IMPROVED FUZZY ARTIFICIAL NEURAL NETWORK (IFANN) CLASSIFIER FOR CORONARY ARTERY HEART DISEASE PREDICTION IN DIABETES PATIENTS

B. Narasimhan Research Scholar PG and Research Department of Computer Science Government Arts College Coimbatore

Dr. A. Malathi* Assistant Professor PG and Research Department of Computer Science Government Arts College Coimbatore *Corresponding Author

ABSTRACT Soft computing techniques and its applications extends its wings in almost all areas which includes data mining, pattern discovery, industrial applications, robotics, automation and many more. Soft computing comprises of the core components such as fuzzy logic, genetic algorithm, artificial neural networks and probabilistic reasoning. In spite of these, recently many bio – inspired computing attracted attention for the researchers to work in that area. Machine learning plays an important role in the design and development of decision support systems, applied soft computing and expert systems applications. This research work aims to build an improved fuzzy logic based artificial neural network classifier for predicting coronary artery heart disease among diabetic patients. Real time data are obtained and the built IFANN classifier is compared with Takagi Sugeno Kang fuzzy classifier and ANN classifier in terms of prediction accuracy, sensitivity, specificity and Mathew's correlation coefficient. The significance of MCC is that to test the ability of the machine learning classifier in spite of other performance metrics. Implementations are done in Scilab and from the obtained results it is inferred that the built IFANN outperforms that that of TSK fuzzy classifier and ANN classifier.

KEYWORDS : soft computing, fuzzy logic, machine learning, CAHD, diabetes, artificial neural network, applications of soft computing.

1. INTRODUCTION

Typically the essential contemplations of traditional computing are accuracy, sureness, and meticulousness. We recognize this as hard computing. Interestingly, the important thought in soft computing is that exactness and sureness convey a cost; sand that calculation, reasoning, and basic leadership should misuse (wherever conceivable) the resistance for imprecision, uncertainty, approximate reasoning, and incomplete truth for getting minimal effort arrangements. This prompts the wonderful human capacity of understanding mutilated discourse, translating messy penmanship, appreciating the subtleties of characteristic language, condensing content, perceiving and arranging pictures, driving a vehicle in thick rush hour gridlock, and, all the more for the most part, settling on normal choices in a domain of uncertainty and imprecision. The test, at that point, is to misuse the resilience for imprecision by concocting strategies for calculation that lead to an acknowledge capable arrangement requiring little to no effort. This, generally, is the core value of soft computing.

There are continuous endeavors to incorporate artificial neural networks (ANNs), fuzzy set theory, genetic algorithms (GAs), rough set theory and other strategies in the soft computing worldview. Hybridization misusing the attributes of these speculations incorporate neuro-fuzzy, rough-fuzzy, neuro-genetic, fuzzy-genetic, neuro-rough, rough-neuro-fuzzy methodologies. Notwithstanding, among these, neuro-fuzzy computing mechanism has gained may researchers' attention in these days.

Heart disease remains the number one cause of death throughout the world for the past decades. In 2015, the World Health Organization (WHO) has estimated that 17.7 million deaths have occurred worldwide due to heart diseases. Heart diseases are the primary cause of death globally: more people die annually from CAHD than from any other causes. If we can predict the CAHD and provide warning beforehand, a handful of deaths can be prevented. The application of soft computing brings a new dimension to CAHD risk prediction.

2. RELATED WORKS

In order to enhance the prediction of heart disease, optimized crow search algorithm [1] was proposed, where it made an attempt to predict the heart disease more accuracy to provide on-time treatment. The results showed that the proposed algorithm is not fit for dataset related to heart disease, where the classification accuracy becomes very low. Two Class Classification [2] was proposed with the framework of machine learning by utilizing artificial neural network classification concept. The classifier works by selecting the spectral features of sub-band. The results show that classifier could not perform well when there are noisy data more than the remarkable range, where the false positive gets increases. Disease Specific Feature Selection strategy [3] was proposed for the purpose of heartbeat classification in a automated manner towards predicting the cardiac attack. It holds 1-vs-1 features idea towards searching for the best feature, where it uses the support vector machine classification concept. The result showed that the feature selection is not suited for this classification, where the results came with false negative rate got increased. Multi Objective Classification method [4] was proposed with the ensemble of particle swarm optimization and genetic algorithms in order to predict the heart disease in a early stage. It calculated the coefficient of polynomial, also the limit of the threshold value which was set for the class and attributes. This calculation was made to decrease the error, but the misclassification error got increased a lot in classifying to the wrong class Automated Classifier based on Support Vector Machine [5] was proposed to classify the electrocardiograms towards predicting the heart disease. It depends on the time period of electrocardiograms, to train the support vector machine to select the feature. The results proved that the classification accuracy went down due to feature selection concept, where the classifier omitted the important feature for classification. Modified version of Ant Colony Optimization [6] was proposed to increase the classification accuracy towards predicting the coronary artery disease, where it uses least square model of regression. The correlation coefficients were calculated for checking the fitness level between the selected features. The result came with low classification accuracy.

Deep Learning Strategy [7] for the classification of ECG towards heart disease prediction was proposed. This strategy was proposed with the target of classifying in a automatic manner. The result showed that the results were not efficient when comparing with the existing algorithms in the term of sensitivity. Fuzzy Classifier [8] was proposed to perform classification on with dynamic electrocardiogram signals with the intention to predict the heart disease in a early stage. It has worked with the dynamic unknown features resulting with very low accuracy. It was also analyzed that the algorithm can work good with known features only. Identification of Heart Disease with a Embedded System [9] was proposed and analyzed viability. It takes the input as electrocardiogram signals for the initial stage clustering and finally used Gustafson Kessel based Fuzzy clustering algorithm for the purpose of classifying and correlating the signals. The result came with increased false negative rate. Hybrid Classifier [10], which was a ensemble of neural network and genetic algorithm, was proposed for the classification of coronary artery disease. Initially neural network was performed and then the genetic algorithm was used. The result showed that the this specific hybrid classifier was not fit for the prediction of coronary artery disease, where the results came with very low classification accuracy.

3. Improved Fuzzy Artificial Neural Network (IFANN) Classifier

IFANN is a supervised mechanism that performs incremental learning to build up information from available training samples. IFANN is composed by two fuzzy components namely ART_a and ART_B (ART stands for Adaptive Resonance Theory) that are connected using a map field named F^{ab}. Every available IFNN consists of three layers of nodes. The nodes are termed as normalization layer $F_0^a(F_0^b)$ where in an M - dimensional input vector, the input layer $F_1^a(F_1^b)$ where in its nodes receive A and the recognition layer $F_2^a(F_2^b)$ where in each node represents a group of information taken out from the recognized input category. The number of recognition nodes increases upon insertion of new nodes to $F_2^a(F_2^b)$ for encoding newly learned information. At the point of time when the training phase is initiated, ART_a obtains an input pattern whereas ART_b obtains the target class of the input pattern. The ART endures a similar pattern-matching cycle which has node selection, similarity test and category search processes. In ART_a, once when the input pattern vector is obtained it has been complement-coded as vector A (where A = (a ; 1 − a) where 1 is the vector of all entries being 1), it is forwarded from $F_1^a(F_2^a)$ where in the activation of each recognition node j is computed using a choice function.

$$T_j = \frac{|A \wedge w_j^a|}{x_a + |W_j^a|} \ldots (1)$$

where $x_a \approx 0$ is the preference bound and w_j^a is the weight of node j . The fuzzy intersection $\wedge$ mentions

$$p \wedge q := (\min(p_i, q_i))_{2M} \ldots(2)$$

and the norm $|\cdot|$ is the l_1 norm:

$$|p| := \sum_{i=1}^{2M} |p_i| \ldots(3)$$

The adaptive fittest rule is referred to identify a fittest node J that responses with the highest activation value. A attention test is conducted to compute the degree of similarity between the fittest prototype w_j^a and A , and compare the result with a attention stricture $\rho_a \in [0,1]$:

$$\frac{|A \wedge w_j^a|}{|A|} \geq \rho_a \ldots(4)$$

If the attention test is failed, a new search cycle will be commenced to find for the next fittest node. The search process is carried on until the identified fittest node could pass in the attention test. If no such node could be found, a new node will be introduced in F_2^a to include A . On the other hand, on presentation of the target vector, ART_b also goes through a similar pattern-matching process to find a node in F_2^b to represent the target class. A map-field attention test is then carried out to evaluate the correctness of the prediction between the two fittest nodes from F_2^a and F_2^b by using the below equation

$$\frac{|y^b \wedge w_j^{ab}|}{|y^b|} \geq \rho_{ab} \ldots(5)$$

where y^b refers to the output vector of $F_2^b; w_j^{ab}$ refers to the weight vector from F_2^a to F^{ab}; and $\rho_a \in [0,1]$ denotes the map-field attention stricture. Normally, ρ_{ab} is set to a value close to 1 , e.g., $\rho_{ab} = 0.95$ If the map-field attention test is failed, it denotes the fittest node in F_2^a has predicted incorrectly the target class in F_2^b A match tracking is then operated to raise from a baseline attention stricture $\bar{\rho}_a$ (where $a\rho$ is a user- defined parameter in a range [0,1]) to

$$\rho_a = \frac{|A \wedge w_j^a|}{|A|} + \delta \ldots(6)$$

where δ is a constant being defined as a small number close to 0 ($e.g.\ \delta = 0.0001$) The purpose of match tracking is to avoid the current fittest node F_2^a in from passing in the ART_a attention test again so that another fittest node could be identified in a new search cycle. The search process is continued until both fittest nodes in F_2^a and F_2^b have made a correct prediction.

Each dimension d of a prototype p in F_2^a has either as $s_{pd} = 0$ or $s_{pd} = 1$ Initi ally all dimensions of a prototype are set to 0. When the prototype dimension is shrunk, its s_{pd} is updated to 1. Further, each F_2^a prototype consists of a reference vector w_j^r . Initially, w_j^r is a zero vector. When IFANN is in the resonance state, apart from the weight vector w_j^a of the J -th fittest node, its w_j^r is also updated iteratively using the below equation.

$$(w_j^r)^{new} = (w_j^r)^{old} + \frac{1}{N_J}\left[A - (w_j^r)^{old}\right] \ldots(7)$$

where N_J represents the latest number of input patterns categorized correctly by the J -th node, $N_J = N_J + 1$.

The prototypes of two fuzzy ARTs and their associations that are established in the map-field during the training phase are utilized to predict an output class on presentation of an unseen pattern during the test phase. The training procedure of IFNN is given below:

1. An M -dimensional input pattern $a \in [0,1]^M$ is complement-coded to a 2M -dimensional vector A in F_0^a; A is then forwarded to F_1^a.
2. A is forwarded to F_2^a through the weight vector, w^a. The activation of each node is calculated using Eq. (1) . The node with the highest activation value is selected as the fittest node J .
3. The prototype of node J is sent backward from F_2^a to F_1^a for evaluation by a attention test as in Eq. (4) .
4. If the attention test is not satisfied, go to Step 3 where a new search cycle for another fittest node is carried out (the same search cycle also happens in ART_b for finding a fittest node).
5. Upon receiving a prediction from F_2^a ($i.e., w_j^{ab}$) and also from F_2^b ($i.e., y^b$) at the map-field F^{ab}, a map-field attention test as in (5) is run.
6. If the map-field attention test is not satisfied, a match tracking as in (6) is exercised. Notably, match tracking only happens, in the ART_a module. Go to Step 3.
7. The weight vectors w_J^a and w^r are adjusted . Likewise, the weight vector w_J^b of the fittest node in ART_b is adjusted by replacing the symbol a with b.

4. The TSK fuzzy inference mechanism applied in IFANN

A rule-based model need to be established prior to implementation of FIS. The rules are typically defined in this format:

$$R_i : if\ u_1\ is\ A_{i1}, ..., and\ u_n\ is\ A_{in}, then\ v_i = f_i^0(a; b_i), i = 1, 2, ..., I \ldots(8)$$

where R_i denotes the i -th rule; $u_{1,...,u_n}$ denote the input variables; $A_{i1},...,A_{in}$ denote the fuzzy sets of the input variables; v_i denotes output value of the i -th rule; a is the input vector; $f_i^0(a;b_i)$ indicates the O -th order of a polynomial function of a with a constant term b_i For any $R_i, u_1, \ , u_n$ repres ent the antecedences whereas v_i the consequence of the rule.

Occasionally, a zero-order TSK model is defined for handling pattern classification problems. In this case (of zero order), the Consequence of $R_i, i.e., f_i^0(a;b_i)$ is a constant, i.e. Hence, v_i is a discrete number re-written as

$$R_i : if\ u_1\ is\ A_{i1}, ..., and\ u_n\ is\ A_{in}, then\ v_i = b_i, i = 1, 2, ..., I \ldots(9)$$

When a data sample x_k is presented to the model, the firing strength of the i -th rule R_i is computed using an AND operator (i.e. a T-norm operator such as min) that combines the membership values between the data sample and the antecedences of R_i , i.e.

$$\xi_i(x_k) := AND(F_1(u_1, x_{k1}), ..., F_n(u_n, x_{kn})) \ldots(10)$$

where $\xi_i(x_k)$ is the firing strength of R_i given $x_k; F_1()...F_n()$ denote input membership functions. The qualified consequence of R_i on the firing strength is $\xi_i(x_k)v_i$. The qualified consequences of all rules based on firing strengths are aggregated, i.e.

$$\sum_{i=1}^{I} \xi_i(x_k)v_i \ldots(1)$$

The output of x_k is the weighted average of all rule outputs, as follows:

$$\hat{y}_k = \frac{\sum_{i=1}^{K} \xi_i(x_k)v_i}{\sum_{i=1}^{K} \xi_i(x_k)} \ldots(12)$$

IFANN is thus modeled for performing the classification task.

5. About the Dataset

The dataset is obtained from cardiac based medical centers. The dataset contains 7525 diabetic patients' records that have data from 4329 males and 3196 females. Totally 17 attributes including class label denoting whether the corresponding patient is likely to have CAHD risk or not. As far as 4329 male diabetic patients' records, 3911 patients owe the CAHD risk and 418 male diabetic patients' do not owe the CAHD risk. As far as 3196 female diabetic patients' records, 2808 patients owe the CAHD risk and 388 female diabetic patients' do not owe the CAHD risk. Scilab 6.0.2 has been utilized for implementation and experiments have been conducted on desktop personal computer with a 3.4 giga hertz Intel Core i7-6700 processor and 8 giga bytes RAM. Table - 1 shows the details of the dataset.

Table - 1. Dataset Details

Number of Attributes	Total Number of patients	Male – 4329		Female – 3196	
		Number of patients with risk of CAHD	Number of patients with no risk of CAHD	Number of patients with risk of CAHD	Number of patients with no risk of CAHD
17	Male 4329 + Female 3911 = 7525 patients	3911	418	2808	388

6. Results and Discussions

Male patients and female patients records are tested separately. Before that, 60% of the patient records (both male and female) are taken for training the classifier. 100% of the patient records are tested for performance evaluation in terms of sensitivity, specificity, prediction accuracy and Matthews correlation coefficient (MCC). The results are portrayed in the Table – 2 and Table – 3 for male and female patients respectively.

Table – 2. Performance Results – Male Patients

Classifiers	TP	TN	FP	FN	Sensitivity (in %)	Specificity (in %)	Accuracy (in %)	Mathews correlation coefficient (in %)
TSK Fuzzy Classifier	3156	339	392	442	87.72	46.37	80.73	33.21
ANN Classifier	3211	393	322	403	88.85	54.97	83.25	42.00
Proposed IFANN Classifier	3352	385	266	326	91.14	59.14	86.32	48.51

Table – 3. Performance Results – Female Patients

Classifiers	TP	TN	FP	FN	Sensitivity (in %)	Specificity (in %)	Accuracy (in %)	Mathews correlation coefficient (in %)
TSK Fuzzy Classifier	2372	205	321	298	88.84	38.97	80.63	28.32
ANN Classifier	2263	360	303	270	89.34	54.30	82.07	44.48
Proposed IFANN Classifier	2385	341	255	215	91.73	57.21	85.29	50.29

Fig.1. shows the performance analysis in terms of CAHD prediction accuracy for male patients. From the results it is inferred that conventional TSK fuzzy classifier obtains 80.73% and conventional ANN classifier obtains 83.25%; whereas the proposed IFANN classifier outperforms than both TSK classifier and ANN classifier by obtaining 86.32%. Fig.1. shows the performance analysis in terms of sensitivity for male patients. From the results it is inferred that conventional TSK fuzzy classifier obtains 80.73% and conventional ANN classifier obtains 83.25%; whereas the proposed IFANN classifier outperforms than both TSK classifier and ANN classifier by obtaining 86.32%. Fig.1. shows the performance analysis in terms of specificity for male patients. From the results it is inferred that conventional TSK fuzzy classifier obtains 80.73% and conventional ANN classifier obtains 83.25%; whereas the proposed IFANN classifier outperforms than both TSK classifier and ANN classifier by obtaining 86.32%. Fig.1. shows the performance analysis in terms of Mathews correlation coefficient for male patients. From the results it is inferred that conventional TSK fuzzy classifier obtains 80.73% and conventional ANN classifier obtains 83.25%; whereas the proposed IFANN classifier outperforms than both TSK classifier and ANN classifier by obtaining 86.32%.

Fig.2. shows the performance analysis in terms of CAHD prediction accuracy for female patients. From the results it is inferred that conventional TSK fuzzy classifier obtains 80.63% and conventional ANN classifier obtains 82.07%; whereas the proposed IFANN classifier outperforms than both TSK classifier and ANN classifier by obtaining 85.29%. Fig.2. shows the performance analysis in terms of sensitivity for female patients. From the results it is inferred that conventional TSK fuzzy classifier obtains 88.84% and conventional ANN classifier obtains 89.34%; whereas the proposed IFANN classifier outperforms than both TSK classifier and ANN classifier by obtaining 91.73%. Fig.2. shows the performance analysis in terms of specificity for female patients. From the results it is inferred that conventional TSK fuzzy classifier obtains 38.97% and conventional ANN classifier obtains 54.30%; whereas the proposed IFANN classifier outperforms than both TSK classifier and ANN classifier by obtaining 57.21%. Fig.2. shows the performance analysis in terms of Mathews correlation coefficient for female patients. From the results it is inferred that conventional TSK fuzzy classifier obtains 28.32% and conventional ANN classifier obtains 44.48%; whereas the proposed IFANN classifier outperforms than both TSK classifier and ANN classifier by obtaining 50.29%.

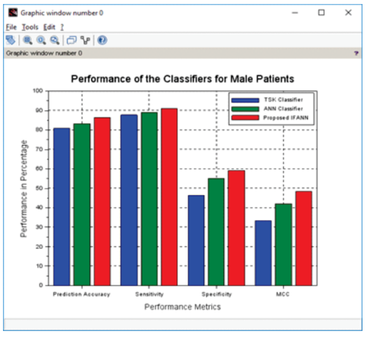

Fig.1. Performance of the Classifiers in Male Patients

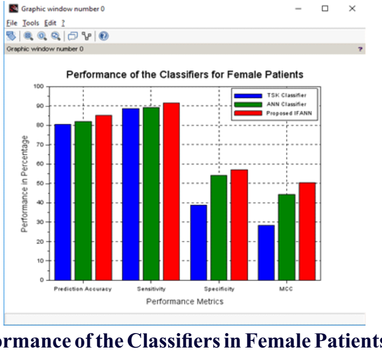

Fig.2. Performance of the Classifiers in Female Patients

7. CONCLUSION

The usage of soft computing techniques in medical domain is more prominent and emerging area of research. Several decision support systems are built for diagnosing diseases among patients. In this research work the aim of the proposed IFANN classifier is to attain maximum prediction accuracy for CAHD among diabetic patients. Both male and female diabetic patient records are obtained from the reputed medical centers along with the class label of CAHD occurrence. The results are promising and it is inferred that 86.32% accuracy is obtained for male diabetic patients and 85.29% accuracy is obtained for female diabetic patients. Yet there is more scope for further improving the prediction accuracy and in the near future some optimization techniques are aimed to be build for attribute selection.

REFERENCES

[1] D. Gupta, S. Sundaram, A. Khanna, A. E. Hassanien, V. H. Albuquerque, Improved diagnosis of Parkinson's disease using optimized crow search algorithm, Computers & Electrical Engineering, Vol 68, pp 412-424, 2018.

[2] P. Samanta, A. Pathak, K. Mandana, G. Saha, Classification of coronary artery diseased and normal subjects using multi-channel phonocardiogram signal, Biocybernetics and Biomedical Engineering, 2019.

[3] Z. Zhang, J. Dong, X. Luo, K. Choi, X. Wu, Heartbeat classification using disease-specific feature selection, Computers in Biology and Medicine, Vol 46, pp 79-89, 2014.

[4] M. Esmaelian, H. Shahmoradi, M. Vali, A novel classification method: A hybrid approach based on extension of the UTADIS with polynomial and PSO-GA algorithm, Applied Soft Computing, Vol 49, pp 56-70, 2016.

[5] V. M. Guerra, J. Novo, J. Rouco, M. G. Penedo, M. Ortega, Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers, Biomedical Signal Processing and Control, Vol 47, pp 41-48, 2019.

[6] L. Hu, C. Yin, S. Ma, Z. Liu, Rapid detection of three quality parameters and classification of wine based on Vis-NIR spectroscopy with wavelength selection by ACO and CARS algorithms, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, Vol 205, pp 574-581, 2018.

[7] G. Sannino, G. D. Pietro, A deep learning approach for ECG-based heartbeat classification for arrhythmia detection, Future Generation Computer Systems, Vol 86, pp 446-455, 2018.

[8] M. H. Vafaie, M. Ataei, H. R. Koofigar, Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals, Biomedical Signal Processing and Control, Vol 14, pp 291-296, 2014.
[9] H. H. Carvalho, R. L. Moreno, T. C. Pimenta, P. C. Crepaldi, E. Cintra, A heart disease recognition embedded system with fuzzy cluster algorithm, Computer Methods and Programs in Biomedicine, Vol 110, pp 447-454, 2013.
[10] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, A. A. Yarifard, Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm, Computer Methods and Programs in Biomedicine, Vol 141, pp 19-26, 2017.