A Review of the Methods for Sudden Cardiac Death Detection

A Guide for Emergency Physicians

Mohammad Karimi Moridani (✉),Shahrzad Marjani
Islamic Azad University, Tehran, Iran
karimi.m@iautmu.ac.ir

Abstract—Sudden cardiac death is an unexpected death of a person with or without knowing cardiac causes are often occurring in less than an hour after the incidence of symptoms. The purpose of this paper is to examine different methods for predicting sudden cardiac death using the ECG signals from 1998 to recent years.

In this paper, studies using various methods to detection sudden cardiac death that has applied the data from the Physionet and MIT-BIH databases with a sampling frequency of 256 samples per second are reviewed. In the field of SCD prediction, various studies have addressed the processing of the ECG signal as well as the HRV signal in different domains, including time, time-frequency, and nonlinear domain.

To evaluate the results of the proposed methods, each of the researchers analyzed the a-few-minute intervals before the SCD. Different classification methods are available to identify the efficiency of the proposed algorithm. The use of features introduced in different domains and different classifiers has led to the observation of different horizons of prediction in various studies. Accordingly, the most prominent of these evaluations is the mixture expert methodology in which the best feature extraction methods are used in a new method for selecting the optimal feature space locally. This method makes it possible to select different features every minute before the event by choosing the optimal features for each one-minute interval of the signal as an episode which increases the prediction time from 4 minutes before the death to 12 minutes and allows the interpretation of clinical symptoms in terms of multiplication of the presence of the features per minute.

The analysis of various studies shows that by approaching the time of death, linear features (time and frequency) can be predictive of death according to the sensible behavior and variations in patients' signal.

Keywords—Sudden Cardiac Death, Heart Rate Variability, Signal Processing, Artificial Neural Network, Prediction Horizon.

1   Introduction

Sudden cardiac death (SCD) is an unexpected event that occurs in the short term, even in people without heart disease. One of the classes which are subject to this risk is the athletes. The most common causes of sudden cardiac death in athletes are hypertrophic cardiomyopathy, anomaly coronary artery [1-3]. In general, this is the most common symptom of cardiovascular disease. Although the underlying cause of SCD is not known, ventricular fibrillation (VF) is considered in 20% of the SCD mechanism which disables the heart unable to pump in blood, resulting in the loss of consciousness and takes the life of the patient within a few minutes. Therefore, the initial prediction of this unexpected risk is very important for people with Ventricular Fibrillation (VF) for timely treatment and increased survival [4-5].

In recent years, extensive studies have been conducted on SCD diagnosis based on the electrocardiogram (ECG) signal or the heart rate variability (HRV), which has been done by feature extraction in linear, nonlinear, time, and time-frequency processing areas and then Individuals are categorized through categories based on the list of extracting features [6-10]. Researchers around the world are testing the heart rate variability and electrocardiogram signal features to identify and select unique features by which they can predict the risk of SCD [11].

Bioelectric signals represent the activities of various organs of the human body. The ECG is one of the most important signals among the bioelectric signals indicating the electrical activity of the heart. The ECG signal contains important information on the heart function and condition, and therefore it serves as an essential tool in the early diagnosis of any disease or heart failure used by doctors [12].

The HRV is variations in the consecutive time intervals of the heart, in other words, it is the time between two R waves in two consecutive heart cycles in the electrocardiogram signal that the variations in HRV time series can be considered as a sign of risk for death after a heart attack [13].

Also, various algorithms such as Wavelet transform, Fourier transform, fundamental component analysis, and artificial neural network are used to analyze and categorize the extracted ECG signal features in the linear, nonlinear, and frequency domain to predict SCD that the studies conducted in this field are discussed below. Voss et al. (1998) extracted the amplitude and frequency parameters from the signal of 26 heart patients in high-risk and low-risk groups. The researchers, using a combination of nonlinear methods and normalized entropy and stepwise discriminant function, divided the patient into two groups.

Similarly, another study was conducted on 325 individuals aged 65 and older who had a history of the disease in 2001 by Huikari et al. [43]. In this study, patients were examined round the clock under the holter. The extracted fractal features and heart rate (HR) were analyzed. During these 10 years follow up, 164 people died, and 71 of them died for cardiac reasons that using the fractal analysis, 95% of cases were predicted, and indicated that short-term fractal changes and HRV reflect the risk of SCD.

Maria Teresa La Rovera et al. (2003) used the data of 202 consecutive patients between 1991 and 1995 with chronic heart infarction (CHF) [42]. These researchers ex-

tracted the HRV parameters of the time and frequency domains of the eight ECGs recorded during respiration control. In these specimens, they measured the low-frequency power (LFP) and the end-diastolic diameter in the validation sample they found that when the LPF reaches 83 milliseconds during the repertory control of 11 ms and the end-diastolic diameter reaches 83 milliseconds, the premature ventricular contraction occurs and accordingly, the SCD can be predicted.

Shen et al., (2007) [14] proposed a method to predict SCD using the violet analysis with LabView software. In this study, in addition to the violet analysis of the artificial neural network, an ECG signal was used.

By Ma'ida Bahrami (2012) [15] conducted a study titled "Predicting SCD with intelligent hybrid systems," Using hybrid methods for feature extraction and combining multiple neural networks to classify data, the system's performance has increased greatly. In this method using the Wavelet transform the extracted non-linear features and HRV signal time-frequency are combined with the features obtained by two non-linear Poincaré and detrended fluctuation analyses and after applying the dimensionality reduction method for increasing the analysis speed of the network the vector of reduced features has been applied to the MLP for classifying the data.

Manis et al. (2013) [44] using the machine learning technique as a powerful tool for classifying the risk of arrhythmia discriminated two groups of high and low-risk patients. Then the HRV was calculated and considered as two inputs in the SVM, and RF categorization and the results showed that the automatic classification of the two groups is possible.

The nonlinear features were extracted from the ECG signal and the use of discrete wavelet transform in Acharya et al. (2014) [9]. In this research, the main focus of the authors was on the classification of the extracted nonlinear signal features by a k-nearest neighbor, support vector machine, and the decision tree that is using their combination with discrete wavelet transform they predicted within 4 minutes before its occurrence.

These other researchers presented another study titled predicting SCD using a heart rate signal. In the previous method, they suggested SCDI (sudden cardiac death index) that could predict the SCD within four minutes before the onset of the event, which was done using the ECG signal, but in the new method, SCD is studied by HRV signal [16]. Similarly, another study was conducted by these researchers in 2017 [17]. In this paper, they addressed automatic SCD prediction using the recurrent quantification analysis (RQA), and using Kolmogorov parameters, and they studied the HRV signal feature. They classified these features by t-test, k-nearest neighbor, decision tree (DT), and the SVM. The performance of this proposed system has been improved by adding more features and stronger classifications.

Jennifer Sheila et al (2014) from the Chennai Lopula Institute of Technology in their research titled "prediction of SCD by SVM" showed that by extracting the HRV features in the time and frequency domain from the ECG signal using the SVM classification and the kernel function the healthy people can be discriminated from the SCD patients [28].

Elias Ebrahimzadeh and Mohammad Pouyan (2011), in their recent study titled "prediction of SCD using electrocardiogram signal processing," have predicted this incident [7-10]. To do this, after producing the HRV signal from the ECG signal, they extracted

the time-frequency domain and nonlinear features. In the next step, to create the most distinction between the two classes, the best feature is chosen and then by applying the principal component analysis (PCA) on the hybrid feature vector, the dimensions of the feature are reduced and eventually the healthy people and the people at risk are classified by MLP.

To evaluate the ability of each analytical method to discriminate individuals, they are compared separately and in combination. On the other hand, these other researchers presented another study titled "prediction of SCD using mixture experts" in 2017 [18]. Due to the existence of different features in different domains, the mixture expert classifier is suggested as a new method for selecting the optimal feature space locally. In this method, each one- min interval of the signal, different features are selected every minute before the event. This increases the prediction time for SCD from 4 minutes to 12 minutes and allows understanding the clinical signs about the multiplication of features.

Hooshyarifar et al. (2014) studied ECG signal classification in the nonlinear domain [19]. By Poincaré feature extraction plot and mapping HRV signal inspection, they predicted SCD 5 minutes before with an accuracy greater than 92%. At this stage, they extracted four features of the inspection chart and three features of the Poincare plot as the rate of inspection, the algebra, entropy, and the mean length of the diagonal line and standard deviation. In the next step, they reduced these features into one by linear discrimination analysis and finally classified the HRV signal using the KNN and SVM.

Mororapan et al. (2015), in their study titled "SCD prediction using the machine learning algorithm from Prince University of Malaysia," showed that by extracting the time-domain feature extraction, one could predict the SCD [20]. In this way, a set of features was extracted in the time domain, and the KNN and fuzzy algorithms were used to classify these data that the maximum mean classification of fuzzy classification and KNN showed that the HRV time-domain features are more efficient to predict SCD at five minutes before its onset. The researchers also presented another article in 2014 [41]. The researchers predicted SCD two minutes before its onset using the discrete wavelet transform method. First, they studied the HRV signal before processing in the time and frequency and nonlinear domain and extracted a total of 34 features from each sample. Finally, they used SVM and PNN for the prediction that their prediction accuracy was 96.36% and 64.93%, respectively. Another paper on SCD linear and nonlinear domain is presented by Reeta Devi et al. In this paper, two groups of SCD and normal patients have been included. The first group and the second group of SCD patients are studied and analyzed by ANOVA method of extracted features one and two hours before the incidence, respectively. Then the data were collected using KNN and SVM classifiers, and the results were analyzed [21].

Mir Hosseini et al. (2016) attempted to improve the accuracy of the early diagnosis of SCD using DT and SVM. In this paper, 22 features are obtained by TreeBagger algorithm and used as inputs to the SVM neural network, and ultimately highly accurate results were obtained. The accuracy of an SVM without significant features was 70.9%, but the SVM accuracy with important features reached 83.24%. [22].

Annmarie G. Raka et al. (2017) predicted SCD using electrocardiogram markers. In this method, features such as heart rate and RR intervals were extracted and classified

using linear discriminant analysis (LDA) and SVM analysis, and then the ideal time interval for SCD was determined [23].

The rest of the paper is structured as follows. The second section describes the database used in articles related to SCD, along with suggested methods in this area. In the third section, all classification systems that have been used in MATLAB software so far are presented, and in Section IV, the summary, accuracy, and validity of all proposed methods will be presented from 1998 to recent years.

2 Materials and Methods

2.1 Database

Most of the papers presented in the SCD prediction have used the data in the Physionet and MIT-BIH database, which has a sampling frequency of 256 samples per second. Data of both genders have normal and abnormal sampling label and the data recording time varies between several seconds to several minutes. In this database, the electrocardiogram signal is taken from twelve men and eleven women within the age of 17-89 who have a type of heart disease [24, 25].

2.2 Proposed method

In the field of SCD prediction, various studies addressed the processing of the ECG and HRV signals in different time-frequency and nonlinear domains. Similarly, they first initiated signal pre-processing, extraction of classical features, and selection of appropriate features. In the time domain process, in general, the statistical features of the time signal, such as the mean and standard deviation of heart rate, the standard deviation of RR intervals, and Root Mean Square of the Successive Differences (RSSD) are used. Also, in the frequency domain using the power spectral density (PSD) method, the signal energy is analyzed in very low frequency, low frequency and, high-frequency bands. Similarly, in the nonlinear domain, features of the Poincare plot, Detrended fluctuation analysis (DFA), common entropy, wavelet transform coefficients, and the features of the recursive graph including Lmax, Lmean, correlation dimension, etc. are used. In all of the proposed algorithms so far, researchers have been trying to alert SCD in a larger interval than the time of death by discriminating the signals into different time periods. Table 1 shows lists of the various features that have been extracted by researchers to achieve a proper result in time, frequency, linear, and nonlinear domains.

Table 1. Feature extraction in different domains

Feature Domain	Data type	Extracted features	Abbreviated features
Time	HRV Signal	Standard deviation	SDNN
		RR means of differences	MNN
		Root Mean Square of the Successive Differences	RMSSD
		Standard Deviation of Successive Differences	SDSD
		Percentage of RR Differences	PNN
Frequency	HRV signal	Estimation of the spectrum, low frequency, very low frequency, and high-frequency bands	PSD,LF,VLF,HF
Non-Linear	HRV signal	Discrete wavelet transform, Point dimension, Detrended fluctuation analysis, Poincare, Hurst, Fractal dimension, Correlation dimension, Sample Entropy, Approximate Entropy	DWT,PD,DFA,POINCARE,H,FD,CD,SampEnt,ApproxEnt
Linear	HRV signal	Geometric method,Statistical methods.Time-frequency methods	Geometric method,Statistical methods,Time-frequency methods

2.3   Signal preprocessing

This section discusses a common pre-processing method that has been reported in many articles. The ECG signal is discriminated at 1-minute intervals before the event. Fig. 1 has presented the electrocardiogram signal of a 42-year-old man who was exposed to sudden cardiac death from two minutes before the moment of cardiac death until some moments after it. The apparent shape of the signal in a pre-incident condition for a person with SCD is not significantly different from that of a healthy person. Similarly, the ECG signals of a healthy person and the one who have been affected by the incident are extracted in a minute before the incident and then it is used to remove noise from the city's electricity by Notch filter and low-frequency disturbances caused by breathing will also be eliminated using the moving average filter. In the following steps, the wave R is obtained by the Pan and Tompkins algorithm [26], and then by putting the spaces of the R wave together, the HRV signal is formed, and these pre-processed signals are used to extract features.
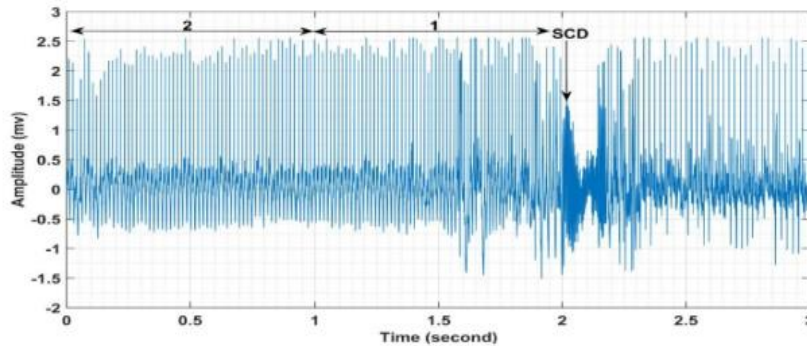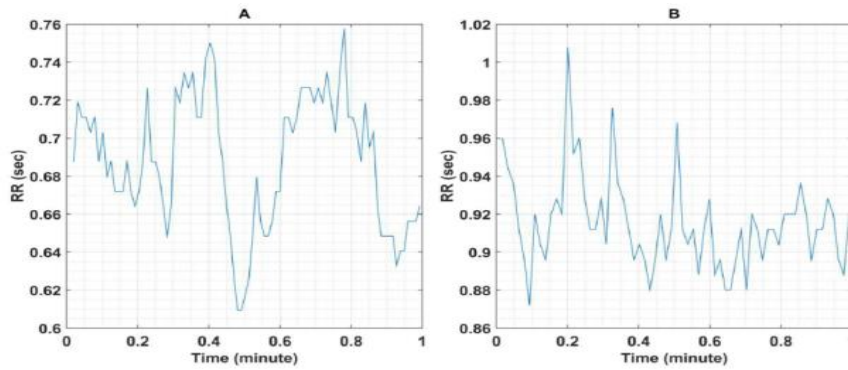
Fig. 1. The individual's ECG signal at the onset of SCD from 2 minutes before the event to some moments after the event [1]

In many cases, the ECG signals of the healthy person and the patient do not differ greatly. This is the case when the HRV signals of these two groups are different, as shown in Fig. 2.

a) One-minute HRV signal of a healthy person

b) The same person's HRV signal in one minute before SCD [1]

Fig. 2. Reducing the Feature Dimensions

Feature selection is one of the important issues in identifying healthy people and patients. Among the features extracted from the signal, the optimal selection of features so that they can achieve better results in the field of detection is very challenging. On the other hand, the removal of features that are not so important or have shared information with other features can reduce the feature vector dimensions, which can help reduce the complexity of computing and speed up the system to achieve the desired response [27]. Generally, feature dimension reduction algorithms are classified into two transformative and selective groups. In the transformative method, mathematical and mapping, methods are used to select the best combination of features. In the selective

method, the best features of a feature vector that had the best performance in the ranking are used to identify the domain [28]. Because the best features are not always useful for correct diagnosis and classification of a disease, the features must be determined in such a way that the combination of features can achieve the desired result. Experience has shown that sometimes, the combination of features with low and high data yields better results than the features that contain more information. Among the well-known transformative methods in the domain of feature, the selection includes the principal component analysis (PCA) and independent component analysis. In these papers, the PCA method is used to select the desired feature vector.

2.4 Feature extraction

The researchers analyzed HRV and ECG signals to discriminate the patients and normal people. In the signal of these people, the time, time-frequency, linear, and nonlinear domain features are used after pre-processing that all features that have been extracted in papers so far are reported in Table 2.

Table 2. Extraction of nonlinear features, time and time-frequency

No	Feature	Abbreviation	Formula
1	Mean RR-intervals	MNN	$MNNRR = \frac{1}{N}\Sigma RRi$
2	Standard Deviation	SDNN	$\sqrt{\frac{1}{N}\sum(RRi - RRm)^2}$
3	Root Mean Square of the Successive Differences	RMSSD	$\sqrt{\frac{1}{N}(RRi + 1 - RRi)2}$
4	Standard Deviation of Successive Differences	SDSD	$\sqrt{\frac{1}{N}\sum_{i=1}^{n}\left(RRdiff - \overline{RRdiff}\right)2}$
5	The percentage of RR differences in each signal section whose difference is more than 50 ms	PNN50	$\frac{[(RRi + 1 - RRi) > 50ms]}{total(RRdiff)}$
6	Fractal dimension	FD	$FD = \frac{\log D}{\log r}$
	rScaling factor and D number of sections		
7	Correlation dimension	CD	$\frac{1}{N^2}\sum_{i,j}^{N} g(R - \Delta R)$
8	Approximate Entropy	AppEntrpy	$Cim(r) = \frac{n_{im}}{N-M+1}$ $ApEn(sN,m,r) = \ln[\frac{cm_r}{cm+1_r}]$
	r defines the similarity criterion. m Specifies the length of the pattern.		
9	Sample Entropy	SampEnt	sampEnt(k,r,N)=-ln(a(k) /b(k-1)) k=0,1,...,m-1 for
	N The length of the time series, m The length of the components to be compared over time series R is the approximate changes to accept a new member		
10	Return rate	Recurrence Rate	$RR = \frac{1}{M^2}\sum_{i,j}^{M} R_{i,j}$

	Rij Trajectory of phase space M is the matrix dimension		
11	Determinism	Determinism	DET=$\frac{\sum_{l=lmin}^{M} lp(l)}{\sum_{i,j}^{M} R_{i,j}}$
	l = The mean length of the diagonal line P (l) is the number of diagonal structures and M is the dimension of the matrix.		
12	Tsallis	ST	ST([q,p])=$\frac{1}{q-1}\sum_{j=1}^{N}[p_j ln p_j^q]$
	P is the time series and where q=1, there is entropy		

3 Classification Systems

All researchers have used the MLP, KNN, SVM, DT, and mixture expert (ME) techniques to diagnose the ECGs of normal people and the one who is at risk of SCD. The main purpose of the classification is to select the best features of the ECG signal in the classification to find results with high accuracy to predict sudden cardiac death. Then performance measurements are examined based on the test results. A summary of each of the classification methods has been reported below.

3.1 Support vector machine

In these papers to discriminate healthy individuals from at-risk people, the support vector machine is used due to the good generalizability, the recognition of input patterns, the selection of the optimal pattern, and learning ability (identifying the optimal structure for the training data set). This algorithm is among the supervised and parametric learning methods. The SVM includes several hyperplanes in a large space. These hyperplanes perform classification and divide the training data into two parts. Since the nonlinear signals cannot be easily discriminated, using kernel functions, they are transmitted to higher-dimensional feature space, and their separation will be much easier. As shown in Fig. 3, using the yellow page the input samples are mapped to a new space where the samples are linearly separable.
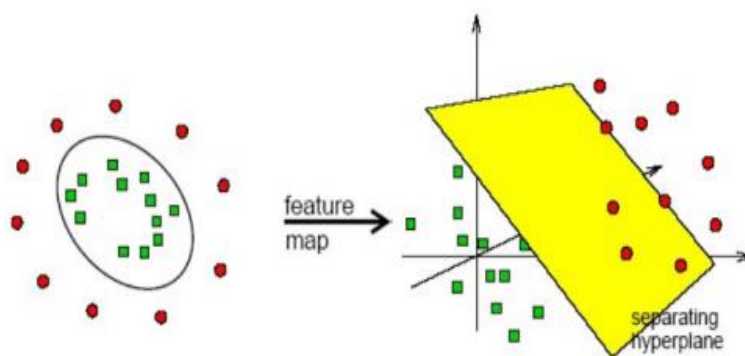

Fig. 3. Data transfer from the input space to the feature space

Today, this approach is frequently applied due to ease of training, the lack of local minima, and the ability to work with high-dimensional, preventing overfitting and compression of information (using SV instead of the entire data) [28, 29].

3.2 Perceptron neural network

Multilayer perceptron neural network is a set of neurons that are sequentially placed in a hidden layer. Input values after the multiplication with the weights in the inter-layer paths and being summed up and after passing through the corresponding function of the network form the output of the neurons. Finally, the obtained output is compared with the desired output, and the error is used to modify the network weights, and this process is called neural network training. The MLP is used to categorize data and discriminate ECG signals of a healthy person from an individual at risk of SCD. In this method, the MLP neural networks are composed of an input layer, one or more hidden layers and an output layer. The number of neurons in the input layer is equal to the dimensions of the feature vector, the number of output layer neurons equal with output classification classes and the number of hidden layers and hidden layers do not follow a particular rule, and optimal values are calculated by trial and error experimentally [24].

Similarly, the backpropagation principle is used for multi-layer perceptron learning, which is a learning algorithm consisting of two forward and return paths. In the forward path, the input vector is applied to the neural network, and its effect is propagated through the middle layers to the output layer. In this path, the network parameters remain constant, and the output values are calculated by the network [30].
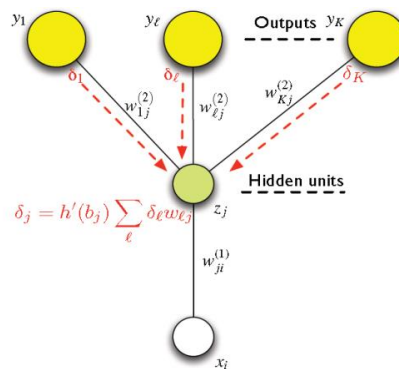

Fig. 4. Propagation of the initial error signals in MLP [31]

In the return path after the generation of output at the forward stage, the difference between the desired output and output calculated by the network is determined. Error signals in the return path of the output layer are redistributed over the entire network, and the network parameters are reset. The above dual process is repeated frequently

until approaching the output of the network to the optimal output. The process of education stops when the error is less than the permissible threshold [30]. Figure 4 shows an example of the initial backpropagation rule of the error signals in the MLP.

3.3   K- Nearest neighbor method

This classifier calculates the minimum distance between test and teaching data, which selects the most common class among the k-nearest neighbor as the desired class.

On the unknown sets given by the k-nearest neighbor, the algorithm is used, and the test sample is chosen among the training samples based on the distance criterion, i.e., Euclidean distance on which the test sample class label is attached. But normalization must be done before selection so that one feature would not dominate the other. Due to the low speed of this classifier in analyzing the test set, the contraction method is used. In this method, the distance between the test set and the database sets in all features or indices is not tested, and it is just applied in some of them. If the distance in this situation is greater than one threshold, the algorithm does its calculation on another set; otherwise, other indicators participate in the distance calculation [32].

3.4   Decision tree method

This classifier creates a tree of educational data using the selected feature [33]. This tree provides rules for classifying and determining the test data classes. The constituent components are the decision node, the branch, and the leaves.

The function of this classification depends on how well the tree is designed. The decision trees sort the samples by sorting them in the tree from the root node to the leaf nodes. Each inner node in the tree tests an attribute of the sample, and each branch removed from the node corresponds to a possible value for that attribute. Also, a category is assigned to each leaf node. Each sample is classified by starting from the root node of the tree and testing the specified attribute by this node and moving in the branch corresponding to the given attribute in the sample. It is possible to extract the set of rules from the decision tree that does not require any knowledge domain or parametric setup, so it is suitable for exploring the basic knowledge that can explore data with a high dimension [34]. Figure 5 shows an example of a decision tree in predicting credit risk.
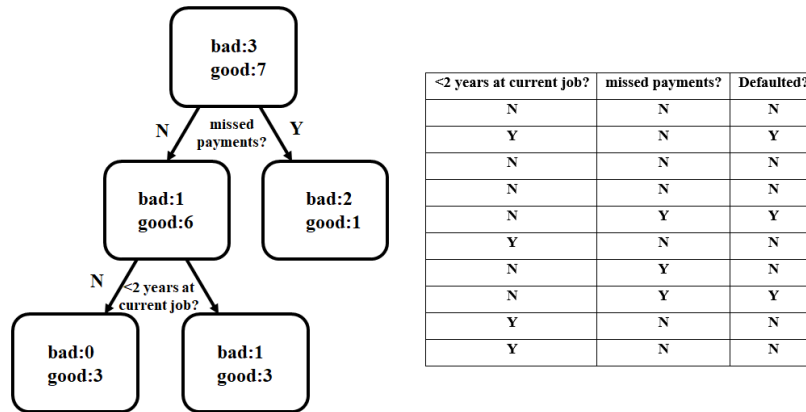
<2 years at current job?	missed payments?	Defaulted?
N	N	N
Y	N	Y
N	N	N
N	N	N
N	Y	Y
Y	N	N
N	Y	N
N	Y	Y
Y	N	N
Y	N	N

Fig. 5. The sample tree of the decision tree in predicting the credit risk

3.5 Mixture expert method

Mixture expert is a new method that selects the feature space locally and due to the presence of different features from different domains, allows choosing varying features every minute before the event by the optimal selection of features in each one-minute interval of the signal as an episode [37]. This not only enhances the prediction time but also allows interpreting the clinical symptoms based on the plurality of features present in each minute. On the other hand, the existence of mixture expert network will lead to a more appropriate decision as to the output for the processing of different domains because the result is the outcome of different decisions from the perspective of different scenarios about this phenomenon and eventually the unanimous outcome is reported [37]. Then to apply the mixture expert method, the linear features are extracted after extracting the HRV signal from the ECG and separating the signal into one-minute intervals according to studies conducted by Elias Ebrahimzade et al. [7]. Then the time-frequency domain feature is extracted, and finally, the nonlinear features are obtained [6-10]. In the next step, the best combination of features based on the highest discrimination between the two classes is selected through the local feature selection method for each one-minute interval. Then the healthy and people at risk of SCD are grouped through mixture experts [45].

Most of the studies that have taken place in the past use the global feature selection, which is not optimal. It means that the features selected in the minute before the event are not necessarily proper features for 1 hour before the event and discriminating people. Therefore, in the mixture expert (ME) method with the purpose of localizing feature selection method, a set of selected features is presented for each one-minute interval that not only increases the prediction time and improves the accuracy of ranking but also presents the type of dominant features at each time interval. To this end, various methods have been evaluated to provide an optimal method for reducing dimension reduction in local structure. Various strategies have been proposed for searching in the vast space of possible subsets of features. Usually, a comprehensive search [35-36] is

not practical because the number of subsets of features grows exponentially with the increasing number of features. On the other hand, the greedy search such as forward selection, backward elimination and, search space restriction provides an approximate response at low-time, but the problem with these methods is being trapped in local extremes. In this research, after proposing a local method, a randomized forward and backward search has been used for feature selection. On the other hand, for searching among the various subsets of the features, it is necessary to define the objective function for optimization. The main criterion for feature selection is to reduce the classification error. There are two evaluation criteria that are theoretically optimal. In other words, the subset of the features that optimizes these criteria minimizes the classification error of an ideal ranking. The first criterion is the Bayesian error rate where the rating error is directly calculated by the ideal ranking based on the formula (1).

$$BER = 1 - maxy\, f\,(x, y)\, dx \tag{1}$$

Where y id the label and x is the input vector with corresponding features. The f is the probability density function. The second criterion is mutual information and is calculated according to formula (2).

$$(i) = \iint (x,y)\, log\frac{f(x.y)}{f(x).f(y)}\, dxdy \tag{2}$$

So the third suggestion in the feature selection section is to use a Bayesian error rate as the target function for optimization; therefore, the best feature is selected based on the Bayesian error. Then, the combination of the best feature selected on the basis of the lby examined with other features to achieve the lowest binary error. Accordingly, this process continues to the point that adding a feature improves the Bayesian error over the previous state. In fact, in this method, the goal is to obtain the least number of features with the least error.

4 Results

To evaluate the results of the proposed methods, each of the researchers analyzed the time intervals before the occurrence of SCD. Different classification methods are available to identify the efficiency of the proposed algorithm, such as a support vector machine, multilayer perceptron neural network, radial base function neural network, k-nearest neighbor, and mixture expert. The use of features introduced in different domains and different classifications has led to the observation of different prediction horizons in various studies. The results of these predictions are often free from the interpretations of clinical symptoms, and their maximum presented time reaches 4 minutes before the event with acceptable validity, which is not an acceptable time for people who have heart attacks outside the hospital. Accordingly, the most prominent of these evaluations is the mixture expert that is using the best feature extraction method a new method is used for optimal feature space selection locally. This method allows choosing varying features every minute before the event by the optimal selection of features in each one-minute interval of the signal as an episode, which increases the prediction

time from 4 minutes before the death to 12 minutes. It also allows the interpretation of clinical symptoms regarding the multiplication of the presence of the features per minute. The accuracy of classifications according to the proposed method, is shown in Fig. 6 in comparison with other methods presented so far [18]. Similarly, in Figure 7, the share of each feature from different processing areas is shown about the initial number of extracted per minute. Here a summary of studies that have been performed to predict sudden cardiac death using ECG signal analysis has been presented in Comparative Tables 3 and 4.
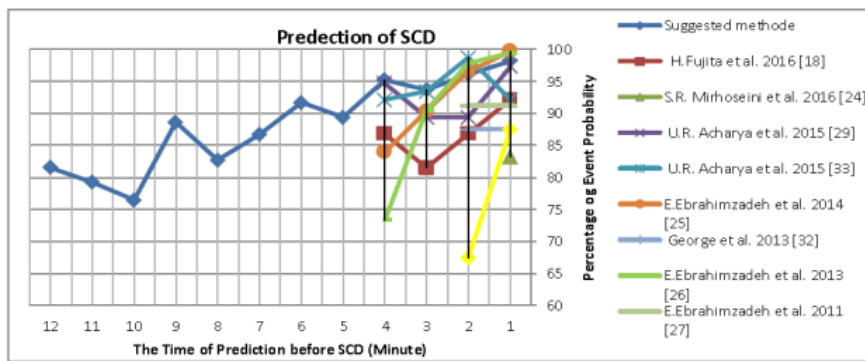

Fig. 6. Comparison of the mixture expert and methods presented in other articles based on the predicted accuracy percentage of SCD and the prediction time (min) [18].
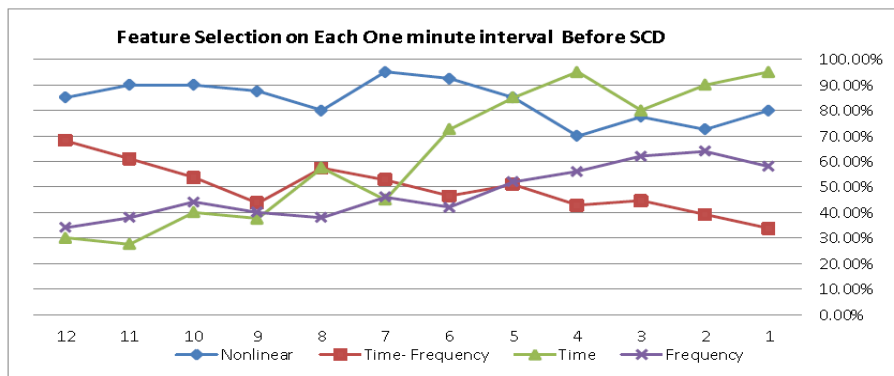

Fig. 7. The share of each selected feature group for each minute relative to the number of features extracted in that domain [18]

Table 3. Summary of studies in the prediction of sudden cardiac death using ECG signal analysis

Author and year of publication	Type of data	Number of data	Method	Classification	Results and findings
Shen et.al (2007)	HRV	20 healthy people 30 patients	Wavelet analysis and analysis of heart rate variation using Fourier transform Total number of extracted features = 4	Leak Square Error, Multilayer Perceptron Neural Network	By wavelet method: Accuracy = 87.5% With Heart Rate Analysis: Accuracy = 67.44%
Combining linear and nonlinear methods					
Voss et.al (1996)	HRV	35 healthy people 26 patients with ventricular fibrillation Total number of extracted features = 23	Linear analysis and nonlinear analysis	Stepwise discriminant function	Accuracy = 96%
George et.al	HRV	40 patients with ischemic heart disease	Linear analysis and nonlinear analysis	SVM	Accuracy of the SVM = 87.5%
Combine linear methods with a conversion function					
Elias et.al (2011)	HRV	25 healthy people 35 patients with a sudden heart attack	The linear analysis in the time and frequency domain and using the Wigner-Veyl transform function Total number of extracted features = 25	Accuracy in one minute before the incident = 99.16% Perceptron, with the KNN method = 96.04% Two minutes before the event = 91.23% with perceptron and KNN = 9.27%	Accuracy 1 minute before event = 99.73%
Elias et.al (2014)	HRV	25 healthy people 35 patients with a sudden heart attack	The linear analysis in the time and frequency domain and use the conversion function and	MLP and KNN	Accuracy 2 minutes before event = 96.59% Accuracy 3 minutes before event = 90.36%
Nonlinear methods					
Acharya et al (2015)	Electrocardiogram	18 healthy people 20 people with SCD	Nonlinear features and sudden cardiac death indicators	SVM and decision tree	Accuracy 1 minute before event = 92.11%

			Total number of extracted features = 18		Accuracy 2 minutes before event = 96.68%
Rohollah et.al (2016)	HRV	19 healthy people 18 patients	Nonlinear, time and frequency features Total extracted features = 22	Poincaré, SVM	Accuracy = 83.24% Validity = 89.5%
	HRV	10 patients with Physionet 10 patients with MIT-BH	RMSSD SDSD RR-interval Heart rate	SVM	Accuracy = 75-80%

Table 4. Comparison of different studies in the SCD detection using ECG and HRV

Author and year	Data	No of data	Method	Features	Classification	Results	Future works
Elis et al., 2011	HRV	25 healthy people 25 patients with a sudden heart attack	The linear analysis in the time and frequency domain and using the Wigner-Veyl transform function	Time features= sdNN, sdaNN1, pNN110, pNN30 Frequency features ULF, VLF, LF, LP/HF, LF/S, HF/S	MLP and KNN	2 minutes before = 91.22% (MLP) (KNN) = 89.27%	Not mentioned
Acharya et al., 2015	ECG	18 healthy people 20 patients with a sudden heart attack	Nonlinear features and sudden cardiac death indicators	Fractal estimation detrended fluctuation analyses Sample entropy	Support vector machine, decision tree, and k-nearest neighbor	Accuracy in 4 minutes before = 92.11%	Improved windows for high accuracy detection in a short time
Acharya et al., 2015	HRV	18 healthy people 20 patients with a sudden heart attack	Non-linear features	Fuzzy Kolmogorov entropy Energy features of discrete wavelet transform coefficients	Support vector machine, decision tree, and k-nearest neighbor	Accuracy in 4 minutes before = 94.7%	The possibility of faster sudden cardiac death prediction

Elias et al., 2017	HRV	35 healthy people 35 people with SCD	Feature = 23 Linear, frequency domain and non-linear	Selecting local features for different times Number of different features	ME	For an average of 12 minutes Accuracy: 88.29%	Not mentioned
Reeta devi et al., 2017	HRV	25 healthy people 25 patients with a sudden heart attack	Features= 26 Linear and non-linear domain	detrended fluctuation Poincaré Frequency and nonlinear features	KNN SVM	Accuracy = 90.91% Sensitivity = 88.89%	Not mentioned

5 Discussion and Conclusion

Due to the nonlinear nature of the HRV signal and the similarity of the ECG signal over many periods of time, the use of the HRV signal is more popular among the researchers. The analysis of various studies shows that by approaching the time of death, linear features (time and frequency) can be predictive of death according to the behavior and evident variations in the signal of patients. Instead, by taking distance from the death range, the chaotic and nonlinear features are more effective. Therefore, a more precise selection of features in this area can be useful for increasing the horizon of prediction of death.

Also, compared with the research conducted on the HRV signal, nonlinear features significantly predict the risk of SCD earlier [16]. By combining the nonlinear method of the discrete wavelet transform with correlation dimension (CD), DFA, FD, and Samp ENT, one can record in the signals of people with SCD and healthy people and predict SCD with a high percentage based on feature selection.

Similarly, Muragappan et al. using the machine learning algorithm and the use of the two KNN and Fuzzy classifications were able to predict the SCD within 5 minutes before it [20]. Validation of extracted features was performed by the ANOVA method which showed that the F values in sdHR had high classification features in comparison with other values. In these results, the KNN has a classification rate of 93.72%, and the results of fuzzy classification are 86.67%. It had a good performance in the diagnosis of SCD and NC with the accuracy of 96.36% by these researchers in 2014 compared with the previous approach [41]. In KNN, the values of k vary from 1 to 10, and the best result is obtained at K = 5. Standard deviation at RR distances provides results with a sensitivity of 92.22% and a positive predictive value of 95.40%. By comparing this method with other colleagues, it can be said that VOSS et al. predicted the SCD 2 minutes before its occurrence with an accuracy of 67.44% using frequency domain and time domain features.

Similarly, Ebrahimzadeh et al. predicted SCD with a precision of 91.23% and83.93% in 1 and 3 minutes before its onset using frequency- time and nonlinear domain method [9, 10].

Also, Obayya et al. using their proposed method was able to predict SCD with 96.36% accuracy, and they increased it into 98.18% by combining the features [40].

Similarly, Fujita et al. predicted SCD within 4 minutes before the event, with an accuracy of 94.7%. This method allows physicians to predict SCD with an alarm mounted on monitoring [16].

Hooshyarifar et al. using their proposed algorithm and SVM and KNN classifiers predicted SCD 4 and 5 minutes before the occurrence with 92.14% (SVM) and 94.10% (KNN) accuracy [19].

Acharya et al. using the P-test and t-test compared the results of all of their selected features, and the results showed that appEnt and sampEnt values had a high ranking. Similarly, they are used as KNN, DT, and SVM inputs that the results of the classification accuracy at the first, second, third, and fourth times are 92.11%, 98.68%, 93.42%, and 92.11% respectively [8]. In another article using the RQA, the researchers predicted the SCD using the KNN and PNN with an accuracy of 86.8%. As shown in the Recurrent plot (RP) in the article, in SCD individuals, diagonal lines and small squares are observed indicating a higher uniformity in the signal and in the normal HRV signal of normal people the points are concentrated in several places in the section, and they seem periodic and rhythmic in some places and diagonal lines do not appear on the normal human plot [17].

Ma'ida Bahrami et al. extracted proper features in time-frequency domains and their composition with nonlinear features [15]. Using extraction of nonlinear time-frequency features with the help of wavelet transform and the extraction of two nonlinear features with detrended fluctuation analysis and Poincaré analysis and combining these two, they provided the proper set of features for network training with MLP. An important advantage of this combined method with other methods is that due to the use of the combination of both wavelet and MLP networks, the weights of the MLP network are corrected, and the weight of the wavelet network is adjusted automatically, and this improves system efficiency compared to other methods. The average discrimination of the linear, time-frequency, nonlinear, and combined methods is 67%, 76%, 83%, and 95%, respectively.

6   References

[1] M. C. Koester. (2001). A review of sudden cardiac death in young athletes and strategies for preparticipation cardiovascular screening. Journal of athletic training, 36, 197–204. https://doi.org/10.5812/asjsm.34818.

[2] Mohammad Karimi Moridani, Seyed Kamaledin Setarehdan, Ali Motie Nasrabadi, Esmaeil Hajinasrollah, A novel approach to mortality prediction of ICU cardiovascular patient based on fuzzy logic method, Biomedical Signal Processing and Control, 45, 160-173. https://doi.org/10.1016/j.bspc.2018.05.019

[3] Jones, J. L., and Tovar, O. H. (1996). The Mechanism of Defibrillation and Cardioversion, Proceedings of the IEEE, 84, 392-403. https://doi.org/10.1146/annurev-bioeng-070909-105305. https://doi.org/10.1109/5.486742

[4] Mohd Zubir Suboh, Rosmina Jaafar, Nazrul Anuar Nayan, Noor Hasmiza Harun. (2019). ECG-based Detection and Prediction Models of Sudden Cardiac Death: Current Performances and New Perspectives on Signal Processing Techniques, International Journal of Online and Biomedical Engineering (iJOE), 15 (15), 110-1260. https://doi.org/10.3991/ijoe.v15i15.11688

[5] Huikuri HV, Castellanos A, Myerburg RJ. (2001). Sudden death due to cardiac arrhythmias. N Engl J Med, 345, 1473-82. https://doi.org/10.1056/nejmra000650

[6] E. Ebrahimzadeh and M. Pooyan. (2011). Early detection of sudden cardiac death by using classical linear techniques and time-frequency methods on electrocardiogram signals. Journal of Biomedical Science and Engineering, 4, 699-706. https://doi.org/10.4236/jbise.2011.411087.

[7] E. Ebrahimzadeh, M. Pooyan, and A. Bijar. (2014). A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals. PloS one, 9, 1-14. https://doi.org/10.1371/journal.pone.0081896

[8] U. R. Acharya, H. Fujita, V. K. Sudarshan, V. S. Sree, L. W. J. Eugene, D. N. Ghista and R. San Tan. (2015). An integrated index for detection of sudden cardiac death using discrete wavelet transform and nonlinear features. Knowledge-Based Systems, 83, 149-158. https://doi.org/10.1016/j.knosys.2015.03.015.

[9] E. Ebrahimzadeh and M. Pooyan. (2013). Prediction of sudden cardiac death (SCD) using time-frequency analysis of ECG signals, Computational Intelligence in Electrical Engineering, 3(4),15-26.

[10] Ebrahimzadeh, E., Pooyan, M. (2011). Early detection of sudden cardiac death by using classical linear techniques and time-frequency methods on electrocardiogram signals, Biomedical Science and Engineering, 4, 699-706. https://doi.org/10.4236/jbise.2011.411087.

[11] Shen, T.W., Shen2, H.P., Lin2, C., and Ou2, Y. (2007). Detection and Prediction of Sudden Cardiac Death (SCD) For Personal Healthcare, Proceedings of the 29th Annual International Conference of the IEEE EMBS, 2575-2578. https://doi.org/10.1109/IEMBS.2007.4352855.

[12] Fishman, G. I., et al. (2010). Sudden Cardiac Death Prediction and Prevention: Report from A National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop, Circulation, 122(22), 2335-2348. https://doi.org/10.3410/f.725393997.793505054.

[13] Shen, T. W., Lin, C. C., Shen, H. P., Ou, Y.L., and Lin, C. H. (2007). A personal Sudden Cardiac Death (SCD) detector based on ECG biometric technology. World Congress on Medical Physics and Biomedical Engineering, 14, 1249-1252. https://doi.org/10.1007/978-3-540-36841-0_302

[14] Shen, T. W., Lin, C. C., Shen, H. P., Ou, Y. L., and Lin, C. H. (2007). A personal Sudden Cardiac Death (SCD) detector based on ECG biometric technology, World Congress on Medical Physics and Biomedical Engineering, 1249-1252. https://doi.org/10.1007/978-3-540-36841-0_302

[15] Maede Bahrami, Mahshid Mahrami, Alireza Taghiabadi. (2013). prediction of sudden cardiac death by mixture intelligent system. 1st Iranian National Conference Electrical Engineering, 1-10.

[16] H. Fujita, U. R. Acharya, V. K. Sudarshan, D. N. Ghista, S. V. Sree, L. W. J. Eugene, and J. E. Koh. (2016). Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index", Applied Soft Computing, 43, 510-519. https://doi.org/10.1016/j.asoc.2016.02.049

[17] U. R. Acharya, H. Fujita, V. K. Sudarshan, D. N. Ghista, W. J. E. Lim, and J. E. Koh. (2015). Automated prediction of sudden cardiac death risk using Kolmogorov complexity and recurrence quantification analysis features extracted from HRV signals, IEEE International Conference on Systems, Man, and Cybernetics (SMC), 141-144. https://doi.org/10.1109/smc.2015.199

[18] S. Amoozegar, M. Pooyan, and E. Ebrahimzadeh. (2013). Classification of brain signals in normal subjects and patients with epilepsy using a mixture of experts, Computational Intelligence in Electrical Engineering, 4,1-8. https://doi.org/10.3233/thc-181579.

[19] V. Houshyarifar, M. Chehel Amirani. (2017). Early detection of sudden cardiac death using Poincare plots and recurrence plot-based features from HRV signals, Turkish Journal of Electrical Engineering & Computer Sciences, 25, 1541-1553. https://doi.org/10.3906/elk-1509-149

[20] L. Murukesan, M. Murugappan, M. Iqbal, and K. Saravanan. (2015). Time Domain Features Based Sudden Cardiac Arrest Prediction Using Machine Learning Algorithms. Journal of Medical Imaging and Health Informatics, 5, 1267-1271. https://doi.org/10.1166/jmihi.2015.1525.

[21] Reeta Devi, Hitender Kumar Tyagi, Dinesh Kumar. (2017). Early Stage Prediction of Sudden Cardiac Death Using Linear and Non-Linear Features of Heart Rate Variability, International Journal of Electronics, Electrical and Computational System, 6, 742-754. https://doi.org/10.1155/2012/219080.

[22] S. R. Mirhoseini, M. R. JahedMotlagh, M. Pooyan. (2016). Improve Accuracy of Early Detection Sudden Cardiac Deaths (SCD) Using Decision Forest and SVM. International Conference on Robotics and Artificial Intelligence, 1-4. https://doi.org/10.1142/S0218213007003163.

[23] Annmarie G. Raka 1, Ganesh R. Naik 2 ID and Rifai Chai. (2017). Computational Algorithms Underlying the Time-Based Detection of Sudden Cardiac Arrest via Electrocardiographic Markers, 7:954-960. https://doi.org/10.3390/app7090954.

[24] http://www.physionet.org/PhysioBank/Signal%20Archives/%20ECG/%20Normal%20Sinus%20Rhythm%20database.

[25] http://www.physionet.org/PhysioBank/Signal%20Archives/%20ECG/Sudden%20Cardiac%20Death%20Holter%20Database.

[26] Pan, J., and Tompkins, W.J. (1985). Real time QRS detection algorithm, IEEE Transactionon Biomedical Engineering, 32(3), 230-36. https://doi.org/10.1109/tbme.1985.325532.

[27] Abdi H, Williams LJ. (2010). Principal component analysis, Comput Statist 2, 433-59. https://doi.org/10.1002/wics.101.

[28] Baudat G. and Anouar F. (2000). Generalized discriminant analysis using a kernel approach. Neural Computation, 12, 2385-2404. https://doi.org/10.1162/089976600300014980

[29] C. Jenefar Sheela and L. Vanitha. (2014). prediction of sudden cardiac death using support vector machine, IEEE Trans. Biomed. Eng. International Conference on Circuit, Power and Computing Technologies [ICCPCT], 377-381. https://doi.org/10.1109/iccpct.2014.7054771

[30] QinghuaGuo, MaggiKelly, Catherine H.Graham. (2005). Support vector machines for predicting distribution of Sudden Oak Death in California, 182, 75-90. https://doi.org/10.1016/j.ecolmodel.2004.07.012.

[31] Hoon Jin1 and Seungcheon Kim. (2015). Performance Evaluations of Diagnostic Prediction with Neural Networks with Data Filters in Different Types. International Journal of Bio-Science and Bio-Technology, 7, 61-70. https://doi.org/10.14257/ijbsbt.2015.7.1.07

[32] Hiroshi Shimodaira. (2015). Multi-Layer Neural Networks, the University of Edinburgh, 1-6.

[33] Daniel O. Mcinerney and Maarten Nieuwenhuis. (2009). A comparative analysis of KNN and decision tree methods for the Irish National Forest Inventory. School of Biology and Environmental Science. 30, 4937-4955. https://doi.org/10.1080/0143116090302 2936.

[34] Uwe Glässer, Sarah Rastkar, Mona Vajihollahi. (2006). Computational modeling and experimental validation of aviation security procedures ISI'06 Proceedings of the 4th IEEE international conference on Intelligence and Security Informatics, 420-431. https://doi.org/10.1007/11760146_37

[35] Lukas Tanner, Mark Schreiber, Jenny G. H. Low, Adrian Ong, Thomas Tolfvenstam, Yee Ling Lai, Lee Ching Ng, Yee Sin Leo, Le Thi Puong, Subhash G. Vasudevan, Cameron P. Simmons, Martin L. Hibberd, and Eng Eong Ooi. (2008). Decision Tree Algorithms Predict the Diagnosis and Outcome of Dengue Fever in the Early Phase of Illness. PLoS Negl Trop Dis, 2, 1-8. https://doi.org/10.1371/journal.pntd.0000196.

[36] P. M. Narendra and K. Fukunaga. (1977). A branch and bound algorithm for feature subset selection, IEEE Transactions on Computers. 26: 917 – 922. https://doi.org/10.1109/tc.1977.1674939

[37] M. Nikravan, E. Ebrahimzadeh, M. R. Izadi and M. Mikaeili, (2016). Toward a computer aided diagnosis system for lumbar disc herniation disease based on MR Image analysis. Biomedical Engineering: Applications, Basis and Communications, 28, 1650042- 1650054. https://doi.org/10.4015/s1016237216500423

[38] S. Amoozegar, M. Pooyan and E. Ebrahimzadeh. (2013). Classification of brain signals in normal subjects and patients with epilepsy using mixture of experts. Intelligent Systems in Electrical Engineering, 4, 1-8. https://doi.org/10.3233/THC-181579.

[39] E. Ebrahimzadeh, S. M. Alavi, A. Bijar and A. Pakkhesal. (2013). A novel approach for detection of deception using Smoothed Pseudo Wigner-Ville Distribution (SPWVD), Journal of Biomedica Science and Engineering. 6: 8-18. https://doi.org/10.4236/jbise.2013.61002.

[40] Voss A, Kurths J, Kleiner HJ, Witt A, Wessel N, Saparin P, Osterziel KJ, Schurath R, Dietz R. (1998). The application of methods of nonlinear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death, cardiovascular research, 31, 419-433. https://doi.org/10.1016/s0008-6363(96)00008-9.

[41] M. Obayya, F. Abou-Chadi. (2008). Data fusion for heart diseases classification using multi-layer feed forward neural network, International Conference on Computer Engineering & Systems (ICCES), 67-70. https://doi.org/10.1109/icces.2008.4772968

[42] L. Murukesan, M. Murugappan, M. Iqbal, and M. Saravanan. (2014). Machine learning approach for sudden cardiac arrest prediction based on optimal heart rate variability features. J. Med. Imaging Health Inf, 4, 521-532. https://doi.org/10.1166/jmihi.2014.1287.

[43] Maria Teresa La Rovera and Franco Cobelli. (2003). Short –Term Heart Rate Variability Strongly Predicts Sudden Cardiac Death in Chronic Heart Failure Patients, Clinical Investigation and Reports, 107, 565-570. https://doi.org/10.1161/01.cir.0000047275.25795.17

[44] Ma kikallio HT, Huikuri HV, Ma¨kikallio A, Sourander LB, Mitrani RD, Castellanos A, Myerburg R. (2001). Prediction of sudden cardiac death by fractal analysis of heart rate variability in elderly subjects. Journal of American college of cardiology, 37, 1395-402. https://doi.org/10.1016/s0735-1097(01)01171-8

[45] Manis G, Nikolopoulos S, Arsenos P, Gatzoulis K, Stefanadis C. (2013). Risk stratification for arrhythmic sudden cardiac death in heart failure patients using machine learning techniques. Comput Cardiol, 40,141-144. https://doi.org/10.3390/jcm7110436.

[46] Hatem A Elaydi. (2019). Personal Health Care System Using IOT; International Journal of Online and Biomedical Engineering (iJOE), 15(7), 16-31. https://doi.org/10.3991/ijoe.v15i07.10265

7   Authors

Mohammad Karimi Moridani received a BS degree in electrical engineering-Electronic in 2006 and his MS and Ph.D. degrees in biomedical engineering-bioelectric in 2008 and 2015, respectively, from Islamic Azad University, Science and Research Branch, Tehran, Iran.

Now he is an assistant professor in the Biomedical Engineering Department at Tehran Medical Science, Islamic Azad University, Tehran, Iran. His current research interests are in the field of biomedical signal and image processing, nonlinear time series analysis, and cognitive science. Particular applications include ECG, HRV, and EEG Signal Processing in detection and prediction of diseases, Epileptic Seizure Prediction, pattern recognition and Image Processing for face and beauty recognition, watermarking, etc. He is with Department of Biomedical Engineering, Faculty of Health, Tehran Medical Sciences, Islamic Azad University, Tehran. Email: karimi.m@iautmu.ac.ir

Shahrzad Marjani received a BS degree in biomedical engineering-clinical in 2017 from Tehran Medical Science, Islamic Azad University, Tehran , Iran.

Her research interests include biomedical signal processing, analysis of nonlinear behavior of ECG signal for cardiac disease detection. Now, she is a Master student in biomedical engineering- bioelectric in Department of Biomedical Engineering, Tehran North Branch, Islamic Azad University, Tehran, Iran.