International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 4 (2015) pp. 8533-8543

http://www.ripublication.com

A Soft Computing Paradigm For A Medical Data Mining Tool To Predict Risk of Coronary Heart Events

[1]Anil George K*, [2]Anitha R
[1]Assistant Professor, Department of Computer Science, St. Thomas' College, Thrissur-1, Kerala, India
[2]Professor, Department of MCA, Muthayammal Engineering College, Rasipuram, Tamilnadu, India
E-mail: [1]anilkonkoth@gmail.com, Phone: +919447518665, Fax: 0487 2421510

Abstract

In this paper, a Fuzzy-Classified Neural learning soft computing tool (FCNL) is proposed for predicting the intensity of risk in Coronary Heart occurrences. The presented model utilizes medical data collected from clinical findings on cardiac patients. The concept of decision trees is employed to classify the attributes that add to the Coronary Artery Disease (CAD). The output obtained as a result is then transformed to crisp if-then rules and then fuzzified into a database of fuzzy rules. A fuzzy-classified neural learning method based on supervised learning is exercised to enhance fuzzy membership functions. The performance and efficiency of the new medical data mining system, in terms of accuracy of prediction is presented against the real-life data.

Keywords: Coronary Heart Disease (CHD), Fuzzy-Classified Tree, Iterative Dichotomiser 3 (ID3) algorithm, Neural Networks, TSK model, Learning.

1. Introduction

Coronary Artery Disease (CAD) or Coronary Heart Disease (CHD) is the most frequent heart disease that leads to heart attacks. It is induced by reduced blood flow to the heart owing to the thinning of the arteries through the deposit of plaques along the inner of arteries. CAD leads to myocardial infarction (MI) by and large followed by sudden cardiac arrest. According to the studies from the American Heart Association, heart ailments are one of the leading causes of deaths in India and all over the world. Statistics show that one fifth of deaths in India are owing to coronary heart disease and it is projected to be one third of all deaths by the year 2020. It is estimated that presently there are about 45 million patients affected with coronary artery disease in India.

The aforementioned alarming situation can be adequately addressed by the timely altering of cardiovascular risk factors with diagnosis and healing at the beginning stages of CAD. Several computer-enhanced models for diagnosing CAD are presented in the literature. The purpose of such medical decision support systems is to enhance, not substitute, a medical expert's ability in the complex and highly discerning process of medical diagnosis.

Podgorelec et al. [1] emphasizes the role of decision tree based algorithms in medicine to provide consistent and effective results with better classification accuracy. Karaolis et al. [2] proposes a system which is based on the C4.5 (decision tree algorithm), used towards the extraction of rules for CAD events which in turn categorize the risks critical for coronary heart disease. A medical decision support system is presented by Tsipouras et al. [3] for the diagnosis of CAD and its better performance is established in comparison with ANNs and ANFIS. An effective Fuzzy inference system has been proposed by Hiremath et al. [4] in the medical field for the detection of follicle in ovaries. In the past, enormous work has been done to exploit the advantages of combining neural networks and decision trees [5]-[6]. Neuro fuzzy techniques have been employed by G. Ranganathan et al. for the estimation of heart rate signals [7]. Ebadian et al. [8] employed a neuro-fuzzy approach on SPECT images that are planar as well as gated, for predicting CAD events. Grossi [9] identified the predictive accuracy and potential advantages of neuro-fuzzy methods for cardiovascular risk assessment. Bensaid et al. [10] derived an ECG classification method for cardiac diagnosis using a hybrid method taking advantages of decision tree, fuzzy logic and neural networks.

In this paper, a fuzzy-classified neural learning system (FCNL), a soft computing tool has been used for diagnosing CAD. The system exploits seven risk factors of CAD patients admitted in coronary critical care unit (CCCU) after having been detected with CAD for the first time or without having a previous history of CAD, as the dataset. The FCNL model is implemented in three phases. In the first phase, a decision tree is constructed from the dataset using modified ID3 algorithm [11] and a set of rules is derived from it. The rules generated in the first phase are in the crisp form. In the second phase, the crisp rules are fuzzified using the Gaussian membership function. Towards the final phase of the model, a supervised backpropagation learning algorithm based on ANFIS [12] is used for optimization of rules and learning process. The system performs significantly well in the diagnosis of CAD which is validated through its performance accuracy.

2. ID3 Decision Tree – A Review

The Decision Trees (DT) are used for simplifying difficult decision-making procedure into a set of rather easy steps of decisions and thus, to provide a result effortlessly [13]. This divide and conquer approach of the DT chooses the attribute that best divides the input-output data and classifies the data according to the information content of the chosen attribute. One of the most efficient decision trees to classify the input data has been the ID3 algorithm [14]. The mathematical descriptions of decision tree generation based on the information measured by ID3 algorithm are as follows:

Let S contain a set of objects that constitute the training set. They belong to different classes and have target classes $\{O_1, O_2,..., O_{K1}\}$ associated with them. A test is conducted based on one attribute at a time. The training set represented as S is now divided into the subsets $S_1, S_2,…, S_{K1}$, where S_j holds those objects in S that have target O_j of the chosen set. Let $O_j(S)$ be the set of data instances in S that are assigned to class O_j. The information, measured in bits, contained in $C_j(S)$ is expressed as follows:

$$\mathrm{Info}(p_j(S)) \quad = \quad -\log_2 \frac{|O_j(S)|}{|S|} \tag{1}$$

Where $1 \le j \le K_1$ and |S| is the number of instances in S. In general, the probability distribution for the data set S of training instances is the information conveyed by this distribution

$$P(S) = \left(p_1(S) = \frac{|O_1(S)|}{|S|}, \cdots, p_{K_1}(S) = \frac{|O_{K_1}(S)|}{|S|} \right) \tag{2}$$

Therefore, entropy of S, known as total entropy, is defined as

$$\mathrm{Entropy}\ (S) = -\sum_{j=1}^{K_1} p_j(S) * \log_2(p_j(S)) \ , \tag{3}$$

When the partition of the instances of S is on the basis of the attribute A into K disjoint class regions $\{C_1, C_2,..., C_K\}$, it can offer enough information for classification of the attribute in comparison with the target classes. For this distribution, the expected information measurement can be found as the weighted sum over the subsets relating to the target attribute:

$$\mathrm{Entropy}\ (A,S) = \sum_{i=1}^{K} \frac{|S_i|}{|S|}\ \mathrm{Entropy}\ (S_i)\ , \tag{4}$$

where S_i, the subset of S, contains the elements in the i^{th} class of the attribute A. The *gain ratio* [15] is used to measure the precision of classification an attribute is able to achieve on the training set in relation to the target:

$$\mathrm{GainRatio}\ (A,S) = \frac{\mathrm{Gain}\ (A,S)}{\mathrm{SplitInfo}\ (A,S)}\ , \tag{5}$$

According to Quinlan [16], *Gain* is defined as the measure of information that is gained by partitioning S in accordance with the attribute A, which is the expected reduction in *entropy*.

i.e., $\mathrm{Gain}\ (A,S) = \mathrm{Entropy}(S) - \mathrm{Entropy}\ (A,S)\ ,$ (6)

where, $\mathrm{SplitInfo}\ (A,S) = -\sum_{i=1}^{K} \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|}$ (7)

On the basis of gain ratio, the ID3 algorithm selects the attribute with the largest value of gain ratio as superior node and assigns it as the root node. The partitioning process is applied on the root node according to the attribute classes $\{C_1, C_2,..., C_K\}$, and then find the root nodes corresponding to these classes by equations (3)-(7).

When this process is applied recursively to each node, the tree is grown in such a way that it starts from the most meaningful attribute and proceeds with all other attributes that are at the lower levels of the tree. In this approach, the task of decision tree classifier is restricted to three aspects [17]: *(i)* selecting a rule for splitting a node, *(ii)* identifying terminal nodes, and *(iii)* allocating each terminal node to a class label.

3. Fuzzy Decision Tree

This segment presents an appropriate meaning of a Fuzzy Decision Tree by improving the traditional ID3 algorithm in situations where ID3 falls short to bring out relevant knowledge from the interconnected region of overlapping classes. An enhanced decision tree algorithm with fuzzy concept, namely Fuzzy-Classified Decision Tree (FCDT) which achieves reduction of the computational complexity of rule extraction yet safeguarding the linguistic character of the decision in rule structure, is introduced below.

3.1 Parameters in Fuzzy Form

Consider the training data set (X, z) of N instances, where X represents the input dataset [x_1, x_2,…, x_n], and z represents the corresponding target output. One can formulate a model which will learn from examples $\{e_i, z_i\}_{i=1}^{N}$, wherein each element e_i and z_i has one-to-one relation R in the pattern pairs (X, z); where $e_i = [x_{i1}, x_{i2}, \cdots, x_{in}]$ such that $X \in R^n$ and $z \in R^n$. Dividing each input and output space into appropriate number of smooth fuzzy sub-regions, there exists overlapping from one fuzzy sub-region to another. Let U stand for input data set termed X. Then the fuzzy set F in the set X with N elements $\{e_{1j}, e_{2j}, \cdots, e_{Nj}\}$ of the j^{th} attribute can be defined as a set of ordered pairs

$$\{(e_{1j}, \mu_F(e_{1j})), (e_{2j}, \mu_F(e_{2j})), \cdots, (e_{Nj}, \mu_F(e_{Nj}))\} \tag{8}$$

Here μ_F stands for the membership function of F. Each attribute x_j can have K linguistic variables LV:

$$\mathrm{LV}(x_j) = \{\mathrm{LV}_1(x_j), \mathrm{LV}_2(x_j), \cdots, \mathrm{LV}_K(x_j)\}, \quad (j = 1, 2, \cdots, n+1) \tag{9}$$

The membership grades of x_j related to these linguistic variables are represented by $\mu_{1j}, \mu_{2j}, \cdots, \mu_{Kj}$ which specify the degree of involvement of the attribute x_j in the associated membership functions. A number of fuzzy sets such as *small* (S), *medium* (M), *large* (L) and *critical* (C) can be defined as the linguistic variables for each of these attributes. Gaussian membership function is used for fuzzification to accomplish classification yet maintaining [18] overlapping among the intersecting areas of linguistic classes as given below:

$$f(\mathrm{x}) \;=\; e^{y} \qquad \text{where } y = -\frac{(\mathrm{x}-\mathrm{c})^2}{2\sigma^2} \tag{10}$$

Here c and σ stand for the centre and width of the fuzzy region respectively. The attributes are expressed by some combinations of overlapping membership functions in terms of their linguistic variables S, M, L and C. When fuzzified, the n-dimensional input $P_i = [\mathrm{x}_{i1}, \mathrm{x}_{i2}, \cdots, \mathrm{x}_{in}]$, where $1 \leq i \leq N$, of N data instances with n attributes is changed into a 4n-dimensional vector P_i [19], which is defined as

$$\begin{aligned}\mathbf{P}_{\mathrm{i}} = [&\mu_{\mathrm{S}}(\mathrm{x}_{\mathrm{i1}}), \mu_{\mathrm{M}}(\mathrm{x}_{\mathrm{i1}}), \mu_{\mathrm{L}}(\mathrm{x}_{\mathrm{i1}}), \mu_{\mathrm{C}}(\mathrm{x}_{\mathrm{i1}}), \cdots,\\ &\mu_{\mathrm{S}}(\mathrm{x}_{\mathrm{in}}), \cdots, \mu_{\mathrm{C}}(\mathrm{x}_{\mathrm{in}})]\end{aligned} \tag{11}$$

where $\{\mu_{\mathrm{S}}, \mu_{\mathrm{M}}, \mu_{\mathrm{L}}, \mu_{\mathrm{C}}\}$ point out membership functions corresponding to the linguistic variables S, M, L and C respectively.

In order to achieve a reasonable input partition, normalization is done on N pattern points. Due to the normalization, each data point is mapped onto a range between [0, 1], and proper centres for all fuzzy variables corresponding to n attributes are chosen. Let $\{\mathrm{c}_{\mathrm{j1}}, \mathrm{c}_{\mathrm{j2}}, \mathrm{c}_{\mathrm{j3}}, \mathrm{c}_{\mathrm{j4}}\}$ be the four centres of the j^{th} attribute. The widths (spreads) from the centres to the left and right are defined as

$$Spread_{jm}\,(=\sigma_{jm}) = \begin{cases} c_{jm} - \lambda_j, & \text{towards left} \\ c_{jm} + \lambda_j, & \text{towards right} \end{cases}, \quad \text{for } 1 \leq m \leq M \tag{12}$$

$$\text{where, } \lambda_j \;=\; \max_m(|c_{jm} - c_{j(m+1)}|) + \delta_j, \quad \text{for } 1 \leq m \leq (M-1) \tag{13}$$

$$\text{where, } \delta_j \;=\; \frac{1}{100}[\max_m(|c_{jm} - c_{j(m+1)}|)], \quad \text{for } 1 \leq m \leq (M-1) \tag{14}$$

On fuzzification, the membership function values of the input attributes corresponding to the linguistic variables can be computed as

$$\mu(x_j)(\mathrm{P}_i) \;=\; \max_m\left(exp\left\{ -\frac{1}{2}\left(\frac{x_{ij} - c_{jm}}{\sigma_{jm}}\right)^2 \right\}\right) \tag{15}$$

where $\mu(\mathrm{x}_{\mathrm{j}})(P_{\mathrm{i}}) = \mu(\mathrm{x}_{\mathrm{ij}})$, j ranges from 1 to n, i from 1 to N and $1 \leq \mathrm{m} \leq 4$, with the variables S, M, L and C. On the basis of the extent of association in the linguistic classes, k membership values for each x_{ij} are generated by the Gaussian function. The membership function grades of the linguistic sets {S, M, L, C} are replaced by their respective class labels {1, 2, 3, 4}. The decision tree that is fuzzy in nature is now generated using the database in the form of class labels.

3.2 Rules Representation

A rule in the proposed scheme is represented by a branch of the FCDT which starts from the root proceeding to the leaf. When the data set at the input phase of the ID3 algorithm is applied, it generates decision trees that represent the fuzzy rules. In order to finalize the comprehensive and final *fuzzy rule set,* those rules corresponding to the unsolved nodes are rejected. This rule reduction technique is crucial since it prevents the possibility of exponential raise in rule generation with the large number of attributes. The final fuzzy rule set is given to the neural networks to incorporate learning to the system.

4. Fuzzy-Classified Neural Learning Architecture

The proposed design is a varied form of ANFIS based on the TSK model which was designed by Takagi, Sugeno and Kang. It is a neural network system with backpropagation having several layers of learning incorporated into it. Fig. 1 depicts the diagrammatic representation of the architecture. The construction of FCNL system is implemented in the following steps: one, the creation of a fuzzy rule set from the FCDT; two, formulation of the parameters of antecedent fuzzy functions; finally, identifying the consequent parameters through learning until the model is optimized.

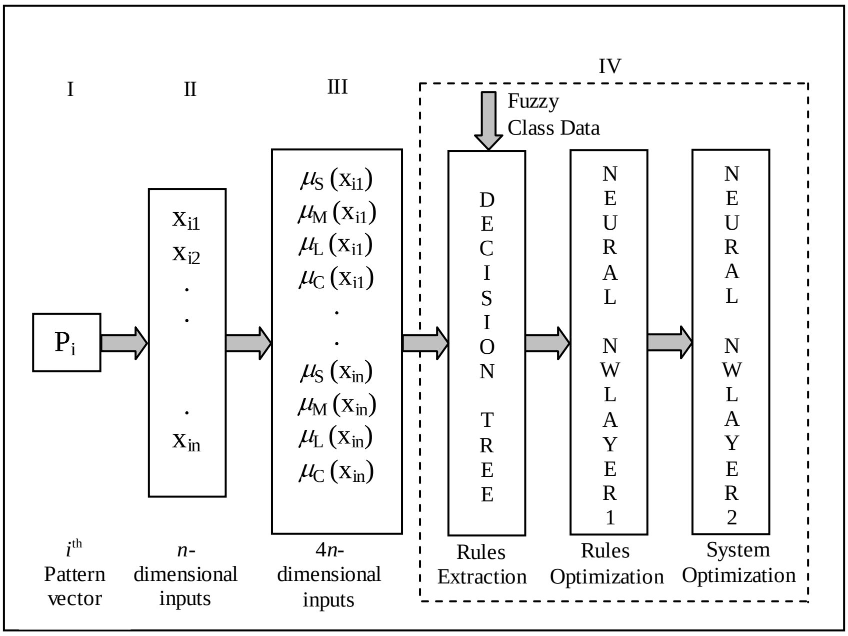

Fig. 1 Block Diagram of the Input Phase of FCNL System

4.1 Fuzzy-Classified Neural Learning Algorithm (FCNL)

The input to the algorithm is the fuzzy rule set obtained from FCDT explained in the previous section. The algorithm is as follows:

1. Fuzzify the input data using Gaussian membership function with appropriate widths and centers as the preliminary values.

$$\mu_{ij} = \mu_i(x_j) \quad = \quad \exp\left\{-\frac{1}{2}\left(\frac{x_j - c_{ij}}{\sigma_{ij}}\right)^2\right\} \tag{16}$$

2. Apply the rule base obtained from FCDT.
3. Calculate the weight (firing strength) assigned to each rule by means of '*min*' operator.
   $$w_i = \min\,(\mu_{ij})$$
4. Compute the overall output of the FCNL model.
   Overall output = z = $\sum_i k\,\overline{w}_i f_i$
   Where k is a scaling factor and f_i is a function of input attributes.
5. Calculate the difference between the actual and target outputs using the function.
   $$E \quad = \quad \sum_{r=1}^{L}(z_r - \tilde{z}_r)^2 \quad = \quad \sum_{r=1}^{L} e_r^2 \tag{17}$$
   where z_r and $\tilde{z}_r$ are the target and actual outputs respectively of the r^{th} training instance and E is the overall error measure.
6. Update antecedent and consequent parameter using gradient descent method.
7. Iterate steps 2 through 6 until the RMSE converges.

As the network uses the backpropagation algorithm, the antecedent (c_{rj}, center and σ_{rj}, width) and consequent parameters (q_{rj}, coefficient) are updated using error-correction by the gradient-descent method [20]. During the forward pass of the network, learning takes place. The backward pass is done to feed the error back and thus make correction between the actual outcome and the expected outcome. The steps are repeated until the system converges upon which de-fuzzification is employed to produce the final crisp output.

4.2 Experimental Studies

For the experiment, a total of 602 samples, with each sample describing seven attributes that add to heart disease, were collected. The attributes (risk factors) considered for the experiment are Age, Blood Pressure (both Systolic - SBP & Diastolic - DBP), Smoking, Blood Sugar (Fasting) represented as DM and Cholesterol (both LDL & Total - TC). Out of the total samples collected, half of the dataset (301 samples) was collected from cardiac patients admitted in the Coronary Critical Care Unit (CCCU) of a hospital and was used as case data for the study. The other half collected from people with the same age group but having no history of CAD was used as control data. The FCDT algorithm is employed on the patient database which results in the formation of rule base. Linguistic functions (S – Small, M – Medium, L – Large and C - Critical) are obtained through fuzzification of case data. The initial MFs are refined through the learning process. Patient data is employed to train the network and devise the system, while the control data is used for verifying it.

4.3 Sample Decision Tree

Consider a set of 10 input-output instances (Table 1) with seven attributes for input and the output attribute gives the risk factor. The input-output data set given here is in fuzzy format represented by their corresponding labels such as *small* = 1, *medium* = 2, *large* = 3 and *critical* = 4.

Sl. No	AGE	SMK	SBP	DBP	LDL	TC	DM	Risk
1.	3	1	3	2	3	2	2	3
2.	3	1	2	2	3	3	2	2
3.	2	3	2	2	2	2	2	2
4.	2	3	1	1	2	2	3	2
5.	2	1	3	3	2	2	3	2
6.	2	2	3	3	3	3	2	3
7.	2	3	2	2	4	4	3	3
8.	2	4	2	2	4	3	2	3
9.	3	1	2	2	3	3	4	3
10.	2	1	2	1	1	1	3	1

Table 1 Sample Data Set

With this CAD database, the ID3 algorithm provides the values of entropy and gain ratios related to each attribute.

Entropy = 1.4166; GainRatio (AGE) = 0.1932;
GainRatio (SMK) = 0.1619; GainRatio (SBP) = 0.2349;
GainRatio (DBP) = 0.1818; GainRatio (LDL) = 0.2897;
GainRatio (TC) = 0.3077 ; GainRatio (DM) = 0.2052

The attribute TC which provides more information than the other attributes is considered as the root for the extension of decision tree. The process is progressed by partitioning TC and finding the superior nodes along each class. Finally, a DT (Fig. 2) is obtained with 14 leaf nodes, where the leaf nodes represent the decision outputs corresponding to the target output. From the figure, it is obvious that there are three "unresolved" nodes which are discarded. The notations D_1, D_2, D_3 and D_4 correspond to the decision outputs, whereas the numbers {1, 2, 3, 4} indicate the classes of the input attributes.

As indicated in Fig. 2, the tree represents a set of rules, recording the test outcomes as antecedents and the leaf-nodes as the consequent.

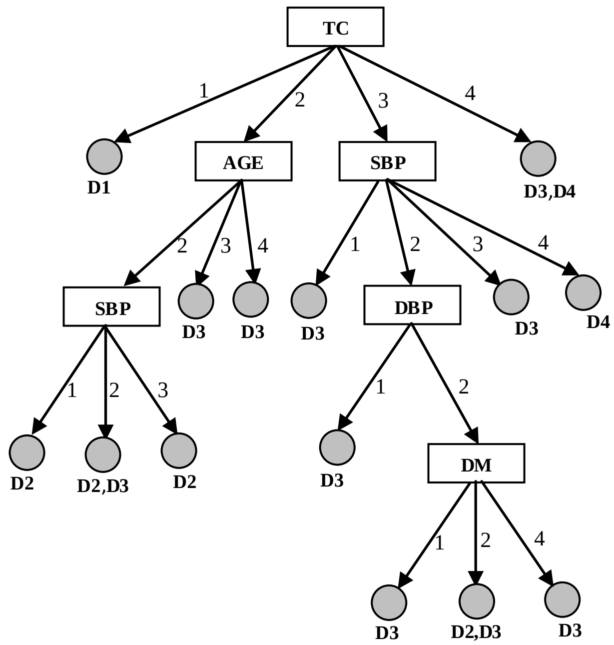

Fig. 2 Sample Decision Tree

5. Results

The FCNL model provides outstanding performance once the optimal number of iterations is completed. Initially, the level of convergence is below 40% with root mean square of sample error is < 0.15. But after training with 25 iterations, the system converges and it achieves 95.67% correct prediction on training data and 93.35% correct prediction on testing data. The higher accuracy is achieved by choosing suitable scaling factors (k = [1.01, 1.07]) and learning rate parameters. The learning rate parameters were empirically proved to be optimal since values above the applied range deliver worse performance. The process of learning of the system occurs by updating widths and centres of the linguistic classes of the risk factors.

Comparing the performance of the model with ANFIS, it is verified that the number of rules extracted from the FCNL model is significantly reduced even when the number of inputs and membership functions are large. For a data set of 301 instances and 7 input attributes, each with four fuzzy classes, the ANFIS theoretically produces 4^7 fuzzy rules that is relatively very large and is not easy to manage. On the other hand, the number of rules extracted using FCNL with FDT (fuzzy decision tree) is only 134 rules, a considerable improvement. The learning that takes place in the model enhances the capability of the system in diagnosing and predicting the level of risk of having CAD by optimizing its membership functions.

6. Conclusion

The FCNL model presented in this paper, aims at predicting the risk in coronary heart events in individuals using medical data. The FCNL integrates the benefits of neural networks, decision tree and fuzzy logic all incorporated into one. The proposed system shows its capability to provide a suitable decision-making tool for medical diagnosis with considerable amount of accuracy. The results show that the system has minimum computational requirements and can achieve high accuracy in diagnosis. Pre-diagnosis of CAD is possible with FCNL model, provided the person in question has risk factors of CAD. The model has employed supervised learning method. That the model has achieved 93.35% convergence confirms that the FCNL system is indeed a potential utility for decision making and pre-diagnosis in the medical field.

Acknowledgement

The authors make grateful acknowledgement to the UGC (University Grants Commission), India for the grant and financial support to the work. The authors acknowledge the support and cooperation rendered by the Cardiac Centre and the Research lab of Amala Medical College, Thrissur, Kerala, for having provided with the necessary facilities to gather clinical data.

References

[1] Podgorelec, V., Kokol, P., Stiglic, B., and Rozman, I., 2002, "Decision trees: an overview and their use in medicine," Journal of medical systems, 26(5), pp. 445-463.

[2] Minas A. Karaolis, Joseph A. Moutiris, Demetra Hadjipanayi, and Constantinos S. Pattichis, 2010, "Assessment of the Risk Factors of Coronary Heart Events Based on Data Mining with Decision Trees," IEEE transactions on Information Technology in Biomedicine, 14(3), pp. 559-566.

[3] Tsipouras, M. G., Exarchos, T. P., and Fotiadis, D. I., 2008, "Automated Diagnosis of Coronary Artery Disease Based on Data Mining and Fuzzy Modeling," IEEE transactions on Information Technology in Biomedicine, 12(4), pp. 447-458.

[4] Hiremath, P. S., and Tegnoor, J. R., 2014, "Fuzzy inference system for follicle detection in ultrasound images of ovaries," Soft Computing, 18, pp. 1353-1362.

[5] Suarez, A., and Lutsko, J. F., 1998, "Globally Optimal Fuzzy Decision Trees for classification and Regression," IEEE transactions on Pattern Analysis and Machine Intelligence, 21(12), pp. 1297-1311.

[6] Janikow, C. Z, 1998, "Fuzzy Decision Trees: Issues and Methods," IEEE Transactions on Systems, Man and Cybernetics, 28(1), pp. 1-14.

[7] Ranganathan, G., Rangarajan, R., and Bindhu, V., 2012, "Estimation of heart rate signals for mental stress assessment using neuro fuzzy technique," Applied Soft Computing, 12(8), pp. 1978 – 1984.

[8] Bagher-Ebadian, H., Soltanian-Zadeh, H., Setayeshi, S., and Smith, S. T., 2004, "Neural network and fuzzy clustering approach for automatic diagnosis of coronary artery disease in nuclear medicine," IEEE Transactions on Nuclear Science, 51(1), pp. 184 – 192.
[9] Grossi, E., 2006, "How artificial intelligence tools can be used to assess individual patient risk in cardiovascular disease: problems with the current methods," Journal of BMC Cardiovascular Disorders, pp. 1-6.
[10] Bensaid, A. M., Bouhouch, N., Bouchouch, R., Fellat, R., and Amri, R., 1998, "Classification of ECG patterns using fuzzy rules derived from ID3-induced decision trees," Proceedings of Conference of the North American Fuzzy Information Processing Society, pp. 34-38.
[11] Quinlan, J. R., 1990, "Decision Trees and Decision Making," IEEE transactions on Systems, Man and Cybernetics, 20(2), pp. 339-346.
[12] Jang, J. S. R., Sun, C. T., and Mizutani, E., 1997, Neuro-Fuzzy and Soft Computing, Prentice-Hall, Michigan.
[13] Mitra, S., Konwar, K. M., and Pal, S. K., 2002, "Fuzzy decision tree, Linguistic rules and Fuzzy knowledge-based network: generation and evaluation," IEEE Transactions on Systems, Man and Cybernetics, 32(4), pp. 328-339.
[14] Umano, M., Okamoto, H., Hatono, I., Tamura, H., Kawachi, F., Umedzu S., and Kinoshita, J., 1994, "Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems," Proceedings of the Third IEEE Conference on Fuzzy Systems, IEEEWorld Congress on Computational Intelligence, pp. 2113-2118
[15] Alhammady, H., and Ramamohanrao, K., 2006, "Using emerging patterns to construct weighted decision trees," IEEE Transactions on Knowledge and Data Engineering, 18(7), pp. 865-876.
[16] Quinlan, J. R., 1986, "Induction of Decision Trees," Machine Learning, 1, pp. 81-106.
[17] Safavian, S. R., and Landgrebe, D., 1991, "A Survey of Decision Tree Classifier Methodology," IEEE Transactions on Systems, Man and Cybernetics, 21, pp. 660-674.
[18] Kochurani, O. G., Aji, S., and Kaimal, M. R., 2007, "A Neuro Fuzzy Decision Tree Model for Predicting the Risk in Coronary Artery Disease", Proceedings of IEEE Multi-conference on Systems and Control, pp.133-145.
[19] Pal, S. K., and Mitra, S., 1999, Neuro-Fuzzy Pattern Recognition: Methods in Soft Computing, New York, Wile.
[20] Yen, J., and Langari, R., 2004, Fuzzy logic: intelligence, control and information, Pearson Education, India.