficiency of the proposed method. In the next section, we will first describe the detail procedure of a case-based clustering method.

3.2. Best parameter setting (SRA)

A set of important indices as shown in Tables 1 and 2 which will affect the judgment of a medical symptom have been provided by the database set. These important input factors will be further selected through stepwise regression analysis (SRA) model in this research.

Fig. 1 describes a series of liver base with input data which will affect the order of a liver function. Stepwise regression analysis is applied to choose input data with significance among these indices. Important input factors selected by stepwise regression analysis can be found in Table 3.

Stepwise regression analysis is applied to determine the set of independent variables that most closely affect the dependent variable. This is accomplished by repeating the variable selection procedure. The step-by-step procedure of the SRA approach is explained in details in the following:

Step 1 Calculate the correlation coefficient (r) of every input variable ($X_1, X_2, \ldots, X_n$), i.e., technical indices, and output data (Y), i.e., trading signal. Set all numbers in a correlation matrix.

Step 2 Choose the largest number of square (r^2) from the correlation matrix (suppose that X_i is the largest one in the current stage), and derive a regression model that is $\hat{Y} = f(X_i)$; then consider the correlation of Y with other input data. Assuming X_j has statistical significance; α value is applied to consider the significance of each input variable.

Step 3 Calculate the partial F value from other input data, as in Eq. (1), and choose the largest correlation coefficient among these input variables (assume it is X_j). Then, derive another regression model $\hat{Y} = f\left(X_i, X_j\right)$ again.

$$\begin{aligned} \mathrm{SSR} &= \sum \left(\hat{Y}_i - \bar{Y}\right)^2 \\ \mathrm{SSE} &= \sum \left(\hat{Y}_i - Y_i\right)^2 \end{aligned} \tag{1}$$

$$F_j^* = \frac{\mathrm{MSR}(X_j|X_i)}{\mathrm{MSE}(X_j, X_i)} = \frac{\mathrm{SSR}/1(X_j|X_i)}{\mathrm{SSE}/(n-2)(X_j, X_i)}, \quad i \in I \tag{2}$$

Step 4 Calculate the partial F value of the original data for input variable X_j. If the value is smaller than a user defined threshold, remove it from the model since X_j is not statistically significant for the output.

Step 5 Repeat step 3 to step 4. If every input variable's partial F value is greater than the user defined threshold, stop. It means that every input value should have significant influences on output value. According to reference [35,36], if F value of a specific variable is greater than the user defined threshold, it is added to the model as a significant factor. When F value of a specific variable is smaller than a user defined threshold, it is removed from the model. The statistical software SPSS for Windows 10.0 is applied for stepwise regression analysis in this research.

3.3. Weighted clustering method

A medical case library from UCI medical library is applied to develop the weighted distance metric. A similarity measure is described in the following.

First assume a medical case library equal to $ML = \{e_1, e_2, \ldots, N\}$. Each case in the library can be identified by an index of corresponding features. In addition, each medical case has an associated action to be made for its current performance and the action is either a positive or negative decision (to judge a patient's symptom). To be exact, we use a collection of features $\{F_j\ (j = 1, \ldots, n)\}$ to represent the cases and a variable V to denote the action. The ith case e_j in the library can be represented as a $n+1$-dimensional vector, i.e., $e_i = \left(x_{i_1}, x_{i_2}, \ldots, x_{i_n}, y_i\right)$. Where x_j corresponds to the value of feature F_j $(1 \le j \le n)$ and y_i corresponds to the action $(i = 1, \ldots, N)$ to be taken and it will be defined later. Suppose that for each j $(1 \le j \le n)$

Table 4
Pseudo code for phase one.

```
Procedure weights_calculate_by_ phase one
  Input parameters: α and the learning rate η
  Initialize weight of each attribute j in each data with random values in [0,1];
  Do
    Compute Δw_j = −η ∂E/∂w_j;
    /* E(w) = 2×[Σ_pq Σ_(q<p) (SM_pq^(w)(1−SM_pq^1)+SM_pq^1(1−SM_pq^(w)))] / (N×(N−1))
    where N is the number of cases in the SL base */
    Update w_j:=w_j + Δw_j
  While not convergent;
  Update each attribute j a new weight;
End
```

a weight wj ($w_j \in [0, 1]$) has been assigned to the jth feature to indicate the importance of the feature, then for any pair of cases e_p and e_q in the library, a weighted distance metric can be defined as

$$d_{pq}^{(w)} = d^{(w)}(e_p, e_q) = \left(\sum_{j=1}^{n} w_j^2 \left(x_{pj} - x_{qj}\right)^2\right)^{1/2} = \left(\sum_{j=1}^{n} w_j^2 x_j^2\right)^{1/2} \tag{3}$$

where $x_j^2 = (x_{pj} - x_{qj})^2$. When all the weights are equal to 1 the distance metric defined above coincides with the Euclidean measure, denote by $d_{pq}^{(1)}$.

By using the weighted distance defined in Eq. (3), a similarity measure between two cases, $SM_{pq}^{(w)}$, can be defined as follows:

$$SM_{pq}^{(q)} = \frac{1}{1 + \alpha \cdot d_{pa}^{(w)}} \tag{4}$$

where α is a positive parameter. When all weighs take value 1 the similarity measure is denoted by $SM_{pq}^{(1)}$.

After introducing the weighted distance metric and the similarity measure, the weighted clustering methodology is further described in the following.

3.3.1. Phase one: finding weighted feature values from input indices

In this step, the gradient method is applied to find the weighted value from medical dataset indices and a feature evaluation function is defined. The smaller is the evaluation value, the better are the corresponding features. Thus we would like to find the weights such that the evaluation function attains its minimum. The detail procedures of phase one are described in Table 4.

3.3.2. Phase two: dividing the ML (medical library database (includes liver disorders and Wisconsin Diagnostic Breast Cancer) into several clusters

This section attempts to partition the medical case library into several clusters by using the weighted distance metric with the weights learned in previous section. Since the features are considered to be in real-value, many methods such as k-means clustering and Kohonen' self-organizing network can be used to partition the case library. However, this paper adopts a typical approach of clustering, by Shiu et al. [49] who use only the information of similarity between cases. This approach first transforms the similarity matrix to an equivalent matrix and then considers the cases being equivalent to each other as one cluster. The detail processes for phase two can be described as in Table 5.

Through a series of weighted and clustering steps, the case library will be divided into different subclusters. The data in each subcluster will have a more similar pattern in terms of the features and action of each data point. These case bases are then ready for further application in medical symptom decisions.

Table 5
Pseudo code for phase two.

```
Procedure divide_case_library_into_clusters_by_ phase two
Give a significant level (threshold) β ∈ (0,1]
  Determine the similarity matrix SM = (SM_pq^(w)) from Eqs. (3) and (4)
  Do compute SM1 = SM. SM = (s_pq) where s_pq = max_k(min(SM_pk^(w), SM_kq^(w)))
    Replace SM with SM1.
    While not (SM1 ⊂ SM)
  Determine clusters
  //Case p and case q in the same cluster if and only if s_pq ≥ β//
End
```

3.4. GAFDT forecasting model

This research combines genetic algorithms and fuzzy decision trees (GAFDT) to develop a classification model for the prediction of medical symptom decisions. The framework of GAFDT is depicted as in Fig. 2.

3.5. Data fuzzification

The fuzzy resolution concept in fuzzy set theory is applied to transform data attribute from continuous to discrete. A decision tree classification method is then further embedded to build a medical symptom decision model. In summary, ID3 decision tree will be applied in our model as a programming tool.

In fuzzy set theory, Membership function is one of the basic concepts, through this concept one will be able to process quantitative fuzzy set data, and dispose of fuzzy message. How to find an apropos membership function to approach quantitative fuzzy set data and dispose of fuzzy message becomes very important in fuzzy set theory. However, one perfect rule does not exist that can be adopted in different kinds of fuzzy set data. Researchers always consider different problems with different membership function; the most used membership function includes Triangle membership functions, Trapezoid membership functions, and Gauss membership functions. This research will adopt Triangle membership functions as our primary membership functions.

3.6. ID3 decision tree

The ID3 decision tree learning algorithm computes the information gain G based on each attribute A, and it is defined as follows:

$$G(S, A) = \text{entropy}(S) - \sum_{v \in \text{values}(A)} \frac{|S_v|}{|S|} \text{entropy}(S_v) \tag{8}$$

where S is the total input space and S_v is the subset of S for which attribute A has a value v. The Entropy (S) over classes is given by $\sum_{i=1}^{c} - p_i \log_2(p_i)$, where p_i represents the probability of class "i." The attribute with the highest information gain, says B, is chosen as the root node of the tree. Next, a new decision tree is recursively constructed over each value of B using the training subspace $S - \{S_B\}$. A leaf node or a decision-node is formed when all the instances within the available training subspace are from the same class. For detecting anomalies, the ID3 decision tree outputs binary classification decision of "0" to indicate normal and "1" to indicate anomaly class assignments to test instances.

3.7. Evolving fuzzy decision tree by genetic algorithm

Genetic algorithm will be used in this stage to improve the accuracy of FDT (fuzzy decision tree) in medical symptom decisions. Genetic algorithms will find the best number of fuzzy terms of every input data (medical feature indices), and then the fitness function

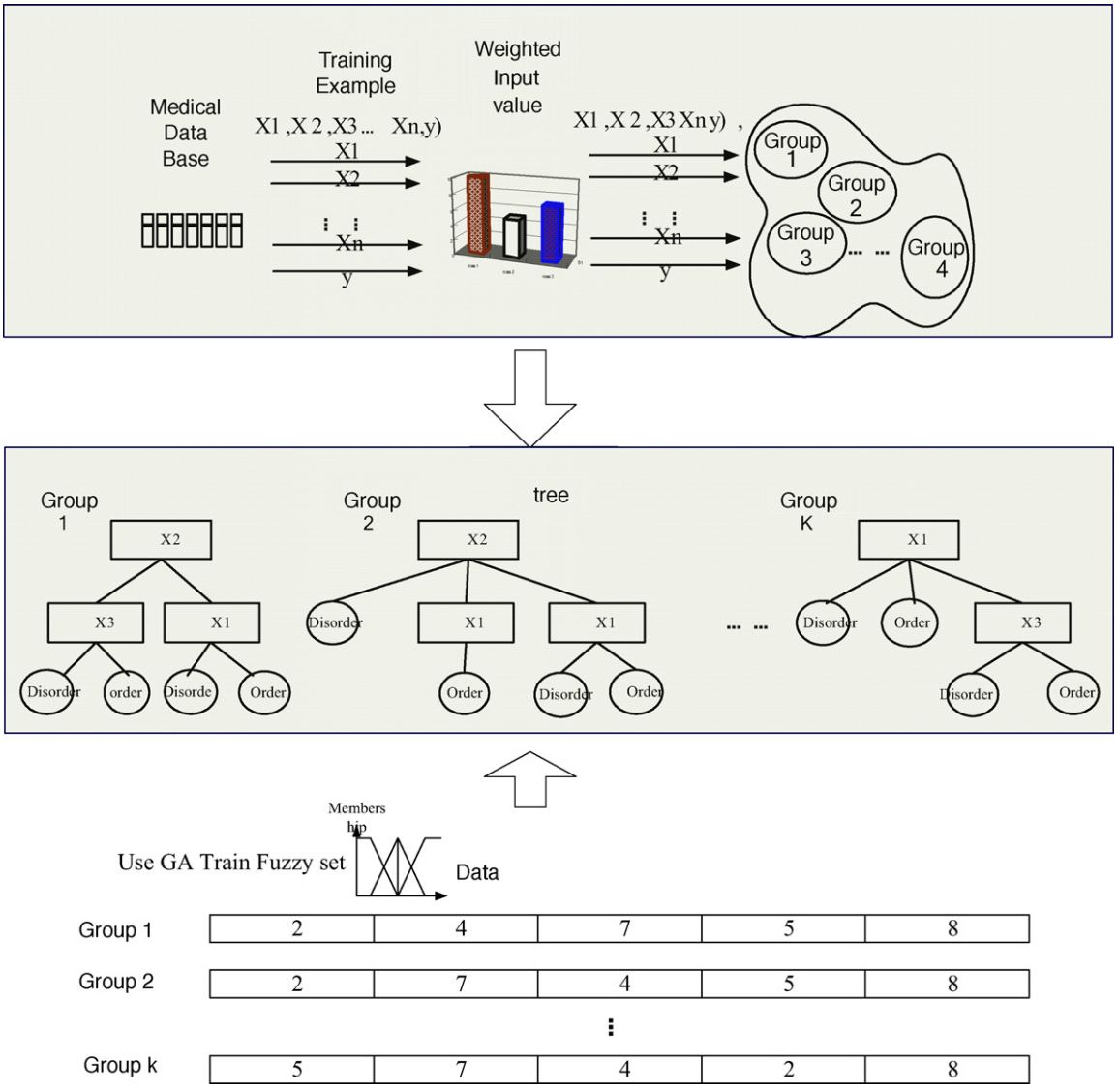

Fig. 2. A Framework of the hybrid model.

will be re-calculated after each new number of fuzzy terms. In this research, a fitness function is the classification accuracy of medical symptom decisions. Next, GA will continue the selection, crossover, and mutation. The process will iteratively repeat until the stopping criteria are satisfied. Detailed procedures of GA are described as follows.

3.7.1. Step 1. Coding

Binary code is adopted here, and for the degree of accuracy, chromosome length needs to be decided after design of experiments. In this research, we will select the four most important features from the input index and then randomly produced 12 ($4 \times 3 = 12$) genes.

3.7.2. Step 2. De-coding (generate initial solutions)

For the first step, two binaries are applied to produce gene, this step needs to transform gene from binary to decimal. Through this transformation, the system would become more workable. Then we can obtain the range of every decimal numbers. After we get the decimal range, we transform this range by Fuzzy set interval. However, binary to decimal number ranges is only between 0 and 7, fuzzy set intervals are between 2 and 9. Therefore, we add two in every transformed data.

3.7.3. Step 3. Calculate the objective value

The accuracy rate of the proposed model in each medical database will be calculated according to the following formula and the accuracy rate is also assumed to be the objective function of GA.

$$\text{accuracy rate} = \frac{\sum_{i=1}^{m} x_i}{n} \times 100\% \quad (9)$$

where, m means the total number of clustering group, x_i means the correct cases in i clustering group, n is the total number of cases. The accuracy rate is calculated by the sum of total correct cases from each clustering group divide by the total number of cases.

3.7.4. Step 4. Representation and selection

The tournament method is applied in this study.

3.7.5. Step 5. Crossover

Two-point crossover method is applied here

3.7.6. Step 6. Mutation

Two-point mutation method is adopted in the research.

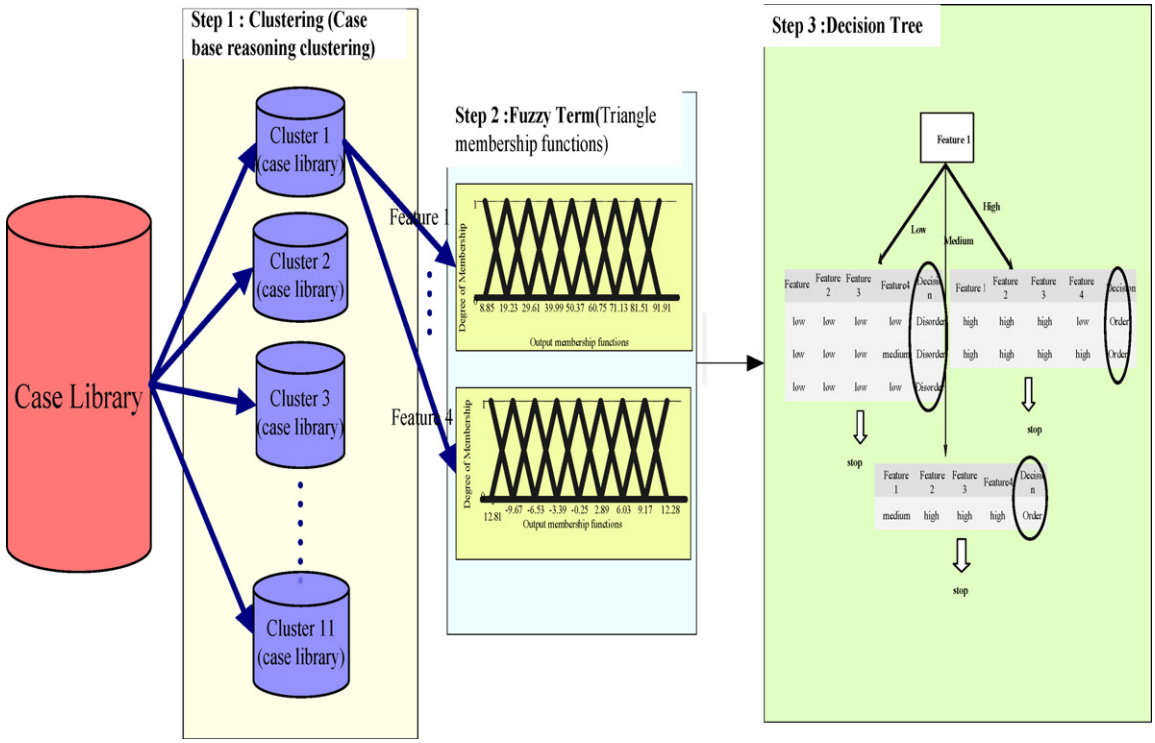

Fig. 3. A Weighted fuzzy clustering method (take liver disorder database for example).

3.7.7. Step 7. Replace

If the objective of current chromosome is better than the current best then replace it with the current one.

3.7.8. Step 8. Terminate

If the stopping criteria are satisfied, the GA process should be terminated and output the best result.

3.8. The judgment of output value

This research mainly applies evolutional fuzzy decision trees to classify the potential medical symptom by comparing the results from a medical dataset library.

4. Experimental results

According to the criteria listed in Tables 1 and 2, there are two medical databases applied to test the efficiency of the proposed model. In addition, the proposed model will be compared with other models developed earlier in the literatures. Another purpose in our research to choose the different kinds of medical database to show that the proposed model have a robust performance even under different type of medical database. Then different input factors for each case are selected according to stepwise regression analysis. Detailed procedures of CBFDT applied in these two different medical databases are explained in the following sections.

Table 6
Best parameter settings for α and β values from experimental designs.

Parameter setting	Liver disorders	Wisconsin Diagnostic Breast Cancer
α	0.6	0.6
Learning rate	0.05	0.05
β	0.18	0.25
Phase one running times	5000	5000
Phase two running times	20	20

4.1. A weighted fuzzy clustering method

Experimental design is applied to decide the best parameter setting. After the experimental tests, the parameter setting is shown in Table 6. In addition, the best number of cases for each stock is also shown in this table.

The final result of medical database after applying CBFDT model can be shown in Fig. 3. The case library is clustered into eight different smaller cases according to a weighted fuzzy clustering method. Fuzzy rules are generated by retrieving training data from each case. These fuzzy rules will be applied to generate prediction signals such as liver disorder or normal. Finally, the accuracy of such prediction can be calculated by comparing these prediction signals with the actual results.

Table 7
Parameter settings for genetic algorithm.

Factors	Liver disorders	Parameter setting	Wisconsin Diagnostic Breast Cancer	Parameter setting
Population size	5	20	5	20
Number of generation	10	100	10	100
Crossover rate	0.6	0.9	0.6	0.9
Mutation rate	0.1	0.3	0.1	0.3

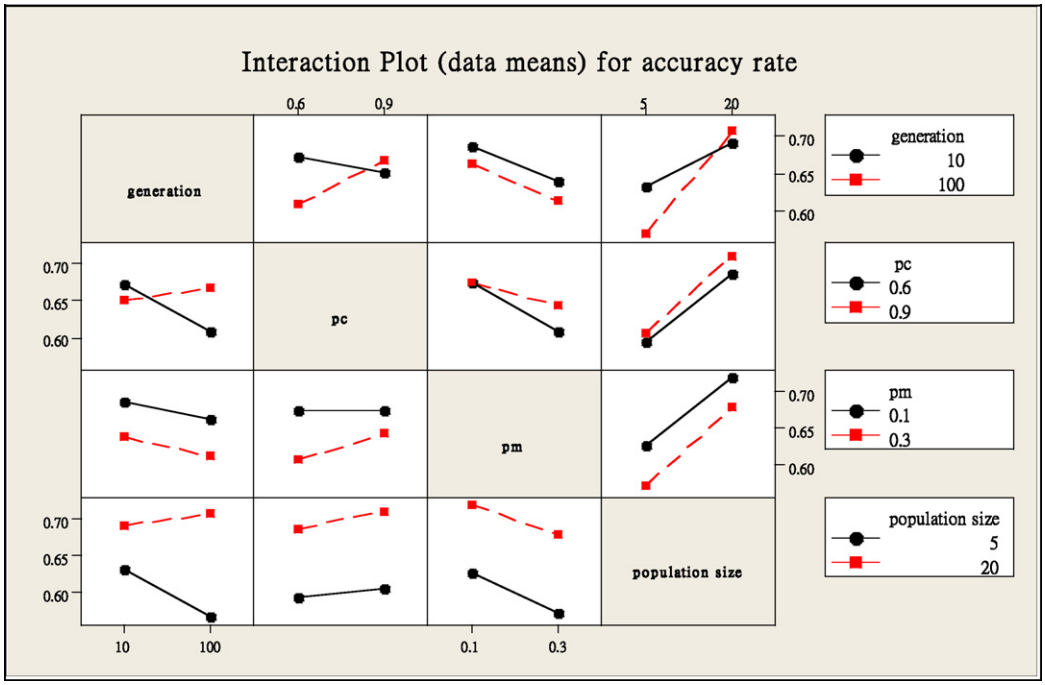

Fig. 4. Interaction plot (data means) for accuracy rate (liver disorders).

4.2. Best parameters settings for genetic algorithms

Genetic algorithms are applied to evolve the fuzzy terms of each factor in this research. Four important factors are selected in this experimental design. They are Population Size, Number of Generation, Crossover rate and Mutation rate. Tournament Selection is applied for chromosome selection from the mating pools. Two-point crossover strategy and two-point mutation strategy are used during the evolving process. Detail parameter settings for GA are shown in Table 7. The interaction plots for factor design are shown in Figs. 4 and 5 and the final parameter settings for these two databases are shown in Table 8.

ID3 decision tree is further applied to these two databases and a set of fuzzy decision trees will be generated as shown in Fig. 6. Following this decision tree from top to bottom is a set of fuzzy rules that can be generated. These fuzzy rules generated can then be applied in those new test data to make a judgment if a potential illness is detected.

4.3. Comparisons of different models

After setting up the parameters of the experiments, we take the output of CBFDT and compare with the output from traditional Classification tools. In our experimental study, medical database set (liver disorder and Wisconsin Diagnostic Breast Cancer) is first pre-processed by stepwise regression and then clustered by the case-based method. The result is then classified by Evolving fuzzy decision tree. As shown in Table 6, 60.1% classification accuracy is obtained from standard fuzzy decision tree while in stepwise regression with preprocessing and case-based method with data clustering, 90.4% (liver disorders) and 98.9% (Wisconsin Diagnostic Breast Cancer) classification accuracy is obtained. In our research, KNN, Naïve Bayes, SVM were chosen to be compared with our method. The CBFDT approach performs much better than traditional classification tools in liver disorder database, but in Wisconsin database, CBFDT shows more similar than traditional SVM method. 75% of the data were randomly chosen for training while 25% of these data is chosen for testing for these models with a total number of 500 execution times. In addition, as shown in Tables 9–11, Figs. 7 and 8, the CBFDT is also compared with other approaches developed in the literature to show the effectiveness of our approach.

Table 8
Final parameter settings for these two database.

Factors	Liver disordersLevel	Wisconsin Diagnostic Breast Cancer Level
Population size	20	20
Number of generation	100	10
Crossover rate	0.9	0.9
Mutation rate	0.1	0.1

4.4. Discussions

As observed in previous table, CBFDT outperforms other traditional methods. The reasons are: (1) a case-based clustering method does split the case library into more homogeneous and smaller cases in data-preprocessing stage. Therefore, fuzzy rules generated from each case can more sensitively react to the related disease. (2) An evolving FDT can be applied to decide the number of fuzzy terms more effectively especially when the number of data is increasingly large. As shown in Table 12, the data amounts and number of fuzzy terms show that with more number of data, more fuzzy terms are generated. To generate effective fuzzy rules, the number of fuzzy terms should be evolved through GA. As a result, the hit rate can be further improved when compares to traditional fuzzy decision methods.

In addition, data distribution is another important factor to be considered since it will affect the number of fuzzy terms clustered in the dataset. For example, sgpt is divided clearly into nine fuzzy terms as shown in Fig. 9 and the number of data in each term is small and distributed evenly. However, if sgpt is divided into three fuzzy terms as shown in Fig. 10, there will be a large number of data in each term and the fuzzy rules generated may not be able to react to the real situation. It may also lead to wrong decisions. Therefore,

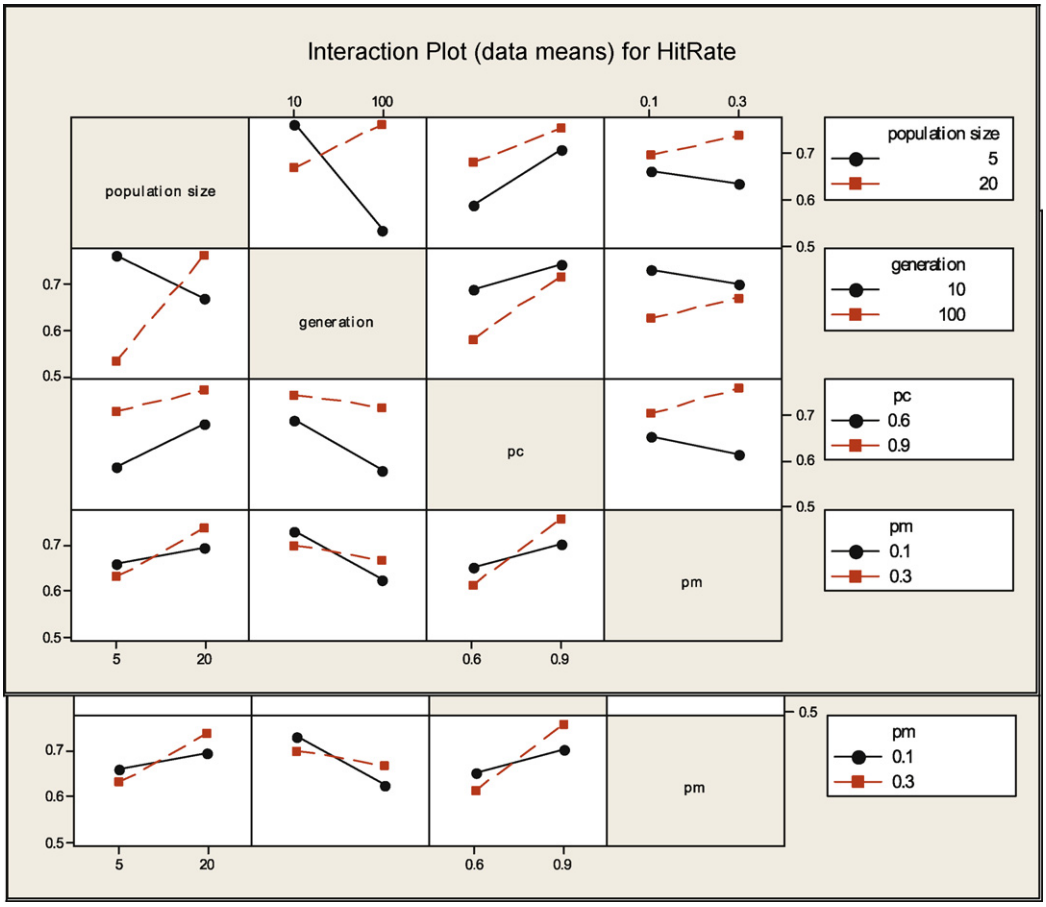

Fig. 5. Interaction plot (data means) for accuracy rate (Wisconsin Diagnostic Breast Cancer).

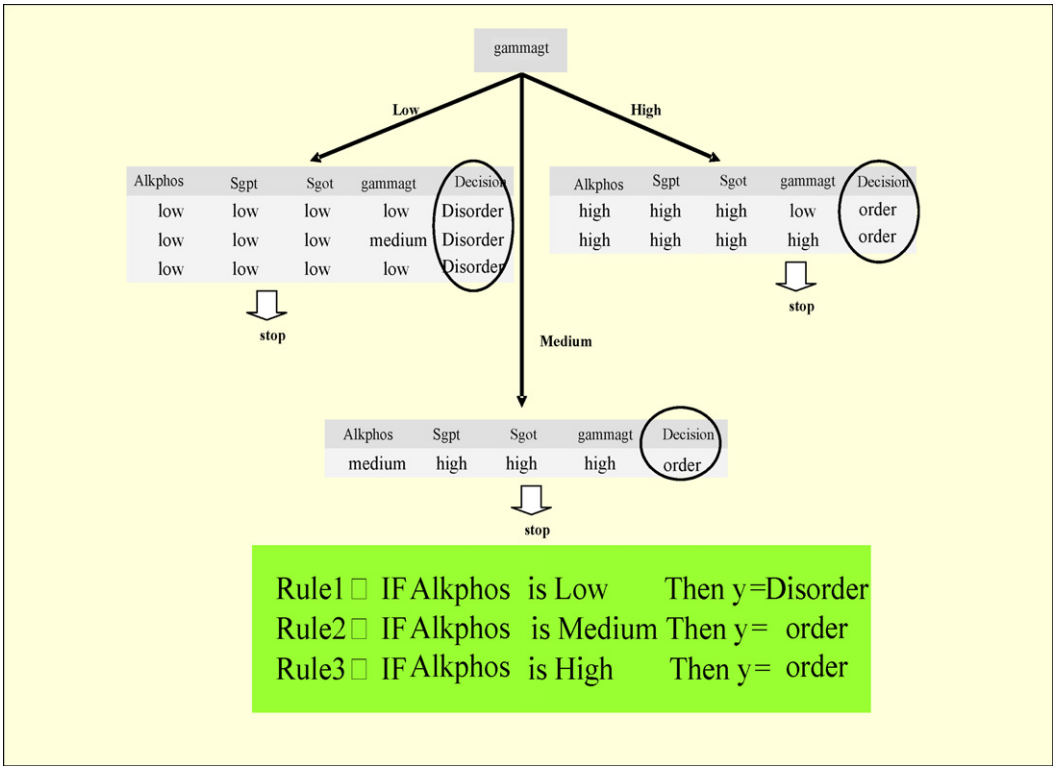

Fig. 6. A set of FDT rules generated from ID3 decision tree.

Table 9
Accuracy rate comparisons of different forecasting models in two different medical databases.

Medical database		SVM	KNN	Naïve Bayes	FDT	CBFDT
		Average diagnosis accuracy ratio (100 times)	Average diagnosis accuracy ratio (100 times)	Average diagnosis accuracy ratio (100 times)	Average diagnosis accuracy ratio (100 times)	Average diagnosis accuracy ratio (100 times)
Liver disorders	Best (100 times)	0.776	0.737	0.702	0.683	0.904
	Average (100 times)	0.693	0.611	0.592	0.601	0.853
	Lowest (100 times)	0.632	0.548	0.581	0.587	0.762
Wisconsin Diagnostic Breast Cancer	Best (100 times)	0.981	0.969	0.914	0.902	0.989
	Average (500 times)	0.932	0.898	0.876	0.862	0.927
	Lowest (500 times)	0.813	0.801	0.853	0.789	0.861

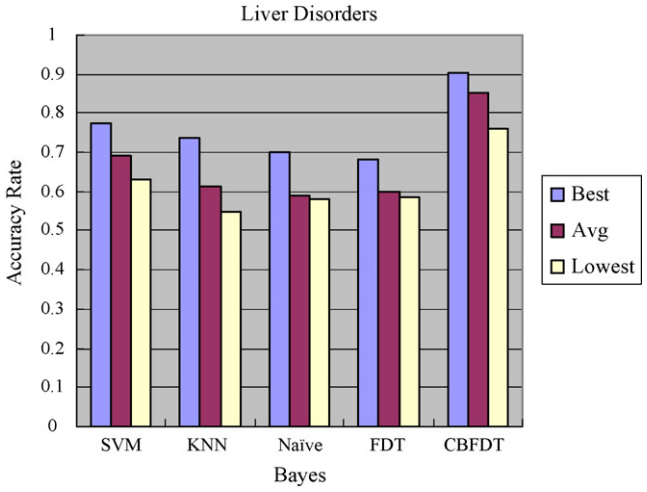

Fig. 7. Best accuracy rate comparison graph for liver disorders.

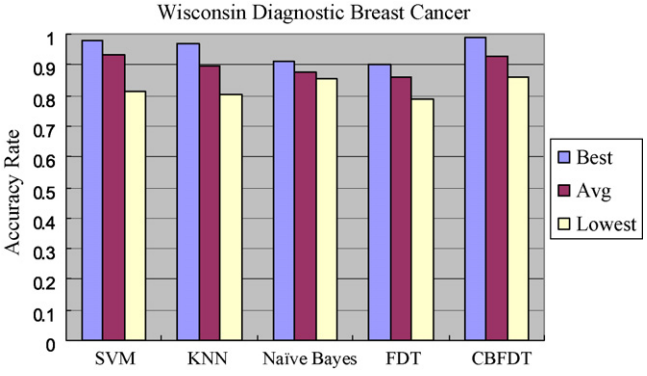

Fig. 8. Best accuracy rate comparison graph for Wisconsin Diagnostic Breast Cancer.

Table 10
Accuracy rate comparisons of CBFDT with other approaches from previous researches in liver disorder medical database.

Author (year)	Method	Classification accuracy (%) (best)
Pham et al. (2000) [10]	RULES-4	55.90
Cheung (2001) [31]	C4.5	65.59
Cheung (2001) [31]	Naïve Bayes	63.39
Cheung (2001) [31]	BNND	61.83
Cheung (2001) [31]	BNNF	61.42
Van Gestel et al. (2002) [51]	SVM with GP	69.70
Lee and Mangasarian, 2001 [54]	SSVM	70.33
Lee and Mangasarian, 2001 [54]	RSVM	74.86
Yalçın and Yıldırım (2003) [30]	MLP	73.05
Yalçın and Yıldırım (2003) [30]	PNN	42.03
Yalçın and Yıldırım (2003) [30]	GRNN	65.55
Gonçalves (2006) [27]	HNFB	73.33
Gadaras and Mikhailov (2009) [16]	Fuzzy rule classification	89.9
Our methods	CBFDT	90.40

Table 11
Accuracy rate comparisons of CBFDT with other approaches from previous researches in wisconsin diagnostic breast cancer medical database.

Author (year)	Method	Classification accuracy (%) (best)
Quinlan (1996) [21]	C4.5	94.74
Hamilton et al. (1996) [15]	RIAC	94.99
Ster and Dobnikar (1996) [2]	LDA	96.80
Bennett and Blue (1997) [25]	SVM	97.20
Nauck and Kruse (1999) [9]	NEFCLASS	95.06
Pena-Reyes and Sipper (2002) [4]	Fuzzy-GA1	97.36
Goodman et al. (2002) [13]	Optimized-LVQ	96.70
Goodman et al. (2002) [13]	Big-LVQ	96.80
Goodman et al. (2002) [13]	AIRS	97.20
Abonyi and Szeifert (2003) [44]	Supervised fuzzy clustering	95.57
Gadaras and Mikhailov (2009) [16]	Fuzzy rule classification	96.08
Our methods	CBFDT	98.90

Table 12
Accuracy rate and number of fuzzy terms for each feature in two different medical databases (take 10 trails for example).

Trail	Data amounts			
	Liver disorders		Wisconsin Diagnostic Breast Cancer	
	Fuzzy term	Accuracy rate	Fuzzy term	Accuracy rate
1	9 8 4 8	0.880546	5 4 8 9	0.869141
2	9 8 3 8	0.849829	5 2 8 9	0.947266
3	9 8 4 8	0.904437	4 9 8 9	0.923828
4	9 8 4 8	0.883959	6 8 8 9	0.923828
5	9 8 3 8	0.904437	5 2 8 9	0.90625
6	9 8 6 8	0.887372	5 4 8 9	0.869141
7	9 8 3 8	0.894198	5 4 8 9	0.835938
8	9 8 5 8	0.883959	5 2 8 9	0.927734
9	9 8 2 8	0.897611	4 9 8 9	0.923828
10	9 8 7 8	0.750853	3 2 6 5	0.989071

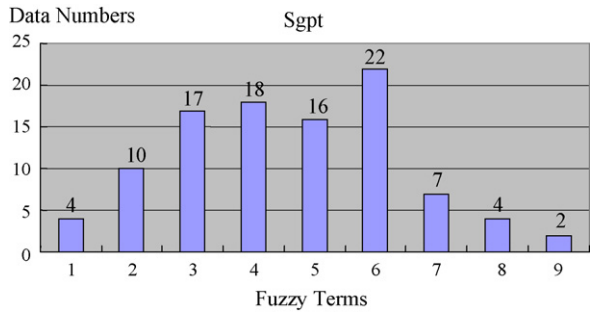

Fig. 9. sgpt splitting clearly into nine fuzzy terms.

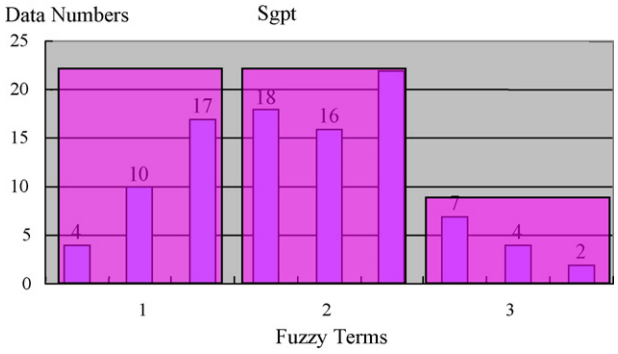

Fig. 10. sgpt fuzzy term in different fuzzy rules.

the numbers of fuzzy terms of each feature do affect the number of rules generated.

5. Conclusions

A considerable amount of research has been conducted to study the behavior of a series of medical symptoms. However, the researcher is more interested in finding potential disease factors. Therefore, we take a different approach by applying a case-based fuzzy decision tree to diagnose the Potential illness symptoms. A stepwise regression (SRA) method is applied to select the most important factors from the set of inputs. Next, a weighted clustering method is adopted to divide the case base into a smaller case. Within each case, a more homogeneous data are grouped into together. Therefore, these data react to the illness symptom detection prominently. Finally, a GA is applied to evolve the fuzzy terms of each factor in order to derive the best fuzzy decision tree from each case. Through a series of experimental tests, the CBFDT outperforms other approaches with an average accuracy rate around 99.5% in breast cancer and 85% in liver disorder, respectively. It is the highest among the literature published up to present. This model can be further applied in classification of other medical disease database to help researcher or doctors to make better decision in medical diagnosis.

References

[1] A. Fernández, M.J. Jesus, F. Herrera, On the influence of an adaptive inference system in fuzzy rule basedclassification systems for imbalanced data-sets, Expert Systems with Applications 36 (2009) 9805–9812.
[2] B. Ster, A. Dobnikar, Neural networks in medical diagnosis: comparison with other methods, in: Proceedings of the International Conference on Engineering Applications of Neural Networks (EANN'96), 1996, pp. 427–430.
[3] C.S. Lee, M.H. Wang, Ontology-based intelligent healthcare agent and its application to respiratory waveform recognition, Expert Systems with Applications 33 (3) (2007) 606–619.
[4] C.-A. Peña-Reyes, M. Sipper, A fuzzy-genetic approach to breast cancer diagnosis, Artificial Intelligence in Medicine 17 (1999) 131–155.
[5] C.-C. Bojarczuk, H.-S. Lopes, A.-A. Freitas, Genetic programming for knowledge discovery in chest-pain diagnosis, IEEE Engineering in Medicine and Biology Magazine 19 (4) (2000) 38–44.
[6] C.E. Floyd, J.Y. Lo, A.J. Yun, D.C. Sullivan, P.J. Kornguth, Prediction of breast cancer malignancy using an artificial neural network, Cancer 74 (1994) 2944–2998.
[7] C.-S. Lee, M.-H. Wang, Ontological fuzzy agent for electrocardiogram application, Expert Systems with Applications 35 (2008) 1223–1236.
[8] D. Furundzic, M. Djordjevic, A.J. Bekic, Neural networks approach to early breast cancer detection, Journal of Systems Architecture 44 (1998) 617–633.
[9] D. Nauck, R. Kruse, Obtaining interpretable fuzzy classification rules from medical data, Artificial Intelligence in Medicine 16 (1999) 149–169.
[10] D.T. Pham, S.S. Dimov, Z. Salem, Technique for selecting examples in inductive learning, in: European Symposium on Intelligent Techniques (ESIT 2000), Aachen, Germany, 2000, pp. 119–127.
[11] D.B. Fogel, E.C. Wasson, E.M. Boughton, Evolving neural networks for detecting breast cancer, Cancer Letters 96 (1) (1995) 49–53.
[12] D.E. Goldberg, Genetic Algorithm in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA, 1989.
[13] D.E. Goodman, L. Boggess, A. Watkins, Artificial immune system classification of multiple-class problems, in: Proceedings of the Artificial Neural Networks in Engineering ANNIE, 2002, pp. 179–183.
[14] E.I. Altam, G. Macro, F. Varetto, Corporate distress diagnosis: comparison using linear discriminant analysis and neural networks, Journal Banking and Finance 18 (1994) 505–529.
[15] H.J. Hamilton, N. Shan, N. Cercone, RIAC: A Rule Induction Algorithm based on Approximate Classification, Technical Report CS 96-06, University of Regina, 1996.
[16] I. Gadaras, L. Mikhailov, An interpretable fuzzy rule-based classification methodology for medical diagnosis, Artificial Intelligence in Medicine 47 (1) (2009) 25–41.
[17] J. Andrés, M. Landajo, P. Lorca, Forecasting business profitability by using classification techniques: a comparative analysis based on a Spanish case, European Journal of Operational Research 167 (2) (2005) 518–542.
[18] J.A. Hartigan, Clustering Algorithms, John Wiley, New York, 1975.
[19] J.B. MacQueen, Some methods for classification and analysis of multivariate observations, in: Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, University of California Press, Berkeley, 1967, pp. 281–297.
[20] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
[21] J.R. Quinlan, Improved use of continuous attributes in C4.5, Journal of Artificial Intelligence Research 4 (1996) 77–90.
[22] J.R. Quinlan, Induction of decision trees, Machine Learning 1 (1986) 81–106.
[23] K. Polat, S. Gunes, A. Arslan, A cascade learning system for classification of diabetes disease: generalized discriminant analysis and least square support vector machine, Expert Systems with Applications 34 (1) (2008) 482–487.
[24] K.C. Lee, S.B. Oh, An intelligent approach to time series identification by a neural network-driven decision tree classifier, Decision Support Systems 17 (1996) 183–197.
[25] K.P. Bennet, J.A. Blue, A Support Vector Machine Approach to Decision Trees, Math Report, Rensselaer Polytechnic Institute, 1997, pp. 97–100.
[26] K.P. Bennett, O.L. Mangasarian, Neural network training via linear programming, in: P.M. Pardalos (Ed.), Advances in Optimization and Parallel Computing, Elsevier, 1992, pp. 56–57.
[27] L.B. Gonçalves, M.M.B.R. Vellasco, M.A.C. Pacheco, F.J. de Souza, Inverted hierarchical neuro-fuzzy BSP system: a novel neuro-fuzzy model for pattern classification and rule extraction in databases, IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews 36 (2) (2006) 236–248.
[28] L. Breiman, J. Friedman, R. Olshen, C. Stone, Classification and Regression Trees, Wadsworth, Belmont, CA, 1984.
[29] M. Setnes, U. Kaymak, Fuzzy modeling of client preference from large data sets: an application to target selection in direct marketing, IEEE Transactions on Fuzzy System 9 (2001) 153–163.
[30] M. Yalcin, T. Yıldırım, Karaciğer bozukluklarının yapay sinir ağları ile tes_hisi, in: Biyomedikal Muhendisliği Ulusal Toplantısı (BIYOMUT 2003), Istanbul, Turkiye, 2003, pp. 293–297.
[31] N. Cheung, Machine learning techniques for medical analysis, B.Sc. thesis, School of Information Technology and Electrical Engineering, University of Queenland, 2001.
[32] P. Langley, W. Iba, K. Thompson, An analysis of Bayesian classifiers, in: Proceedings of the Tenth National Conference on Artificial Intelligence, AAAI Press, San Jose, CA, 1992, pp. 223–228.
[33] P. Pulkkinen, H. Koivisto, Identification of interpretable and accurate fuzzy classifiers and function estimators with hybrid methods, Applied Soft Computing 7 (2007) 520–533.
[34] P. Wilding, M.A. Morgan, A.E. Grygotis, M.A. Shoffner, E.F. Rosato, Application of backpropagation neural networks to diagnosis of breast and ovarian cancer, Cancer Letters 77 (1994) 145–153.
[35] P.C. Chang, C.-Y. Fan, A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting, IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews 38 (6) (2008) 802–815.
[36] P.C. Chang, C.-H. Liu, A TSK type fuzzy rule based system for stock price prediction, Expert Systems with Applications 34 (1) (2008) 135–144.
[37] P.C. Chang, Y.W. Wang, Fuzzy Delphi and back-propagation model for sales forecasting in PCB industry, Expert Systems with Applications 30 (4) (2006) 715–726.

[38] P.C. Chang, T.W. Liao, Combing SOM and fuzzy rule base for flow time prediction in semiconductor manufacturing factory, Applied Soft Computing 6 (2) (2006) 198–206.
[39] P.C. Pendharkar, J.A. Rodger, G.J. Yaverbaum, N. Herman, M. Benner, Association, statistical, mathematical and neural approaches for mining breast cancer patterns, Expert Systems with Applications 17 (1999) 223–232.
[40] R.K. Lai, C.-Y. Fan, W.-H. Huang, P.-C. Chang, Evolving and clustering fuzzy decision tree for financial time series data forecasting, Expert Systems with Applications 36 (2–2) (2009) 3761–3773.
[41] R. Setiono, Generating concise and accurate classification rules for breast cancer diagnosis, Artificial Intelligence in Medicine 18 (2000) 205–219.
[42] R. Setiono, H. Liu, Symbolic representation of neural networks, Computer 29 (3) (1996) 71–77.
[43] R. Setiono, L. Huan, Understanding neural networks via rule extraction, in: Proceedings of the International Joint Conference on Artificial Intelligence, Morgan Kauffman, San Mateo, CA, 1995, pp. 480–487.
[44] János Abonyi, Ferenc Szeifert, Supervised fuzzy clustering for the identification of fuzzy classifiers, Pattern Recognition Letters 24 (14) (2003) 2195–2207.
[45] R.S. Youssif, C.N. Purdy, Combining genetic algorithms and neural networks to build a signal pattern classifier, Neurocomputing 61 (2004) 39–56.
[46] R. Setiono, Extracting rules from pruned neural networks for breast cancer diagnosis, Artificial Intelligence in Medicine 8 (1) (1996) 37–51.
[47] S. Greco, B. Matarazzo, R. Slowinski, Rough sets theory for multicriteria decision analysis, European Journal of Operational Research 129 (1) (2001) 1–47.
[48] S.C. Johnson, Hierarchical clustering schemes, Psychometrika 2 (1967) 241–254.
[49] S.C.K. Shiu, C.H. Sun, X.Z. Wang, D.S. Yeung, Maintaining case-based reasoning systems using fuzzy decision trees, in: Proceedings of the 5th European Workshop on Case-based Reasoning—EWCBR2K 5, Trento, Italy, 2000, pp. 285–296.
[50] T. Kervahut, J.Y. Potvin, An interactive-graphic environment for automatic generation of decision trees, Decision Support Systems 18 (2) (1996) 117–134.
[51] T. Van Gestel, J.A.K. Suykens, G. Lanckriet, A. Lambrechts, B.D. Moor, J. Vandewalle, Bayesian framework for least squares support vector machine classifiers, Gaussian processes and kernel fisher discriminant analysis, Neural Computation 14 (5) (2002) 1115–1147.
[52] X.G. Chang, J.H. Lilly, Evolutionary design of a fuzzy classifier from data, IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics 34 (3) (2004) 1894–1906.
[53] X.-N. Song, Y.-J. Zheng, X.-J. Wud, X.-B. Yang, J.-Y. Yang, A complete fuzzy discriminant analysis approach for face recognition, Applied Soft Computing 10 (2010) 208–214.
[54] Y.-J. Lee, O.L. Mangasarian, SSVM: a smooth support vector machine for classification, Computational Optimization and Applications 20 (1) (2001) 5–22.
[55] Y.-Z. Wu, M.-L. Giger, K. Doi, C.J. Vyborny, R.A. Schmidt, C.E. Metz, Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer, Radiology 187 (1993) 81–87.