Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 176 (2020) 2940–2948

Procedia Computer Science

www.elsevier.com/locate/procedia

24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

A big – data classification tree for decision support system in the detection of dilated cardiomyopathy using heart rate variability

G.Silveri[a,*], M.Merlo[b], L.Restivo[b], M.Ajčević[a], G.Sinagra[b], A.Accardo[a]

[a]*Dep. Engineering and Architecture, University of Trieste,via Valerio10, Trieste 34127, Italy*
[b]*Cardiovascular Department, University of Trieste, via Valdoni7, Trieste 34149, Italy*

Abstract

Dilated cardiomyopathy (DCM) is a heart muscle disease characterized by left ventricular (LV) or biventricular dilatation and systolic dysfunction in the absence of either pressure or volume overload or coronary artery disease sufficient to explain the dysfunction. The use of heart rate variability (HRV) analysis as well as of some machine learning algorithms, proved to be a valuable support in the diagnosis of cardiovascular disease. However, till now, only single beats or electrocardiogram segments of subjects affected by DCM were identified using machine learning techniques applied to HRV parameters. In this study, we used linear and non-linear HRV parameters and some clinical parameters (age, sex and left ventricular ejection fraction) evaluated on a large cohort of 972 subjects to early identify subjects suffered from DCM and to find which features could be selected as predictors for a correct diagnosis. By using principal component analysis and stepwise regression, we reduced the original parameters used as inputs for a series of classification and regression trees (CART). The highest accuracy of 97% and Area Under the Curve (AUC) of 95% were achieved using the ratio between low frequency and high frequency (LF/HF), sex and left ventricular ejection fraction (LVEF) parameters as inputs of the classifier.

Keywords: Classification and Regression Tree, Dilated cardiomyopathy, HRV parameters;

* Silveri Giulia. Tel.:+39 0405587130;
giulia.silveri@phd.units.it

1877-0509 © 2020 The Authors. Published by Elsevier B.V.

10.1016/j.procs.2020.09.209

1. Introduction

Dilated cardiomyopathy (DCM) is a progressive disease of the heart muscle defined by the chamber enlargement and contractile dysfunction of the left ventricle in the absence of chronic pressure and or/volume overload; this disease is the third most common cause of heart failure [1]. If no cause is discovered the cardiomyopathy may be referred as idiopathic, characterized by an advanced stage of left ventricular dilatation and dysfunction. The DCM could be diagnosed in asymptomatic individuals during routine medical screening in which the electrocardiogram (ECG) presented abnormalities ranging from isolated T wave changes to septal Q waves. In this patient population, increased QT variability has been independently associated with occurrence of major arrhythmic events, including sudden cardiac death [2].

A useful tool extracted from ECG is the variation over time of the period between consecutive heartbeats (RR intervals) called heart rate variability (HRV). This measurement is used to assess quantitatively the cardiac autonomic activity as the result of the interaction between sympathetic and parasympathetic activity [3].

In literature, some studies used HRV measurement as an instrument to provide additional valuable insight into physiological and pathological conditions, to enhance risk stratification and to be a predictor of the risk of mortality [4-10]. In particular, linear (time and spectral) and non-linear features of HRV have been evaluated for detection of cardiovascular disease like heart failure. Depressed HRV parameters indicated an impairment of the autonomic nervous system that was observed for DCM [6-10]. The time and frequency HRV parameters had higher values in normal than in DCM [6-9], except for normalized high frequency (HFn) that showed higher values in DCM patients [7]. However, due to the non-linear nature of heart signal, the features over time and spectral domain analysis do not characterize enough subjects suffering from cardiovascular disease. Moreover, as the noise in signal increases, the effectiveness of spectral domain analysis will decrease, so some authors introduced non-linear features extracted from HRV signal like sample entropy and Poincarè parameters that presented lower values in DCM [10].

Nowadays, to support decision-making, many machine learning-based methods are widely applied in the field of medicine to solve clinical problems as well as to reduce time of pre-screening process. Thus, several different machine learning methods [11-29] was used in ECG-classification research particularly toward arrhythmia identification such as ventricular tachycardia and ventricular fibrillation [11,12]. In particular, the literature reports the application of methods like Kth nearest-neighbors (KNN) [13], clustering technique [14], linear discrimination analysis (LDA) [15,16], fuzzy analysis [17-19] artificial neural network (ANN) [20-23] support vector machine (SVM) [24-26], recurrent neural network [27] and decision tree [28,29].

Some of these algorithms are also be used to classify cardiomyopathy [22,23,26-29]. In particular, Adetiba et al [22] developed an automated heart detection model using 400 ECG data segments recorded in 40 athletes and ANNs in order to differentiate normal and cardiomyopathy heart conditions. Some statistical parameters (mean, median, mode, variance and standard deviation) calculated on each data segment were used as inputs of the ANNs with10 hidden layers neurons, achieving an accuracy of 98%. Megat Ali et al [23] presented a cardiomyopathy detection approach using a multilayer perceptron network based on the width of P-wave, QRS-complex and T-wave characteristics of 600 beats, from PTB Diagnostic ECG Database, achieving an accuracy of 98.9%. Ghosh et al [26] developed support vector machine (SVM) to classify normal and cardiomyopathy segments using ECG signal, recorded for few minutes on 5 normal subjects and 5 patients taken from [1]PTB Diagnostic ECG Database. The SVM was used to classify the features extracted through a continuous wavelet transform of the ECG signals, obtaining an accuracy of 92%. Shukuri et al [27] proposed the use of recurrent neural network for detecting cardiomyopathy in ECG beat samples, 200 waveforms from 52 normal and others 200 waveforms from 18 cardiomyopathy subjects, taken from [1]PTB Diagnostic ECG Database, too. They used five hidden neurons and four different learning algorithms, obtaining an accuracy of 90%.

As these studies focused their attention on the characteristics of single beats that could be strongly affected by noise or artifacts and since cardiomyopathy produces changes also in the variation over time of the interval between consecutive heartbeats, other authors [28,29] turned their attention toward parameters extract from Heart Rate Variability (HRV) signals. In particular, Thirugnaman et al [28] studied 25 linear and non-linear HRV parameters calculated on ECG samples, 16 from DCM and 6 from normal subjects, taken from Physionet database [www.physione.org]. They developed a stacked-ensemble classifier using SVM, KNN and decision tree to identify cardiomyopathy, achieving a sample detection accuracy of 99.9%. Moreover, Mahesh et al [29] used a combination of tree structure and logistic regression on 15 linear and non-linear HRV parameters to identify 13 cardiomyopathy and 410 normal segments extracted from [1]Physionet database. Considering separately different groups of parameters, they achieved a maximum accuracy of 95.61%.

Although these studies reached interesting results, none of them produced a system capable of classifying subjects as rather single beat or segment belonging to some normal subject or with DCM. Furthermore, data coming from a very limited number of DCM subjects were examined and complex decision support systems were developed.

In order to focus mainly on the identification of subjects with cardiac disease rather than on single beats, the aim of this study is to verify if HRV derived parameters applied to a simple Classification and Regression Tree (CART) [30,31] could be able to distinguish DCM patients from normal subjects in large cohort of cases. With the aim to evaluate which features, applied to CART, could discriminate DCM patients with high accuracy, some combinations of linear and non-linear parameters extracted from HRV together with age and sex were examined. To reduce the number of features preserving the variance, principal component analysis and stepwise regression were used too. Finally, we evaluate whether the addition of a clinical parameter, i.e. the left ventricular ejection fraction (LVEF), that is a prognostic indicator in DCM patients [1,2], could improve the classification performance.

2. Methods

2.1. Study population

A total of 972 subjects consecutively enrolled from December 2016 to October 2018 at the Cardiovascular Department, ASUGI, Trieste were examined. Of these, 773 (374 males, aged 63±19 and 399 females, aged 60±19) were normal subjects and 199 (126 males, aged 57±14 and 73 females, aged 62±15) suffered from DCM. The disease is characterized by left ventricular (LV) dilatation and LV systolic dysfunction with normal wall thickness in the absence of abnormal loading conditions. Following guidelines [2], DCM is defined by echocardiography as the presence of an ejection fraction below 45% and/or a fractional shortening less than 25% and a LV end-diastolic dimension greater than 112% of the predicted value corrected for age and body surface area. Normal subject did non present neither peripheral artery disease, thyroid disorders or history of myocardial revascularization. The study was performed according to the Declaration of Helsinki and all patients gave written consent.

2.2. HRV parameters

All subjects performed a 24h ECG Holter monitoring session using three channels tracking record (Sorin Group, Italy). The ECG signal was sampled at 200Hz and RR intervals were automatically extracted from records by using SyneScope analysis software (Sorin Group, Italy) that automatically identify QRS complexes. Data were analyzed by using proprietary MATLAB (MathWorks, USA) program that examined segments of 300s each. In order to reduce the influence of noise and artifacts a pre-processing of RR times series was carried out. RR segments were accepted only if the longest ectopic beats sequence or the longest artifact in the examined segment was shorter than 10s and if the total duration of artifacts and ectopic beats was shorter than 20% of the segment duration [32]. In this pre-processing phase, it has to be considered that RR series are a sequence of varying intervals that occur at non equidistant sampling times and that refused RR must be appropriately replaced. With the aim to perform analysis requiring time series sampled at a constant frequency, RR sequences were interpolated by using cubic spline based on valid RR values and resampled at 2 Hz. Moreover, since the original recordings started at different times, the subjects' RR time series were aligned to have a common start time at 10:00am.

To describe quantitatively HRV signal, we calculated on each segment and subject several linear and non-linear HRV parameters successively averaged along the 24h. The linear parameters exploited time and frequency domains. In the former the mean (meanRR) and the standard deviation (SDNN) of the RR intervals, the root mean square of the squared differences of successive RR intervals (RMSSD), the number of differences of successive RR intervals greater than 50ms (NN50), and the proportion of NN50 divided by the total number of RR intervals (pNN50) were calculated. In the frequency domain, the absolute powers in high (HF) and low (LF) frequency bands were examined. The power in high frequency band (from 0.15 Hz to 0.4Hz) is an indicator of the vagus nerve control on the heart rhythm while the power in low frequency band (from 0.04Hz to 0.15Hz) is due to the joint action of the vagal and sympathetic components on the heart beats modulation, with a predominance of the sympathetic one [33]. These powers were also expressed in normalized units (LFn and HFn) which represent the relative value of each power component in proportion to the total power. The LFn and HFn parameters emphasize the controlled and relative behavior of the two branches of the autonomic nervous system (ANS). Finally, we calculated the ratio between LF and HF (LF/HF) which reflects the balance between the two components of the ANS.

Since non-linear mechanisms are also involved in the genesis of heart rate dynamic [34], some non linear HRV parameters were considered, nominally Poincaré plot parameters, Fractal Dimension (FD) and beta exponent. The Poincaré plot [35] is a quantitative-visual technique that represents the projection of the attractor in the phase space in a two-dimensional space obtained by plotting each point of the time series against the next one. The dispersion of points perpendicular to the line of identity reflects the level of short-term variability (SD1) of the RR sequences while the dispersion of points along the line of identity indicates the level of long-term variability (SD2) of the RR sequences. SD1, SD2 and their ratio (SD1/SD2), representing the relationship between these components, were the three parameters considered. Moreover, fractal dimension, quantifying the fractal-like behavior of a time series and the complexity of the system producing the signal, was evaluated by using Higuchi's algorithm [36]. The last non-linear parameter, the beta exponent, measuring the relation between frequency and power spectrum of RR interval, was calculated as the slope of the regression line of the log(power) versus log(frequency) relation in the 10^{-4} – 10^{-2} Hz frequency range, providing an index for long-term scaling characteristics of HRV signal [37].

2.3. Features selection

In order to classify DCM we developed six different classifiers using three different subsets of normalized HRV features with age and sex and adding, in half of them, the left ventricular ejection fraction (LVEF). The first combination of features (TREE1) was obtained applying the stepwise regression to all the 17 parameters selecting those presenting significant p-value ($p<0.05$). In the second case (TREE2), we selected the first principal components explaining at least the 90% of system variability while in the last case (TREE3) we considered all the 17 parameters. Since the LVEF proved to be a non-invasive clinical parameter able to differentiate patients with heart failure [1,2], we added it to each the first three combinations obtaining others three combinations of parameters (TREE4, 5 and 6). In all cases, the training and test sizes were respectively 75% and 25% of the total number of data.

2.4. Cart method

The CART algorithm [30] is based on two steps: tree growing and tree pruning. In the former stage, the tree grows by selecting from all the possible splits, those that generate the nodes that contains elements of only one class. The large-tree is the outcome of this step. Successively, this tree is pruned according to a minimal cost-complexity function in which both the number of nodes and the misclassification probability are considered. The outcome of this step is the best sub-tree that minimizes the cost-complexity function. In the tree growing, to measure the impurity of each node, we adopted the Gini index criterion computed in this way:

$$\text{Gini index}(t) = \quad 1 - \left(\frac{n_i}{n}\right)^2 - \left(\frac{n_j}{n}\right)^2 \tag{1}$$

where *t* is the considered node, *i* and *j* are the two class labels, n_i and n_j are the number of subjects present at the node belonging either to the class *i* or *j,* respectively, and *n* is the total number of subjects at the node.

2.5. Performance measurements

To evaluate the performance of each classifier, we used the confusion matrices [38] and calculated the accuracy (ACC), sensitivity (SEN), specificity (SPE) and precision (PRE) as in Tab.1. Furthermore, the receiver operating characteristics curve (ROC) was used to depict trade-offs between hit rate (sensitivity) and false alarm rate (1.0-specificity) and the area under the curve (AUC) was evaluated [39]. In order to assess the influence of data included in the training and test sets, for each of the six situations, the classifiers were simulated 100 times changing randomly the training and test dataset. The maximum level for pruning was set to six. We selected the classifier producing the highest accuracy. The CART was implemented by using the methods and construction of the MATLAB class *fitctree*.

Table 1. Classification performance measures. TP: true positive, TN: true negative, FP: false positive, FN: false negative.

Measure	Definition
Sensitivity	$\frac{TP}{TP + FN}$
Specificity	$\frac{TN}{TN + FP}$
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$

Table 2. Input features utilized in the six cases. In bold the features selected by CART algorithm presented in the best sub-tree.

Tree Classifier	Input features
TREE1	**pNN50, Beta exponent, sex**, age
TREE2	1°Comp, 2°Comp, 3°Comp, **4°Comp**, **5°Comp**
TREE3	meanRR, SDNN, RMSSD, **NN50**, pNN50, LF, HF, LF/HF, LFn, HFn, **Beta exponent,** SD1, SD2, SD1/SD2, FD, sex, **age**
TREE4	**LF/HF**, sex, **LVEF**
TREE5	**1°Comp, 2°Comp, 3°Comp, 4°Comp, 5°Comp**
TREE6	meanRR, SDNN, RMSSD, NN50, pNN50, LF, HF, LF/HF, LFn, HFn, **Beta exponent,** SD1, SD2, SD1/SD2, FD, sex, age, **LVEF**

Table 3. Classification performances on test set of the three models not including LVEF (TREE 1-3) and of the other three models considering LVEF (TREE4-6). In bold the highest values. SEN: sensitivity, SPE: specificity, PRE: precision, ACC: accuracy, AUC: area under the curve.

	TEST SET				
	SEN (%)	SPE (%)	PRE (%)	ACC (%)	AUC (%)
TREE1	**18**	**99**	**89**	**84**	**67**
TREE2	15	99	59	79	62
TREE3	29	95	61	82	64
TREE4	**90**	**99**	**95**	**97**	**95**
TREE5	67	98	91	92	87
TREE6	90	98	93	96	94

3. Results

Table 2 shows the selected features used in the six considered classification trees. The first three classifiers refer to parameters without LVEF that was considered in the last three classifiers. In TREE1, we applied the stepwise regression to all the fifteen HRV parameters together with age and sex obtaining in the final multilinear model only four features. In TREE2, the principal component analysis highlighted that five components accounted for 91% of the variance in the dataset. Finally, the third classification tree used as inputs all the seventeen parameters. The addition of LVEF and the repetition of the selection techniques produced three new parameters for TREE4 and five principal components for TREE5. In TREE6, all the eighteen parameters were considered as inputs of the classification tree.

In Table 2, we reported in bold the features used by each best tree found in the sixth conditions presenting the highest accuracy among 100 different combinations of data for training and test.

Table 3 shows the classification performances together with the area under the curve (AUC) in the test phase of all the six classification trees. If the LVEF was added to the input, higher performances values (about 13% for accuracy) than those obtained without this clinical parameter, were achieved. The best accuracy, both in case with or without LVEF, was achieved by using parameters selected with stepwise regression (TREE1 and TREE4). The TREE4 tree produced the best performance with an accuracy of 97%. Figure 1 shows the ROC curves of TREE1 and TREE4 in the test phases with AUC values of 67% and 95%, respectively. Figures 2 and 3 show the TREE representations of the TREE1 and TREE4 highlighted that only two parameters (LF/HF, LVEF) are sufficient to distinguish with very high accuracy DCM patients from normal subjects.

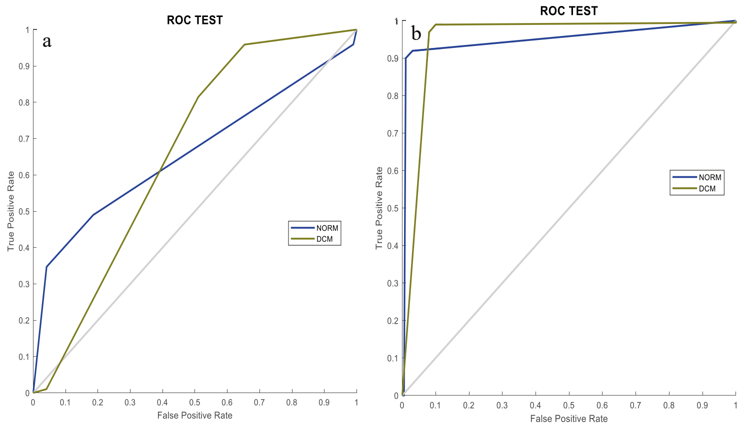

Fig. 1. (a) ROC Test of TREE1 AUC=67%; (b) ROC Test of TREE4 AUC=95%.

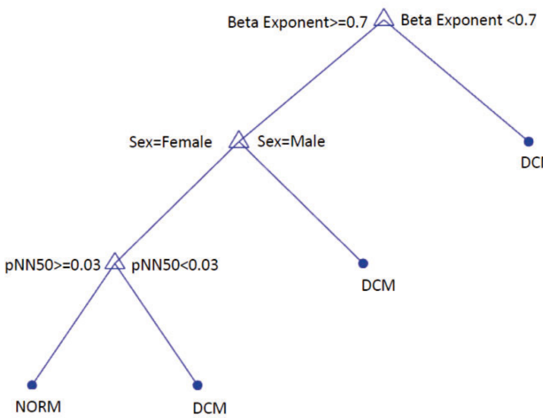

Fig.2. The model of the best tree not considering LVEF (TREE1) with ACC=84%.

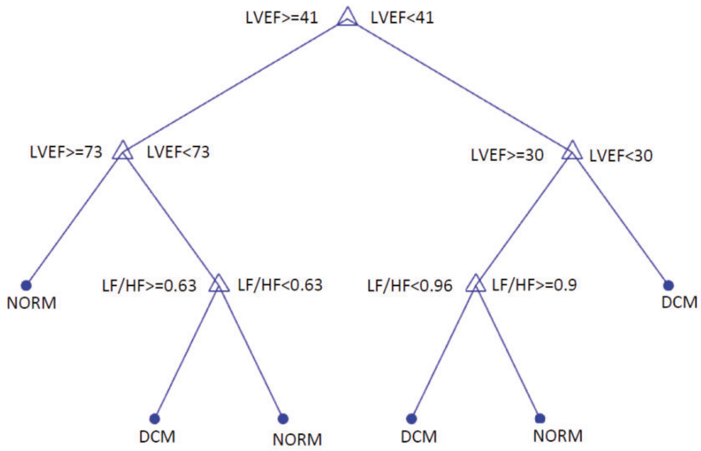

Fig.3. The model of the best tree considering LVEF (TREE4) with ACC=97%.

4. Discussion

Heart rate variability is reduced in subjects suffered from cardiovascular disease and particularly in patients affected by dilated cardiomyopathy presenting lower values of linear and non-linear HRV parameters [6-10]. Nowadays, different mathematical approaches for decision support systems have been proposed for the automatic classification of heartbeats and cardiovascular diseases [13-27], among them the classification and regression tree analysis represents a non-parametric decision tree methodology that has the ability to efficiently segment populations into meaningful subgroups.

Until now authors identified only segments or beats of dilated cardiomyopathy type using HRV parameters as input of complex classifier in which the CART was combined with other machine learning techniques [28,29], producing high accuracies (greater 95%). In order to identify subjects suffered from DCM and not only single segments or beats, in this paper, we evaluate in a large cohort of subjects, the performance of some classification trees considering as inputs the features obtained by using pre selection techniques decreasing the complexity of the model maintaining the same variance. At first, we considered all the linear and non-linear HRV parameters together with age and sex and successively we added the LVEF, a prognostic indicator of DCM patients.

Without considering LVEF, the best performance was achieved by TREE1 that had as inputs pNN50, Beta exponent, sex and age with an accuracy of 84% and an AUC of 64% in the validation test.

Considering the LVEF, the TREE4 with the three parameters selected by stepwise regression (LF/HF, sex and LVEF), presented the highest accuracy (97%) and AUC (95%) in the test phase. Analyzing the performances of the TREEs in which the features were selected by principal component analysis, we observed lower accuracy values suggesting that this pre-selection technique is worse than stepwise regression in the discrimination between DCM and normal subjects. On the contrary, considering all the parameters, we noted performances quite similar to those achieved using stepwise regression especially if LVEF was considered. Finally, comparing the performances of TREE1-TREE3 with TREE4-TREE6, we observed an increase in accuracy of 13% and in AUC of 28% if the LVEF clinical parameter was added, supporting the relevance of this non-invasive clinical parameter.

The proposed technique highlighted the powerful capability of selected linear and nonlinear features with age, sex and LVEF as inputs of CART technique in differentiating normal from DCM patients. Moreover, for the specific large dataset examined, we achieved a very accurate classification of subjects (97%) considering as inputs only LVEF, sex and LF/HF.

Our future work would examine other machine learning technique such as ANN or single vector machine applied to pre-selected parameters to evaluate their ability to discriminate normal and subjects that have different pathologies.

Acknowledgements

Work partially supported by Master in Clinical Engineering, University of Trieste.

References

[1] Perry, Elliott. (200) "Diagnosis and management of dilated cardiomyopathy," *Heart* **84**:106-112.

[2] Bozkurt, Biykem, Monica Colvin, and Jennifer Cook. (2016) "Current diagnostic and treatment strategies for specific dilated cardiomyopathies. A scientific statement from the American Heart Association." *Circulation* **134**:e579-e646.

[3] Acharya U., Rajendra, Joseph K. Paul and Natarajan Kannathal. (2006) "Heart rate variability: a review." *Medical Biological Engineering and Computing* **44(3151):**1031–1051.

[4] Guzzetti, Stefano, Renata Magatelli and Ester Borroni. (2001) "Heart rate variability in chronic heart failure." *Autonomic Neuroscience Basic and Clinical* **90(64)**:102–105.

[5] Hadase, Mitsuyoshi, Akihiro Azuma and Kan Zen. (2004) "Very low frequency power of heart rate variability is a powerful predictor of clinical prognosis in patients with congestive heart failure." *Circulation Journal* **68(49):**343–347.

[6] Yi, Gang, Jonathan H. Goldman and Philip J. Kelling. (1997) "Heart rate variability in idiopathic dilated cardiomyopathy: relation to disease severity and prognosis." *Heart* **77**:108-114.

[7] Carvajal, Raul, Niels Wessel and Montserrat Vallverdu. (2005) "Correlation dimension analysis of heart rate variability in patients with dilated cardiomyopathy." *Computer Methods and Programs in Biomedicine* **78**:133-140.

[8] Wessel, Niels, Christine Bruckner and Horst Malberg. (1999) "Long-term symbolic dynamics for heart arte variability analysis in patients with dilated cardiomiophaty." *Computers in Cardiology* **26**; 253-256.

[9] Malberg, Hagen, Robert Bauernschmitt and Udo Meyerfeldt. (2003) "Short-term heart rate turbulence analysis versus variability and baroreceptor sensitivity in patients with dilated cardiomyopathy." *German Journal of Cardiology* **92**:547-557.

[10] Voss, Andreas, Steffen Schulz and Rico Schroeder. (2009) "Methods derived from nonlinear dynamics for analyzing heart rate variability." *Philosophical transaction of the royal society* **367**: 277-296.

[11] Taye Getu Tadele, Eun Bo Shim and Han-Jeong Hwang (2019) "Machine learning approach to predict ventricular fibrilation based on QRS complex shape" *Frontiers in Physiology* **10**:1193.

[12] Li Qiao, Cadathur Rajagopalan and Gari D Clifford "Ventricular fibrillation and tachycardia classification usign a machine learning approach" *IEEE Transaction on biomedical engigneering* **61**:1607-13.

[13] Christov, Ivaylo, German Gómez-Herrero, and Vessela Krasteva. (2006) "Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification," *Mediacla Engineering & Physics* **28**:876–887.

[14] Rodríguez-Sotelo, Jose Luis German Castellanos-Domínguez and Carlos D. Acosta-Medina. (2012) "Recognition of Cardiac Arrhythmia by Means of Beat Clustering on ECG-Holter Recordings", in Intech (eds) *Advances in Electrocardiograms—Methods and Analysis*.

[15] Llamedo, Mariano and Juan P. Martinez. (2011) "Heartbeat classification using feature selection driven by database generalization criteria," IEEE *Transaction Biomedical Engineering* **5**:616–625.

[16] Jekova, Ivaylo, Giovanni Bortolan and Irena Christov. (2004) "Pattern recognition and optimal parameter selection in premature ventricular contraction classification," *Compututers in Cardiol*ogy **31**:357–360.

[17] Behadada, Omar and Mohammed A. Chikh. (2013) "An interpretable classifier for detection of cardiac arrhythmias by using the fuzzy decision tree," *Artificial Intelligence Research* **2**:45–58.

[18] Wieben, Oliver, Valtino X. Afonso and Willis J. Tompkins. (1999) "Classification of premature ventricular complexes using filter bank features, induction of decision trees and a fuzzy rule-based system" *Medical and Biological Engigneering Computing* **37**:560–565.

[19] Kannathal, Nataraja, Puthusserypady Sadasivan, and Choo Min Lim. (2005) "Cardiac state diagnosis usign adaptive neuro-fuzzy technique" *Proceedings IEEE Engigneering in Medicine and Biology* **1**:3864-3867.

[20] Jiang, Wei and Kong Seong G. (2007) "Block-based neural networks for personalized ECG signal classification" IEEE *Transactions on Neural Networks* **18**:1750–1761.

[21] Acharya, Rajendra U, Sathish Kumar,and Subbanna Bhat (2004) "Classification of cardiac abnormalities usign heart rate signals" *Medical and Biological Engigneering and Computing* **42**: 288-293.

[22] Adebita, Emmanuel, Veronica C. Iweanya and Segun I. Popoola.(2017) "Automated detection of heart defects in athletes based on electrocardiography and artificial neaural network." *Cogent Engineering* **4**:1411220

[23] Megat Ali, Amin, Chep Z. A. Che Zainal, C. C. and Afif. (2012) "Detection of cardiomyopathy using multilayered perceptron network", *Proceedings 2012 IEEE International Colloquium on Signal Processing and its Applications* **1**:436-440.

[24] Khazaee, Ali and Ata Ebrahimzadeh. (2010) "Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features," *Biomedical Signal Processing and Control* **5**: 252–263.

[25] Acir, Nurettin (2006) "A support vector machine classifier algorithm based on a perturbation method and its application to ECG beat recognition systems," *Expert Systems with Applications* **31**:150–158.

[26] Ghosh Digvijay, Bijoy L. Midya and Chitanjib Koley. (2005) "Wavelet aided SVM analysis of ECG signals for cardiac abnormality detection", *Proceedings 2005 Annual IEEE India Conference* **1**:9-13.

[27] Shukri M.H. Ahmad, Amin Megat Ali and Mah Z Noor. (2012) "Investigation on Elman neural network for detection of cardiomyopathy", *Proceedings IEEE Control and System Graduate Research Colloquium* **1**:328-332.

[28] Mythili Thirugnanam, and Megana P Santhoshi. (2020) ‘Hybrid feature extraction and stacking based ensemble classifier model for cardiomiopathy classification” *International Journal of Advanced Science and Technology* **29**:1396-1413.

[29] Mhesh, Venkatachalam, Arumugam, Kandaswamy and Sathish Balakrishnan (2010) “Cardiac disease classification usign heart rate signals” *International Journal of Electronic Healthcare* **5**:211-230.

[30] Breiman, Leo, Jerome Fredman and Charles J Stone (1984) “Classification and regression trees”, in Wadsworth International Group (eds) *The Wadsworth Statistics/Probability Series.*

[31] Roche, Federic, Jean-Michel Gaspoz, and Isabelle Court-Fortune. (1999) “Screening of obstructive sleep apnea syndrome by heart rate variability analysis” Circulation **100(121)**:1411–1415.

[32] Fornasa, Elisa, Agostino Accardo, and Marco Cinquetti. (2014) “HRV Spectral and Fractal Analysis in Heart Failure Patients with Different Aetiologies” Computing in Cardiology **41**:421-424.

[33] Pagani, Massimiliano, Federico Lombardi, and Stefano Guzzetti. (19849 “Power spectral density of heart rate variability as an index of sympatho-vagal interaction in normal and hypertensive subjects” Journal of Hyperthension **2**:S383-5.

[34] Goldberger, Ary L. (1996) “Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside” Lancet **347**:1312-14.

[35] Woo, Mary A., Willima G. Stevenson and Debora K. Moser. (1992) “Patterns of beat-to-beat heart rate variability in advanced heart failure.” America Heart Journal **123**:704-710.

[36]Higuchi, Toshiro. (1988) “Approach to an irregular time series on the basis of the fractal theory.” Physica D. **31**:277-283.

[37]Cusenza, Monica, Agostino Accardo and Giovanni D'Addio. (2010) “Relationship between fractal dimension and power-law exponent of heart rate variability in normal and heart failure subjects” Computing in Cardiology **37**:935-938.

[38] Provost, Foster and Ron Kohavi. (1998) “On applied research in matching learning.” Machine Learning **30**:127-132.

[39] Hajian-Tilaki, Karimollah. (2013) “Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation,” Caspian journal of internal medicine **4**:627–635.