International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-11 Issue-1, November, 2021

Optimal Neuro-Fuzzy Classification for Arrhythmia Data Driven System

Hela Lassoued, Raouf Ketata, Hajer Ben Mahmoud

Abstract: ***This paper presents a data driven system used for cardiac arrhythmia classification. It applies the Neuro-Fuzzy Inference System (ANFIS) to classify MIT-BIH arrhythmia database electrocardiogram (ECG) recordings into five (5) heartbeat types. In fact, in order to obtain the input feature vector from recordings, a time scale method based on a Discrete Wavelet Transform (DWT) was investigated. Then, the time scale features are selected by applying the Principal Component Analysis (PCA). Therefore, the selected input feature vectors are classified by the Neuro-Fuzzy method. However, the ANFIS configuration needs mainly the choice of an initial Fuzzy Inference System (FIS) and the training algorithm. Indeed, two clustering algorithms which are the fuzzy c-means (FCM) and the subtractive (SUBCLUST) algorithms, are applied to generate the initial FIS. Besides, for tuning the ANFIS membership function and rule base parameters, Gradient descent and evolutionary training algorithms are also evaluated. Gradient descent consists of the backpropagation (BP) method and its hybridization with the least square algorithm (Hybrid). However, the evolutionary training methods involve the Particle Swarm Optimization (PSO) and the Genetic Algorithm (GA). Therefore, eight (8) ANFIS are configured and assessed. Accordingly, a comparison study between their obtained Root Mean Square Error (RMSE) is analyzed. At the end, we have selected an optimal ANFIS which uses the SUBTRUCT algorithm to generate the initial FIS and the GA to tune its parameters. Moreover, to guarantee the effectiveness of this work, a comparison study with related works is done.***

Keywords: Data driven system, classification, arrhythmias Neuro-Fuzzy, gradient descent, evolutionary algorithms, optimization.

I. INTRODUCTION

Data driven systems have significantly improved decision making in various fields. They are mostly used in manufacturing [1], telecommunications [2], engineering [3], and medicine [4]. For instance, in medicine, it is difficult to make the right decision instantaneously due to massive amounts of knowledge. Therefore, experts need an intelligent tool to make more informed decisions at a quicker pace. In fact, artificial intelligence algorithms are used to develop the data driven system. They are based on the human brain neurons to perform computation. Accordingly, artificial neural networks (ANN) are considered as popular supervised learning approaches which are configured mainly for their capacity of training data and their ability to adjust weights by reducing the error between desired and predicted data. These networks have been used in medical decision making. They have produced high accuracy and quick response time in many medical tasks [5]. They are also regarded as an alternative tool to find patterns in large data sets [6].

Nevertheless, the decision making process needs to interpret and act as a human brain. This aspect is offered mainly by fuzzy logic systems which are famous for their human reasoning. Therefore, several fuzzy methods are also used to develop data driven systems [7]. However, applying Fuzzy Inference System (FIS) to analyze input/output data, assumes that discerning membership functions parameters by looking at data can be difficult or impossible. In these cases, rather than choosing the parameters associated with a given membership function arbitrarily, the Sugeno fuzzy inference systems can be tuned using neuro-adaptive learning techniques, such as, the artificial Neuro-Fuzzy inference system (ANFIS) [7]. Its main potential is the power of adaptation and the rapid learning capability. It is proposed to overcome the disadvantages in neural and fuzzy processes and combine their powerful criteria [8]. In fact, it is a Neuro-Fuzzy system that trains data without the need for an expert knowledge rule base. It deals with both numerical and linguistic knowledge.

The ANFIS as the ANN is able to identify patterns and classes. However, it leads to less memorization errors and has a greater clarity. Neuro-Fuzzy systems can be trained with derivative-based methods like gradient descent [7], or with evolutionary algorithms (EAs) such as genetic algorithms (GA) and swarm intelligence [9]. EAs have less probability to stick in a local optimum. The ANFIS has been applied to many decision-making systems in medical fields including brain tumor detection [10], Alzheimer detection [11] the COVID classification [12], the breast cancer [13], thyroid disorder [14] and the arrhythmia classification [7] and [15].

The Neuro-Fuzzy arrhythmia classification is also evaluated in this study. It is still a challenging task since cardiovascular arrhythmias necessitate an early diagnosis to prevent death of patients. Therefore, eight ANFIS models are investigated and evaluated regarding the selected initial FIS and the selected training algorithms.

Manuscript received on November 23, 2021.
Revised Manuscript received on November 28, 2021.
Manuscript published on November 30, 2021.

* Correspondence Author

Hela Lassoued*, National Institute of Applied Science and Technology INSAT, Tunis, Tunisia. Email: laswed.hela@gmail.com

Raouf Ketata, National Institute of Applied Science and Technology INSAT, Tunis, Tunisia. Email: ketataraouf@gmail.com

Hajer Ben Mahmoud, National Institute of Applied Science and Technology INSAT, Tunis, Tunisia. Email: hajerbenmahmoud@hotmail.com

They are evaluated in order to select the most accurate ANFIS. In this paper, we investigate a number of ANFIS models for classifying ECG recordings into five arrhythmia classes. In order to obtain an ANFIS, it is required to generate its corresponding FIS. For this purpose, two algorithms which are the Fuzzy c-means (FCM) and the subtractive clustering algorithm (SUBCLUST) were applied [16]. Then, to adjust the obtained FIS membership functions, an optimization approach is selected. Therefore, Gradient descent and evolutionary optimization algorithms are evaluated. Regarding gradient algorithms, two algorithms were used, including, the backpropagation (BP) and its combination with the least square. However, the Particular Swarm optimization (PSO) and the Genetic Algorithm (GA) were selected as the evolutionary algorithms [17]. Then, a comparative analysis, between the obtained results of the different ANFIS, is done and discussed in order to select the most appropriate network for this arrhythmia classification task. A comparative study with related works is also done in order to highlight the effectiveness of the selected optimal ANFIS. The following parts of this paper are organized as follows. Section2 presents the materials and methods. Section3 describes the experimental work. Section4 shows and discusses the obtained results. At the end a conclusion is achieved.

II. MATERIALS AND METHODS

The block diagram of the proposed work is illustrated by Fig.1. In fact, forty eight (48) electrocardiogram (ECG) recordings, collected from MIT-BIH arrhythmia Database, are firstly preprocessed and then classified into five (5) heartbeat classes [18]. The preprocessing stage includes feature extraction and selection parts in order to get the input feature vectors. In this study, a time scale pre-processing block was applied to extract the features [18]. Then, a Principal Component Analysis (PCA) was used to select the most pertinent features [19-20]. Once the preprocessing phase is achieved, the selected input vectors will be treated by a Neuro-Fuzzy classification approach. This stage includes the initial FIS configuration and the ANFIS Designing.

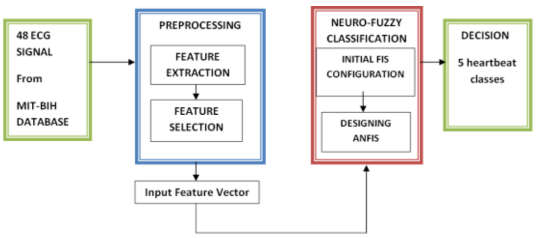

Fig. 1.Arrhythmia classification block diagram.

As it is described in Fig.2, the selected input feature vectors are used as inputs by the NEURO-FUZZY block. In fact, multiple ANFIS are evaluated. These ANFIS classifiers are different among their clustering and training algorithms. Indeed, their configuration demands firstly the generation of the corresponding FIS by using either the FCM algorithm or the SUBCLUST one. Then, for their designing, it is essential to select a training algorithm. These algorithms are mainly used for tuning the ANFIS membership functions and rule base parameters. For instance, gradient descent and evolutionary algorithms are analyzed. At the end, an optimal ANFIS model which achieves accurate performances will be picked.

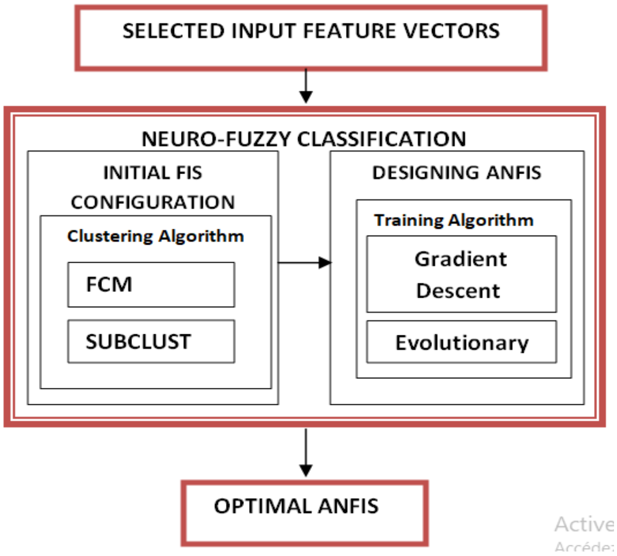

Fig. 2.Neuro-Fuzzy classification methodology.

A. Dataset

In the present study, we have analyzed the MIT-BIH arrhythmia database accessible from Physionet.org [21]. The database contains 48 Electrocardiogram (ECG) recordings, as it is shown in Fig.3. It represents a single heartbeat which consists of three electrical wave shapes. The P wave represents the contractions that pump blood in the upper chambers of the heart. It is named atrial depolarization. However, the QRS complex and the T wave represent activity in the lower chambers of the heart. They involve the ventricular depolarization and ventricular repolarization, respectively [22-23].

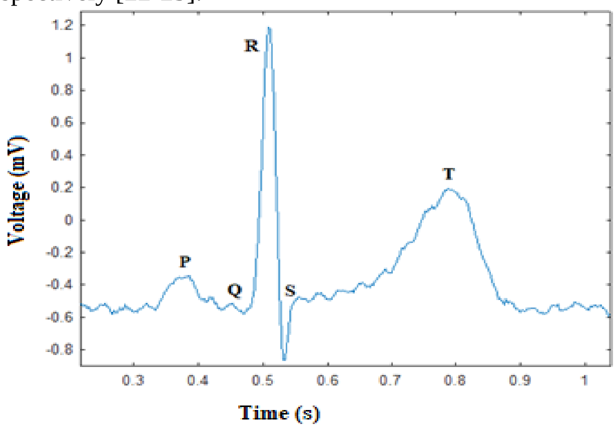

Fig. 3.Normal ECG heartbeat.

Arrhythmias are represented in Fig.4. The Fig.4 (a) and Fig.4 (b) describe PVC and P beats, respectively. The PVC beats correspond to premature contraction of one of the ventricles. The paced beats are generated by a pacemaker and they are detected by the enlarged RR intervals (>280 ms). However, the Fig.4 (c) and Fig.4 (d) represent the LBBB and RBBB arrhythmias, respectively.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

In fact, both of them are identified by the presence of a widened QRS complex (> 120ms). Particularly, LBBB arrhythmia is detected by a wrong displacement of ST and T waves. Indeed, they are mostly opposite to the major deflection of the QRS complex. However, the RBBB arrhythmia is identified by a slurred S waves [22].

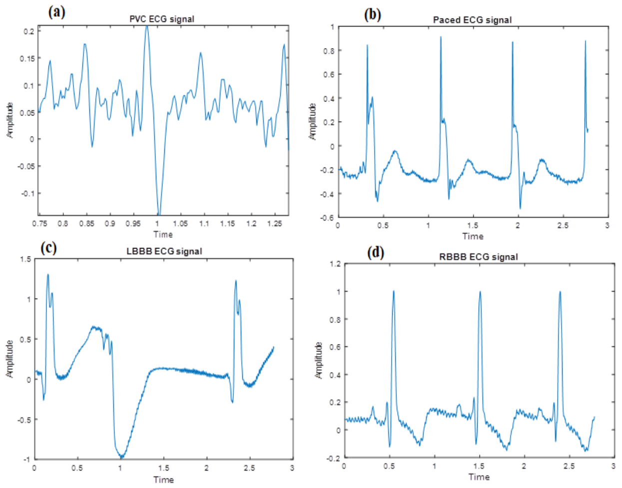

Fig. 4.Arrhythmia heartbeats.

By the way, classification needs feature extraction and selection stages. Accordingly, it is important to extract and select relevant features from data sets. In fact, the discrete wavelet transform was applied for ECG feature extraction. Indeed, it decomposes the original signal into low and high frequency components. The high-frequency components at low scales are called details and the low frequency components at high scale are known as approximations. Accordingly, the mother wavelet Daubechies 6 (db6) was applied to decompose the signal up to eight (8) levels [24]. Therefore, the obtained approximations and details wavelet coefficients were collected for each ECG recording. However, a high dimensional feature vector size was acquired. In order to reduce its dimension, the arithmetic mean (Moy), variance (Var) and standard deviation (Std) are computed for each detail (d_i where i is in {1,2,3,4,5,6,7,8}) and approximation (a_j, where j is in {1,2,3,4,5,6,7,8})at each level. Accordingly, 48 time-scale coefficient features for each ECG recording are collected. Then, the Principal Components Analysis (PCA) is applied to select the most relevant features [25]. Hence, the input matrix (48x48) becomes a reduced matrix (10x48) including ten (10) features for forty eight (48) ECG recordings. The selected features are described in Table-I.

Table- I: Selected Features.

Features	Mean	Std
Var a_6	7959	430.764
Var a_4	3982.688	213.290
Var a_5	5631.917	302.092
Var a_3	2815.396	151.287
Var a_1	1407.458	75.598
Std a_2	97.818	65.071
Std a_1	69.350	45.944
Std a_3	136.028	92.283
Std d_5	105.341	75.924
Std a_4	177.064	132.796

Based on the selected features, we estimated whether the ECG signal is in one of the four arrhythmia cases which are the PVC, P, LBBB and RBBB or in the normal one. The dataset characteristics are given in Table- II.

Table- II: Dataset Description.

Dataset	
Instances number	48
Features number	10
Features type	Real, Integers
Task	Classification
Learning algorithm	Supervised
Machine learning approach	ANFIS

In this study, the selected dataset was analyzed by using MATLAB and was later divided into 70% for training and 30% for testing dataset. The division respects the Association for the Advancement of Medical Instrumentation (AAMI) recommendations [26].

B. Neuro-Fuzzy Inference System

Adaptive Neuro-Fuzzy systems (ANFIS) are the result of the simultaneous use of neural and fuzzy systems. They are able to take advantage of the learning capabilities of the neural networks and the interpretation of the fuzzy systems [7]. Furthermore, ANFIS system is a FIS which is put in the form of a multilayer network. It consists of a Sugeno fuzzy inference model with multi-layer Neural Network [7]. It endorses five (5) layers which manipulate adaptive and non-adaptive nodes instead of synaptic weights. This model is based on a fuzzy set theoretical mechanism. In fact, the membership values vary between 0 and 1. A fuzzy number represents the associated set of possible values, where each value ranges between 0 and 1. The range is called as a membership function. Linguistic variables such as low, very low, medium, high and very high are used to illustrate the complex expressions. In order to explain the ANFIS architecture, we assumed that there are two inputs: x and y.

Two fuzzy if–then rules for a first-order Sugeno fuzzy model can be expressed as follows [9]:

Rule 1: If x is A_1 and y is B_1, then $f_1 = p_1x + q_1y + r_1$,

Rule 2: If x is A_2 and y is B_2, then $f_2 = p_2x + q_2y + r_2$,

where A_i and B_i are the fuzzy sets, f_i is the output, and p_i, q_i, and r_i are the design parameters that are determined during the training process.

The ANFIS architecture which uses the two rules is shown in Fig. 5.

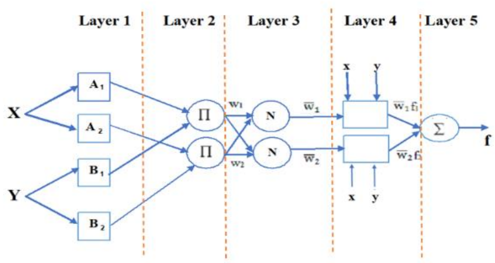

Fig. 5.ANFIS model architecture with two inputs, one output, and two rules

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The ANFIS architecture includes five layers as follows. Where: O_{ij} represents the output of the i^{th} node and j^{th} layer.

In Layer 1, all nodes are adaptive and represent the input membership degrees, as defined in the equations (1) and (2).

$$o_{1,i} = \mu_{A_i}(x), i = 1,2 \quad (1)$$

$$o_{2,i} = \mu_{B_i}(x), i = 1,2 \quad (2)$$

Where μ_{A_i} and μ_{B_i} can adopt several membership function types.

For example, if the membership function is Gaussian then its parameters $\mu_{A_i}(x)$ are calculated by equation (3), with a_i and c_i are the membership parameters.

$$\mu_{A_i}(x) = \exp\left[-\left(\frac{x - c_i}{a_i}\right)^2\right] \quad (3)$$

In layer 2, the nodes are all fixed. This layer uses the operator (AND) to fuzzify the inputs. The output of this layer is represented by equation (4) with w_i represents the degree of activation of each rule.

$$\overline{w}_i = \mu_{A_i}(x) * \mu_{B_i}(y), i = 1,2 \quad (4)$$

In layer 3, the nodes are also fixed, labeled by N. They calculate the normalized value of the product.

The output of this layer can be represented by equation (5).

$$\overline{w}_i = \frac{w_i}{w_1 + w_2}, i = 1,2 \quad (5)$$

In layer 4, the nodes are adaptive. The output of each node in this layer is simply the product of $\overline{W}_i$ and a first order polynomial. The output of this layer is given by equation (6) with: $\overline{W}_i$ is the output of layer 3 and (p_i, q_i and r_i) are the parameters of the consequence represented by the equation (6).

$$\overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i), i = 1,2 \quad (6)$$

In layer 5, there is only one fixed node. This node performs the sum of the outputs of the previous layer. The overall result of the ANFIS model is given by equation (7).

$$\sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}, i = 1,2 \quad (7)$$

The fuzzy membership functions and rules parameters are usually identified by a training algorithm. In fact, the most used algorithms are the derivative based algorithms, including, the Gradient descent with BP algorithm or with the hybridization between the least square approach and back propagation algorithm (Hybrid). The evolutionary algorithms become also popular. In this study both of them are evaluated and tested.

C. Neuro-Fuzzy Inference Process

The initial FIS generation demands the selection of a number of linguistic variables, the membership functions type, the fuzzy model, the fuzzy rules and many others. FIS is established either by building a knowledge base based on fuzzy rules provided by an expert, or by fuzzy clustering techniques. In this paper, the two algorithms (FCM and SUBCLUST) are used to generate the corresponding FIS for the ANFIS model.

According to Fig.6, for each input vector, an ANFIS model is generated from an initial FIS. Then, the generated ANFIS model is evaluated during the learning phase by comparing its Root Mean Square Error (RMSE) to a RMSE threshold (RMSE_optimal). When the obtained RMSE is close to or equal to RMSE_optimal, the ANFIS evaluates the examples of the whole test. Otherwise, a readjustment of the input parameters will be established. Otherwise, the ANFIS process stops.

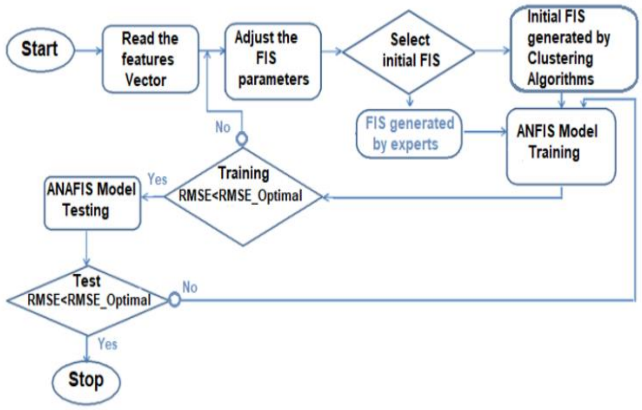

Fig. 6. ANFIS process.

D. Clustering Algorithms

The purpose of clustering is to identify natural groupings from a large data set to produce a concise representation of the data. Clusters are identified by applying either FCM or SUBCLUST within input/output training data. Therefore, the resulting cluster information is used to generate a Sugeno-type fuzzy inference system to model the data behavior.

Generally, the obtained Sugeno FIS best models the data behavior using a minimum number of rules. Its main properties are mentioned in Table-III.

Table-III: FIS Properties Description.

Sugeno FIS properties	
Output number	Single
Defuzzification Method	Weighted average
Output membership functions	Linear or constant and equals to the number of rules

- Fuzzy c-means

The Fuzzy c-means (FCM) algorithms, initially introduced by Jim Bezdek in 1981 [25], is a data clustering technique that belongs each data point to a cluster specified by a membership degree.

FCM was an improvement on earlier clustering methods. The FCM algorithm begins with an initial estimate for the cluster centers. Centers mark the mean location of each cluster. FCM assigns every data point a membership degree for each cluster.

Retrieval Number: 100.1/ijitee.A96281111121
DOI: 10.35940/ijitee.A9628.1111121
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

By iteratively updating the cluster centers and the membership grades for each data cluster centers to the right location within a data set. This iteration is based on minimizing an objective function J_m (equation (8)) that represents the distance from any given data point to a cluster center weighted by that data point's membership grade [25].

$$J_m = \sum_{i=1}^{D}\sum_{j=1}^{N}\mu_{ij}^{m}\left\|x_i - c_j\right\|^2 \quad (8)$$

Where:
D is the number of data points.
N is the number of clusters.
m is fuzzy partition matrix exponent for controlling the degree of fuzzy overlap, with m > 1.
Fuzzy overlap refers to how fuzzy the boundaries between clusters are, that is the number of data points that have significant membership in more than one cluster.
x_i is the i^{th} data point.
c_j is the center of the j^{th} cluster.
μ_{ij} is the degree of membership of x_i in the j^{th} cluster.
For a given data point, x_i, the sum of the membership values for all clusters is one.
The FCM performs the following steps during clustering:

1. Randomly initialize the cluster membership values, μ_{ij}.
2. Calculate the cluster centers by using equation (9).

$$c_j = \frac{\sum_{i=1}^{D}\mu_{ij}^{m}x_i}{\sum_{i=1}^{D}\mu_{ij}^{m}} \quad (9)$$

3. Update μ_{ij} according to equation (10).

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{N}\left(\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_k\right\|}\right)^{\frac{2}{m-1}}} \quad (10)$$

4. Calculate the objective function, J_m.
5. Repeat steps 2–4 until J_m improves by less than a specified minimum threshold or until after a specified maximum number of iterations.

As a result, a list of cluster centers and several membership grades for each data point are generated. These lists are then used to build a FIS whose membership functions represent the fuzzy qualities of each cluster.

- Subtractive Clustering

The subtractive clustering algorithm (SUBCLUST) is applied for estimating the number of clusters and the cluster centers for a set of data [27], when it is not clear how many clusters there should be for a given set of data. The returned clusters can be used to initialize iterative optimization-based clustering methods (FCM) and also model the ANFIS classifier. To reduce the number of rules, consider using the subtractive clustering method. Subtractive clustering assumes that each data point is a potential cluster center [28].
The algorithm does the following steps:

1. Calculate the likelihood that each data point would define a cluster center, based on the density of surrounding data points.
2. Choose the data point with the highest potential to be the first cluster center.
3. Remove all data points near the first cluster center. The vicinity is determined using cluster Influence Range.
4. Choose the remaining point with the highest potential as the next cluster center.
5. Repeat steps 3 and 4 until all the data is within the influence range of a cluster center.

E. Gradient Descent Optimization Algorithms

Training of an ANFIS structure is a special kind of optimization problem. Therefore, the gradient and evolutionary algorithms can be used to train an ANFIS structure. In other words, they are applied to tune its corresponding FIS membership parameters. In this study, after creating the initial FIS structure, the BP and the Hybrid learning algorithms are investigated to train the ANFIS.

- Backpropagation Algorithm

The BP algorithm is based on the Gradient Descent rule which tends to adjust the associated parameters with the input and output membership functions.
It can be summarized in four fundamental steps:

1. Initialize the connection weights with random values.
2. Compute the output of the ANN by propagating each input pattern through the network in a forward direction.
3. Compute the Mean Square Error E between the desired output O^d and the produced output O by the ANN via equation (11).

$$E = \frac{1}{2}\sum_{i=1}^{N}\left(O_i^d - O_i\right)^2 \quad (11)$$

4. Adjust the connection weights according to equation (12) where η is the learning rate and $\frac{\partial E}{\partial w(t)}$ is the gradient.
5. $$w(t+1) = w(t) - \eta\frac{\partial E}{\partial w(t)} \quad (12$$

The above process is repeated until a stopping criterion is met which can be a desired minimum error or a maximum number of iterations.

In the case of the BP learning rule, the central part concerns how to recursively obtain a gradient vector in which each element is defined as the derivative of an error measure with respect to a parameter.

- Hybrid Learning algorithm

Hybrid method consists of applying the (BP) for the parameters associated with the input membership functions, and least squares estimation for the parameters associated with the output membership functions. Initially, it is assumed that the input membership functions are symmetrically spaced over the entire universe of discourse. Accordingly, some initial values for the center and the spread of each input membership function are assumed, whereas, in the case of output for each rule,

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

all initial values are assumed to be zero. Then, the input parameters are optimized by the error BP algorithm and the output constants are optimized by least square method. The tuned ANFIS thus obtained is then used to obtain a stable output.

F. Evolutionary Methods

In this study, the evolutionary ANFIS training is investigated. An initial ANFIS structure is initially built. Then the (GA) and (PSO) algorithm is used to train the ANFIS.

- Particle Swarm Optimization

Particle Swarm Optimization (PSO) is an optimization algorithm, developed by Kennedy and Eberhart in 1995 [29]. It is based on a stochastic descent model which defines the social behavior of birds and fish. The stochastic descent begins with an initial solution. Then, it compares it to all its neighbors, keeping each time the best result. PSO algorithm saves the best previous local position and neighbors' temporary position. Unfortunately, it adjusts the current particle according to its best updated local position. Hence, the population can coalesce around one location, or can coalesce around a few locations, or can continue to move. Thus, it generally displaces the swarm to the global minimum.

The PSO algorithm steps are as follows.

1. Creating the initial particles and assigning them initial velocities.
2. Evaluation of the objective functions at each particle location.
3. Determination of the lowest function value and the best location.
4. Selection of new velocities based on the current velocity, the particles' individual best locations and the best locations of their neighbors.
5. Updating iteratively the particle locations (the new location is the old one plus the velocity, modified to keep particles within bounds), velocities, and neighbors.
6. Iterations proceed until the algorithm reaches a stopping criterion.

- Genetic Algorithms

The genetic algorithms are investigated for optimization and unsupervised learning problems. They simulate a fictional environment based on the theory of evolution [18]. The following outline summarizes how the genetic algorithm works:

1. The algorithm begins by creating a random initial population.
2. The algorithm then creates a sequence of new populations. At each step, the algorithm uses the individuals in the current generation to create the next population. To create the new population, the algorithm performs the following steps:

1. Scores each member of the current population by computing its fitness value. These values are called the raw fitness scores.
2. Scales the raw fitness scores to convert them into a more usable range of values. These scaled values are called expectation values.
3. Selects members, called parents, based on their expectation.
4. Some of the individuals in the current population that have lower fitness are chosen as elite. These elite individuals are passed to the next population.
5. Produces children from the parents. Children are produced either by making random changes to a single parent—mutation—or by combining the vector entries of a pair of parents—crossover.
6. Replaces the current population with the children to form the next generation.
7. The algorithm stops when one of the stopping criteria is met. See Stopping Conditions for the Algorithm.

An example of a binary chromosome is illustrated in Fig.7. It represents the Gaussian and rule base parameters.

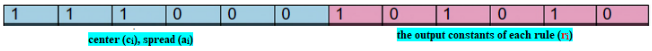

Fig. 7. Binary Chromosome Example.

III. EXPERIMENTAL WORK

In this work, our goal is to select the more appropriate ANFIS which reaches the reduced RMSE for classifying 48 ECG signals into 5 cardiac classes. In fact, eight (8) proposed ANFIS architectures are investigated. However, they are different from each other regarding the optimization or training algorithm and also the clustering approach adopted for the initial FIS generation. In fact, either the FCM or the SUBCLUST clustering algorithm is applied to generate the initial FIS. Moreover, either the gradient optimization algorithms or the evolutionary ones are applied to adjust the FIS membership and rule base parameters. Table- IV collects the configuration of the eight (8) proposed ANFIS. Subsequently, a comparison study between the proposed ANFIS results is discussed in order to select the more appropriate one for this arrhythmia task.

Moreover, the curve representing the difference between the desired and the computed output value for the selected ANFIS design will be analyzed and discussed. Among the eight ANFIS defined in Table- IV, one of them will be selected as the most accurate model.

Table- IV: the proposed eight ANFIS.

ANFIS	Initial FIS		Training Algorithm			
			Gradient		*Evolutionary*	
	FCM	*SUBCLUST*	BP	Hybrid	GA	PSO
ANFIS_1	X		X			
ANFIS_2	X			X		
ANFIS_3	X				X	
ANFIS_4	X					X
ANFIS_5		X	X			
ANFIS_6		X		X		
ANFIS_7		X			X	
ANFIS_8		X				X

Retrieval Number: 100.1/ijitee.A96281111121
DOI: 10.35940/ijitee.A9628.1111121
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

A. Initial FIS Structure

Before training the ANFIS model, an initial constant order Sugeno FIS structure has to be generated. It is obtained by using the FCM algorithm or the SUBCLUST one.

- FCM configuration

The FCM configuration is essentially based on the choice of the number of classes (Class_Number). Therefore, we've evaluated the influence of the number of classes on the RMSE value and the regression (R^2). The choice of the most efficient FIS corresponds to the minimum value of RMSE and maximum value of R^2.

Table- V: FCM Configuration.

Class_Number	RMSE	R^2
10	0.120	0.610
9	**0.010**	**0.740**
8	0.020	0.770
7	0.010	0.400
6	0.010	0.520
5	0.050	0.240
4	0.030	0.170
3	0.140	0.140
2	0.210	0.390

Referring to Table- V, it is clear that the choice of nine (9) classes generates the most accurate FIS. In fact, it reaches the lowest value of the RMSE (RMSE = 0.01) and the closest regression value (R^2 =0.74).

- SUBCLUST configuration

The configuration of SUBCLUST consists of adjusting a set of parameters including the radius of the class (Radi), the maximum number of iterations, the desired error, the initial step value, the step size increment rate (Δ+) and decrement rate (Δ-).

➢ Radi parameter

Table- VI shows that when the value of Radi varies from 0.3 to 1, the value of RMSE remains unchanged, the number of fuzzy rules (Rule_Number) of the generated FIS decreases from 37 to 15 and the rate of the regression (R^2) is unstable in the interval [0.61 0.92]. Consequently, the value 0.7 of the radius of the class was considered as the most adequate value, since the FIS reaches a maximum regression value (R^2= 0.92). Therefore, we notice that when the radius of the class is small, many small classes and many rules will be generated. However, when it is large, a few large classes and fewer rules will be generated.

Table- VI: Radi impact.

Radi	RMSE	Rule_Number	R^2
0.30	0.00	37	0.81
0.40	0.00	37	0.73
0.50	0.00	37	0.82
0.60	0.00	35	0.91
0.70	**0.00**	**29**	**0.92**
0.80	0.00	24	0.69
0.90	0.00	16	0.61
1	0.00	15	0.76

➢ Initial step

According to Table-VII, the initial step size (0.010) is selected regarding the reduced regression value (R^2=0.920).

Table- VII: Initial Step Impact.

Radi	Initial Step	RMSE	Rule_Number	R^2
0.70	**0.010**	**0.0**	**29**	**0.920**
	0.100	0.00	27	0.720
	1	0.00	33	0.540

➢ Increment and decrement steps

Table- VIII presents the influence of the incrementation (Δ +) and decrementation (Δ-) steps of the initial step on the performance of the generated FIS.

Table- VIII: Increment and decrement initial impact.

Radi	Intial step	(Δ -)	(Δ +)	RMSE	Rule_ Number	R^2
0.70	**0.010**	**0.900**	**1.100**	**0**	**31**	**0.920**
		0.800	1	0	29	0.640
		0.700	0.900	0	29	0.880
		0.600	0.800	0	31	0.810
		0.500	0.700	0.010	29	0.590
		0.400	0.600	0	27	0.760
		0.300	0.500	0	30	0.760
		0.200	0.400	0	31	0.660
		0.100	0.300	0	32	0.590

By referring to Table- VIII, it is deduced whatever the adjusted values of the two parameters (Δ-) and (Δ+), RMSE remains zero, but the regression value deteriorates. As a consequence, we kept the default values ((Δ-) = 0.9 and (Δ+) =1.1) which generate the highest regression value (R^2 = 0.92).

IV. DISCUSSION

Hence, we have concluded that the performance of the generated initial FIS by the FCM algorithm depends on the input vector. Moreover, the FCM algorithm has certain limits, namely, the random selection of centers, the iterative reasoning of the algorithm likely to converge towards local minima and the need to know beforehand the number of classes. Therefore, to overcome these limits and improve the obtained performance, the SUBCLUST algorithm which automatically selects the appropriate number of classes, could be better used to generate the initial FIS.

As a result, we have eliminated the four configured ANFIS using the initial FIS by FCM algorithm.

We will focus on the most accurate training algorithm in the case of using an initial FIS configured by the SUBCLUST algorithm.

A. Designing ANFIS

In order to design the ANFIS model, many steps have to be done. Firstly, we have selected the training, checking and testing data. Training data presents 70% of the entire selected dataset. The checking data presents 30% of the training data and the testing data consists of 30% of the entire selected dataset. It is an additional set that is not used for training or checking data.

Then, the initial FIS has to be chosen, either which is generated by the (FCM) or the (SUBCLUST) algorithm. In this study, we have chosen the configured initial FIS by SUBCLUST algorithm. Afterwards, we have adjusted the number of fuzzy sets on three (3) for all input variables. We have also selected the input and output membership functions.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

In fact, Gaussian membership functions and constants are used to define the degree of membership of the input variables and the output, respectively. Finally, the optimization approach has to be selected among four different optimization approaches, including, BP, Hybrid, GA and PSO. They are used in order to optimize the Gaussian input membership functions parameters (center (c_i) and spread (a_i)) and the output constants of each rule (r_i).

B. Training ANFIS

The obtained FIS were used to generate ANFIS models. These models are trained by applying two optimization categories. They consist of the Gradient descent and evolutionary algorithms.

The first category includes the (BP) and (Hybrid) algorithms. However, the second category contains the GA and the PSO. In fact, the training process of these algorithms tunes the membership function parameters associated with the input and output of the generated initial FIS. Therefore, regarding the optimization algorithm, several ANFIS models were analyzed in this study. Their properties description is resumed in Table- IX.

Table- IX: ANFIS Properties Description.

ANFIS models properties	
Initial FIS	SUBCLUST
Optimization Method	BP, Hybrid, PSO, GA
Number of training epochs	50
Training error (RMSE)	10-6
Initial Step size	0.010
Increment Step-size rate	0.9
DecrementStep-size rate	1.1

V. RESULTS AND DISCUSSION

During the learning and testing phases, the Mean Square Error (MSE) will be determined by calculating the difference between the predicted ANFIS output and the desired value. Next, the RMSE will be computed by finding the square root of the MSE value. In addition, the mean of the MSE (Mean_error) and its standard deviation (Std_error) will be assessed. The ANFIS model is considered efficient when its errors (MSE and RMSE) are low.

A. Backpropagation Training ANFIS

In this section, BP and Hybrid optimization algorithms are applied to optimize the parameters of the inputs membership functions and the constants of the output, by minimizing the RMSE cost function. The learning and testing performances for each algorithm are presented in Table- X. As it is shown in Table- X, BP optimization algorithm underperforms the Hybrid one. It gives the best performances during learning (MSE=2.35e-14, RMSE=1.53e-07, std-error=1.47e-07) and testing phases (MSE=0.3163, RMSE=0.1473, std-error=0.09).

Table- X: Backpropagation Training ANFIS performances.

	Data	MSE	RMSE	Mean_error	Std-error
BP	Learning	0.0015	0.0387	0.0060	0.9972
	Testing	0.1932	0.4396	- 0.0848	0.4359
Hybrid	Learning	**2.35e-14**	**1.53e-07**	**4.73e-08**	**1.47e-07**
	Testing	**0.3163**	**0.1473**	**-0.0548**	**0.09**

B. Evolutionary Training ANFIS

In this section, GA and PSO algorithms are applied to optimize the Gaussian input membership functions parameters (center (c_i) and spread (a_i)) and the output constants of each rule (r_i). The optimization uses the RMSE as a cost function [30].

Their properties are summarized in Table- XI and Table-XII, respectively.

Table- XI: GA Properties.

GA properties	
Initial FIS	SUBCLUST
Maximum Iteration	1000
Number of population	25
Cost function	RMSE
Crossover Percentage	0.4
Mutation Percentage	0.7
Mutation Rate	0.15
Selection Pressure	8

Table- XII: PSO Properties.

PSO properties	
Initial FIS	SUBCLUST
Maximum Iteration	1000
Number of population	25
Cost function	RMSE
Inertia weight	1
Inertia Weight Damping Ratio	0.99
Personal Learning Coefficient	1
Global Learning Coefficient	2

The learning and testing performances are presented in Table- XIII. It shows that the ANFIS model optimized by the PSO is powerful during learning. Indeed, during this phase, the error is minimal (MSE = 0.009, RMSE = 0.098) and its standard deviation is very reduced (0.098).However, during the test phase, the most effective optimization method is the GA (MSE=0.424, RMSE=0.651, Std-error=0.678).

Table- XIII: Evolutionary Training ANFIS performances.

Algorithm	Data	MSE	RMSE	Mean_error	Std-error
GA	Learning	0.027	0.165	0.017	0.167
	Testing	**0.424**	**0.651**	-0.002	**0.678**
PSO	Learning	**0.009**	**0.098**	-0.017	**0.098**
	Testing	2.587	1.60	0.812	1.445

C. Comparison study between ANFIS

In order to highlight the advantages of the proposed approach in this paper, a comparative study was carried out by comparing the obtained ANFIS performances. Table- XIV summarizes the performances of the two selected ANFIS models with all data of the selected dataset (10x48). ANFIS models are which use the Hybrid and GA for the optimization task.

Table- XIV: Comparison study between ANFIS performances.

Data	Algorithm	MSE	RMSE	Mean_error	Std-error
All data	Hybrid	0.1932	0.4396	-0.0848	0.4359
	GA	**0.0424**	**0.2060**	**0.0383**	**0.2046**

Retrieval Number: 100.1/ijitee.A96281111121
DOI: 10.35940/ijitee.A9628.1111121
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Referring to Table- XIV, it is clear that ANFIS model which uses the GA (MSE = 0.0424, RMSE = 0.2060, Std-error= 0.2046) to optimize its FIS parameters is most accurate than which uses the Hybrid algorithm, when we evaluate all the feature dataset examples. However, its performances can be improved further by applying two major modifications. The first one consists of modifying the FIS properties, such as, the FIS type, the number of inputs, the number of linguistic variables, and the types of membership functions. The second one is to increase the number of iterations which increases the duration of the optimization process, but can be a good solution. However, this alternative can increase the validation error due to the over-fitting of the ANFIS model with the training data.

The learning and testing error curves of the ANFIS, optimized by the GA, are described in Fig.8 and Fig.9.

Fig.8 illustrates the difference between the predicted outputs (Output) of the model and those desired (Target) during the learning phase. From which, we noticed that the two outputs are similar. This proves the accuracy of the learning phase. Fig.9 describes the difference between the predicted outputs (Output) of the SUBCLUST_ANFIS model and those desired (Target) during the test phase. From which, it is clear that the curve of the predicted output does not perfectly follow that of the desired output for some examples of the whole test.

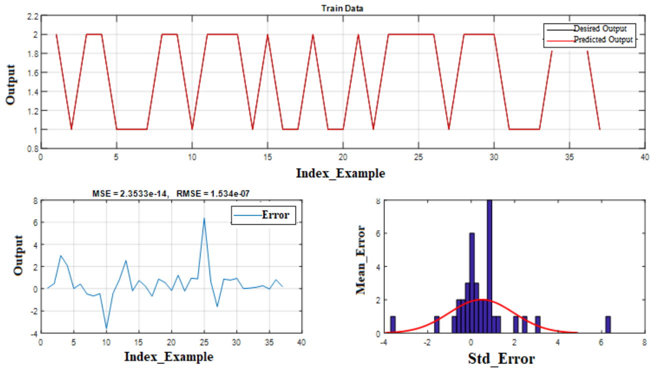

Fig. 8. Genetic training performances.

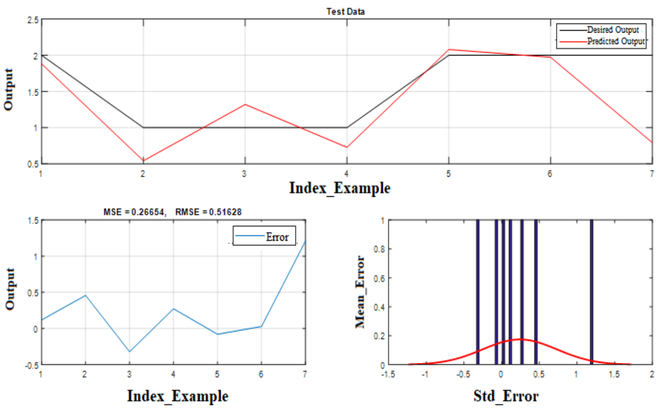

Fig. 9.Genetic testing performances.

D. Discussion

Hence, we have concluded that the performances of the ANFIS model trained by the GA are better than those obtained by the PSO algorithm. Nevertheless, the effective ANFIS configuration depends on several choices that subjectively influence its performances. Indeed, it is essential to choose the input FIS, the assessment sets (for learning, validation and testing), and the parameters of the GA. The initial FIS is in turn dependent on certain major design parameters such as the input and output universe range, the number of linguistic terms, the type of the membership function, and the premises of the rules, the consequences of the rules, the logical operators and the defuzzification methods.

E. Comparative study with related works

A comparative study is also done to evaluate the effectiveness of this paper. It is summarized in Table– XV. It recapitulates the obtained MSE of five similar works, including the current work and other classifiers [18-19- 20-24]. The common subject is the use of machine learning algorithms (ANN, FIS, and ANFIS) and time-scale features for ECG arrhythmia classification. However, each work has a number of input features, output classes, and a configured classifier according to its proper methodology. In fact, in our case of study, selected features by PCA algorithm are used to classify ECG recordings into five classes. Consequently, all these works are effective to the tune of about (0.250 e-30 to 0.383) of MSE value. Besides, the proposed work is considered as an acceptable method with a reduced MSE value 0.042. Hence, the precision depends mostly on the input feature vector as well as the configuration of the machine learning algorithm.

Table- XV: Comparative study with related works.

Reference	**Characteristics**	**Classifier**	**MSE**
[18]	6 input feature 5 output	FIS	0.383
[20]	10 input feature 2 output	Probabilistic Neural Network (PNN)	0.162
		Radial Basis Function (RBF)	**0.250 e-30**
		Multilayer Perceptron (MLP)	0.064
[19]	62 input feature 5 output	MLP	0.027
[24]	10 input feature 2 output	MLP	0.045
Current Study	10 input feature 5 output	**ANFIS**	**0.042**

VI. CONCLUSIONS

In this work, a Neuro-Fuzzy model is selected for the configuration of a data driven system based on an arrhythmia classification task. The model is able to classify ECG recordings into five heartbeat classes. However, classification requires an input feature vector. Indeed, time-scale features were extracted by a wavelet transformation and were selected by applying static metrics and PCA algorithm. The obtained input feature vectors were used by FCM and SUBCLUST algorithms to generate the initial FIS. Hence, we have concluded that the SUBCLUST algorithm is better to generate the initial FIS. Then, we have saved the initial FIS obtained by the subtractive algorithm.

Retrieval Number: 100.1/ijitee.A96281111121
DOI: 10.35940/ijitee.A9628.1111121
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

However, to design the Neuro-Fuzzy model, it is essential to choose the most relevant training algorithm. Therefore, four optimization algorithms, including, the BP, Hybrid, PSO and GA were evaluated. They are mainly studied in order to optimize the Gaussian input membership functions parameters (center (c_i) and spread (a_i)) and the output constants of each rule (r_i). Hence, we have obtained that the best performances are produced by the ANFIS which applies the GA as a training algorithm. However, the curve of its testing error is not ideal. It demands to adjust the initial FIS properties, such as the number of linguistic variables, the membership function type, the input feature vectors and many others. Accordingly, the application of an automated method for the optimization of these parameters is of major interest and is envisaged.

REFERENCES

1. Yadegaridehkordi, Elaheh, et al. "Influence of big data adoption on manufacturing companies' performance: An integrated DEMATEL-ANFIS approach." Technological forecasting and social change 137 (2018): 199-210.
2. Semenova, Olena O., et al. "Genetic ANFIS for scheduling in telecommunication networks." Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2018. Vol. 10808. International Society for Optics and Photonics, 2018.
3. Mir, Mahdi, et al. "Applying ANFIS-PSO algorithm as a novel accurate approach for prediction of gas density." Petroleum Science and Technology 36.12 (2018): 820-826.
4. Husein, A. M., et al. "The implementation of two stages clustering (k-means clustering and adaptive neuro fuzzy inference system) for prediction of medicine need based on medical data." Journal of Physics: Conference Series. Vol. 978. No. 1. IOP Publishing, 2018.
5. Alizadehsani, Roohallah, et al. "Handling of uncertainty in medical data using machine learning and probability theory techniques: A review of 30 years (1991–2020)." Annals of Operations Research (2021): 1-42.
6. Thakkar, Harshil, et al. "Comparative anatomization of data mining and fuzzy logic techniques used in diabetes prognosis." Clinical eHealth 4 (2021): 12-23.
7. Mittal, Kanika, et al. "A comprehensive review on type 2 fuzzy logic applications: Past, present and future." Engineering Applications of Artificial Intelligence 95 (2020): 103916.
8. Ahmadi, Hossein, et al. "Diseases diagnosis using fuzzy logic methods: A systematic and meta-analysis review." Computer Methods and Programs in Biomedicine 161 (2018): 145-172.
9. Ojha, Varun, Ajith Abraham, and Václav Snášel. "Heuristic design of fuzzy inference systems: A review of three decades of research." Engineering Applications of Artificial Intelligence 85 (2019): 845-864.
10. Kumarganesh, S., and M. Suganthi. "An enhanced medical diagnosis sustainable system for brain tumor detection and segmentation using ANFIS classifier." Current Medical Imaging 14.2 (2018): 271-279.
11. Bali, Bulus, and Etemi Joshua Garba. "Neuro-fuzzy Approach for Prediction of Neurological Disorders: A Systematic Review." SN Computer Science 2.4 (2021): 1-11.
12. Alabool, Hamzeh, et al. "Artificial intelligence techniques for containment COVID-19 pandemic: a systematic review." (2020).
13. Kour, Haneet, Jatinder Manhas, and Vinod Sharma. "Usage and implementation of neuro-fuzzy systems for classification and prediction in the diagnosis of different types of medical disorders: a decade review." Artificial Intelligence Review 53.7 (2020): 4651-4706.
14. Kar, Samarjit, Sujit Das, and PijushKanti Ghosh. "Applications of neuro fuzzy systems: A brief review and future outline." Applied Soft Computing 15 (2014): 243-259.
15. Lyon, Aurore, et al. "Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances." Journal of The Royal Society Interface 15.138 (2018): 20170821.
16. Chen, Zhijia, et al. "Self-adaptive prediction of cloud resource demands using ensemble model and subtractive-fuzzy clustering based fuzzy neural network." Computational intelligence and neuroscience 2015 (2015).
17. Chen, Wei, Mahdi Panahi, and Hamid Reza Pourghasemi. "Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling." Catena 157 (2017): 310-324.
18. Lassoued, Hela, and Raouf Ketata. "Genetic fuzzy logic based system for arrhythmia classification." Control Theory in Biomedical Engineering. Academic Press, 2020. 105-127.
19. Lassoued, Hela, and Raouf Ketata. "ECG multi-class classification using neural network as machine learning model." 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET). IEEE, 2018.
20. Lassoued, Hela, Raouf Ketata, and Slim Yacoub. "ECG decision support system based on feedforward neural networks." International Journal on Smart Sensing and Intelligent Systems 11.1 (2018): 1-13.
21. Moody, George B., and Roger G. Mark. "The impact of the MIT-BIH arrhythmia database." IEEE Engineering in Medicine and Biology Magazine 20.3 (2001): 45-50.
22. Lassoued, Hela, and Raouf Ketata. "Hybrid Two Stage Neuro Genetic System for Arrhythmia Diagnosis." IJCSNS 18.9 (2018): 31.
23. Shen, Tsu-Wang, W. J. Tompkins, and Y. H. Hu. "One-lead ECG for identity verification." Proceedings of the second joint 24th annual conference and the annual fall meeting of the biomedical engineering society][engineering in medicine and biology. Vol. 1. IEEE, 2002.
24. Lassoued, Hela, and Raouf KETATA. "Artificial Neural Network classifier for heartbeat arrhythmia detection." Proceedings of Engineering and Technology-PET. Vol. 22. 2017.
25. Ringnér, Markus. "What is principal component analysis?." Nature biotechnology 26.3 (2008): 303-304.
26. Association for the Advancement of Medical Instrumentation. Diagnostic electrocardiographic devices. Association for the Advancement of Medical Instrumentation (AAMI), 1986.
27. Buse, Richard, Zhi-Qiang Liu, and Jim Bezdek. "Word recognition using fuzzy logic." Ieee transactions on fuzzy systems 10.1 (2002): 65-76.
28. Dhanachandra, Nameirakpam, KhumanthemManglem, and YambemJinaChanu. "Image segmentation using K-means clustering algorithm and subtractive clustering algorithm." Procedia Computer Science 54 (2015): 764-771.
29. Kennedy, James, and Russell Eberhart. "Particle swarm optimization." Proceedings of ICNN'95-international conference on neural networks. Vol. 4. IEEE, 1995.
30. Mostapha Kalami Heris, Evolutionary ANFIS Training in MATLAB (URL: https://yarpiz.com/319/ypfz104-evolutionary-anfis-training), Yarpiz, 2015.

AUTHORS PROFILE

Hela Lassoued, was born in Gafsa, Tunisia in 1987. She received her Ph.D degree in Industrial Computing in 2020 and her Degree in Industrial Computing and Automatic Engineer in 2013, from the “Applied Science and Technology Institute (INSAT), Carthage University, Tunisia”. Her current research interests are in the Machine Learning fields especially the Artificial Neural Networks, the Fuzzy Logic Systems and the Neuro Fuzzy Systems.

Raouf Ketata, was born in Sfax-Tunisia in 1961. He received the Engineering Diploma in Electrical Engineering and Instrumentation from the National School of Engineer (ENIS) in Sfax-Tunisia in 1987, the Master degree in Automatic, Industrial Computer Science and Signal Treatment in 1988, the Ph.D. degree in Automatic and Industrial Computer Science in 1992 from National Institute of Applied Science (INSA) in Toulouse-France and the University Habilitation in Industrial Computer Science from the National Institute of Applied Science and Technology (INSAT) in Tunis-Tunisia. From 1988 to 1992, he had a Contract Assistant Post in National Institute of Applied Science (INSA) in Toulouse-France. From 1992 to 1994, he had a Conference Professor Post in the Atlantic Superior School in Electrical Engineering (ESA IGELEC) in Saint-Nazaire-France. Between 2010 and 2012, he was appointed Director of the department of Physic and Instrumentation Engineering at the National Institute of Applied Science and Technology (INSAT) in Tunis – Tunisia. From 2018 to 2020, he was elected Director of the National Institute of Applied Science and Technology (INSAT) in Tunis – Tunisia. He is currently a Conference Professor in the National Institute of Applied Science and Technology in Tunis- Tunisia and member of the Remote Sensing and Spatial Reference Information Systems (LTSIRS) laboratory.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

His current research interests lie in fuzzy logic, neural networks, genetic algorithms, intelligent systems and their applications in machine learning, classification, and intelligent control.

Hajer Ben Mahmoud, was born in Tunis, Tunisia in 1979. She received the Engineering Diploma in industrial computer science and automatic, the Master degree in industrial computer science and automatic in 2006 and the Ph.D degree in Industrial Computing in 2013 from the National Institute of Applied Science and Technology (INSAT) Tunis-Tunisia. Currently, she occupies a post of Assistant Professor at the National Institute of Applied Sciences and Technology of Kairouan -Tunisia. Her current research interests lie in fuzzy logic, genetic algorithms, intelligent systems modeling, and control.