

ECG Signals Classification Using Overlapping Variables to Detect Atrial Fibrillation

I. G. ZICCARDI^{1*}, A. A. REY² and W. E. LEGNANI³

Received on September 27, 2021 / Accepted on March 24, 2022

ABSTRACT. In the present work a method for the detection of the cardiac pathology known as atrial fibrillation is proposed by calculating different information, statistics and other nonlinear measures over ECG signals. The original database contains records corresponding to patients who are diagnosed with this disease as well as healthy subjects. To formulate the dataset the Rényi permutation entropy, Fisher information measure, statistical complexity, Lyapunov exponent and fractal dimension were calculated, in order to determine how to combine this features to optimize the identification of the signals coming from ECG with the above mentioned cardiac pathology. With the aim to improve the results obtained in previous studies, a classification method based upon decision trees algorithms is implemented. Later a Montecarlo simulation of one thousand trials is performed with a seventy percent randomly selected from the dataset dedicated to train the classifier and the remaining thirty percent reserved to test in every trial. The quality of the classification is assessed through the computation of the area under the receiver operation characteristic curve (ROC), the F1-score and other classical performance metrics, such as the balanced accuracy, sensitivity, specificity, positive and negative predicted values. The results show that the incorporation of all these features to the dataset when are employed to train the classifier in the training task produces the best classification, in such a way that the largest quality parameter is achieved.

Keywords: Rényi entropy, statistical complexity, Fisher information, Lyapunov exponent, fractal dimension, atrial fibrillation, decision trees.

1 INTRODUCTION

The atrial fibrillation (AF) can be characterized as a heart condition that is associated with an irregular and often abnormally fast heart rate. Under normal conditions, the heart beats are com-

*Corresponding author: Ignacio Gabriel Ziccardi – E-mail: iziccardi@frba.utn.edu.ar

¹Centro de Procesamiento de Señales e Imágenes (CPSI), Facultad Regional Buenos Aires, Universidad Tecnológica Nacional, Av. Medrano 951, 1179, Ciudad Autónoma de Buenos Aires, Argentina – E-mail: iziccardi@frba.utn.edu.ar
<https://orcid.org/0000-0003-3804-0334>

²Centro de Procesamiento de Señales e Imágenes (CPSI), Facultad Regional Buenos Aires, Universidad Tecnológica Nacional, Av. Medrano 951, 1179, Ciudad Autónoma de Buenos Aires, Argentina – E-mail: arey@frba.utn.edu.ar
<https://orcid.org/0000-0002-9185-1382>

³Centro de Procesamiento de Señales e Imágenes (CPSI), Facultad Regional Buenos Aires, Universidad Tecnológica Nacional, Av. Medrano 951, 1179, Ciudad Autónoma de Buenos Aires, Argentina – E-mail: walter@frba.utn.edu.ar
<https://orcid.org/0000-0002-6949-0728>

posed by the contraction and release of its muscular walls, in such a way that to force the blood to circulate in the whole body. After pump, the heart goes into relax stage to fill itself with blood again. This process is repeated every time the heart beats. In the presence of atrial fibrillation, the heart's upper chambers (atria) contract randomly and sometimes so fast that the heart muscle cannot relax properly between contractions. This condition reduces the heart's efficiency and its performance in general. When the electrical impulses trigger fire the atria become at abnormal rate, the AF takes place in the dynamics of the heart. This affect the heart mechanics and may develop into a serious clinical situation. Both imaging and electrophysiological studies can be used to monitor AF. It is very important that once AF has been detected, the patient is followed up so that the physician can determine the most appropriate therapy to normalise the functioning of the atrial chamber of the heart. The main objective of the present research is to formulate a feature space that contributes with the work of the physician in the diagnose of the AF. To achieve this objective some magnitudes are calculated from electrocardiogram (ECG) records of the Physionet database to constitute a proper dataset, later used by classification algorithms based on decision trees. Parameters calculated from the ECG records can be joined in two distinct groups, one associated with informational measures such as the statistical complexity, Rényi entropy and Fisher information measure, and other one formed by characteristics usually employed in nonlinear signal analysis as it is the case of the Lyapunov exponent and the fractal dimension.

With the intention to avoid assumptions on the distribution of the data, signal ordinal patterns are applied to define the probability density function, since their use only requires the comparison of neighbouring values taken from the signal. In this way, the calculation is made directly on actual data without the need of calibration or pre-filtering. For some chaotic dynamic systems it is demonstrated that the calculations based on entropies behave similarly to the Lyapunov exponent and are particularly useful in the presence of dynamic or observable noise. The great advantage of computing different types of signal characteristics is that they bring the possibility to detect abrupt changes as well as local variations.

During the investigation, an evident overlapping of the features appears, that looks like ruins the differentiation between the two groups of normal ECG records and those with presence of AF. However, this fact has been overcome by the classification stage very satisfactorily.

As mentioned above, the methodology applied to classify the signals is based upon decision trees, that for its simplicity, easy implementation, and no necessity to adjust any parameter, results very appropriate to be selected. In many papers the area under the receiver operation characteristic curve, the F1-score and the accuracy are employed as quality classification parameters, in this work the balanced accuracy is incorporated due to the imbalanced number of samples in the database of the normal sinus rhythm and AF. This research is organized in the following sections: introduction (this section), section 2 is devoted to describe atrial fibrillation, section 3 contains a description of the magnitudes used to construct the feature space and presents the classification method applied, section 4 is dedicated to analyze the results, section 5 summarises the conclusions, and finally the end of the work contains the bibliography.

2 ATRIAL FIBRILLATION

One of the common arrhythmias is known as AF which is considered as a significant cause of mortality, especially among elderly people. It occurs in about 10% of those persons whose age is over 75 years old. In presence of AF, the electrical activity of the atria is irregular and the depolarization occurs at a frequency of 300-600 bpm. This behavior does not result in effective atrial contraction, but it simply means a wave effect on the muscle, named fibrillation. Ventricular activity is also affected, since the impulses are conducted sporadically through the atrioventricular node, involving a considerable time for the ventricles to be filled between each beat. As a consequence, there exists a marked irregularity in a characteristic pulse in terms of frequency and volume. According to the type of AF, they can be grouped into persistent, permanent, or paroxysmal. This abnormal heart rhythm is usually caused by mitral valve disease, ischemic heart disease, thyrotoxicosis, hypertension, or alcoholism. The development of thrombi (blood clots) in the left atrium is predisposed by the lack of effective atrial contraction and the resulting blood stasis, which can lead to the passage of emboli that may cause ischemic strokes. Due to AF is the main cause of stroke in the elderly, its study should not be overlooked. Some patients with AF have palpitations or may even have dizziness or syncope (fainting) [4]. There are two possible treatments for this disease, one consists of the control of frequency with the purpose to reduce the ventricular frequency, while the other one is focused on the carefulness of rhythm in order to recover the sinus rhythm. None of them is agreed to be the best selection among the medical community.

The ECG was first introduced into clinical practice about one hundred years ago by Einthoven, and it is basically a linear record of the electrical activity of the heart that develops along the time. For every cardiac cycle, both atrial and ventricular depolarization wave, as well as a ventricular repolarization wave, are successively recorded. They are respectively known as P wave, QRS complex and T wave. The intervals between the waves within two successive cycles vary depending on the heart rate. The ECG is the technique of choice for the study of patients with precordial pain, syncope, palpitations and acute dyspnea. On the other hand, it is extremely important for the diagnosis of cardiac arrhythmias, conduction disturbances, pre-excitation syndromes and channelopathies. Likewise, it is fundamental to evaluate the evolution and response to treatment of all types of heart conditions and other diseases, as well as different situations such as electrolyte imbalances, administration of drugs, effects and results of sport, surgical evaluation, among others. It is also useful for epidemiological and control studies (clinical checks). Even though the mentioned high utility of ECG, an incorrect diagnosis may be made when doctors rely in excess on normal ECG registers, since almost the ten percent of acute coronary syndromes present a normal record, especially at the beginning of the disease. Even more, some subtle alterations of the ECG can be observed without evidence of heart disease. In those cases, the involved professional should be cautious in the ruling out of some diseases such as ischemic heart disease, channelopathies (*e.g.* long QT and Brugada syndrome) or pre-excitation syndromes, before considering a non-specific alteration. Therefore, a good practise requires to read the ECG signal in consideration with the clinical context and possibly the necessity to make additional sequential

records [2]. Besides, normal variations can be also observed in the ECG record that are related to constitutional habit, chest wall malformations or age condition. Even more, transitory alterations can be detected due to a series of diverse causes as it is the case of hyperventilation, hypothermia, glucose regulation or alcohol intake, ionic alterations or the effect of certain drugs [2].

3 MATERIAL AND METHODS

In the present work, the MIT-BIH arrhythmia database of PhysioNet is the source of data which is available at <http://www.physionet.org>. This set is composed by ECG signals from healthy patients and many kinds of arrhythmia (included AF). The database provides 283 normal sinus ECG signal fragments taken from 23 patients and 135 with AF from 6 patients.

These ECG study consisted of 3600 non-overlapping samples taken from a thousand randomly selected ECG fragments acquired with a sampling frequency of 360 Hz and a gain of 200 adu/mV at the main ECG position, corresponding to forty five individuals, nineteen women (in the range of twenty three and eighty nine years old) and twenty six men (in the range of thirty two and eighty nine years old) [15].

The aim of the present work is to classify the database mentioned above in two classes: “Fibrillation” and “Normal” according respectively, to the presence or not of detected atrial fibrillation in the ECG recorded. Thus, a crucial task is to determine the features from which the classifier is going to learn. In what follows, the five signal characteristics to be considered are briefly described.

Lyapunov exponent. The behavior of a time series can be characterized in part by the corresponding Lyapunov exponent (LE) denoted by λ . This parameter brings information about the exponential divergence of the orbits in the phase space in the case of a chaotic process. Many of the methods usually applied to compute the LE require the estimation of the attractor embedding dimension as well as the time delay to reconstruct it, and other similar parameters. In this work it was selected one of the methods that not demands any other parameters than the sampling frequency of the signal. This is the case of the method proposed by Kantz in 1994 [9] which provides the maximal LE in a robust way only using the time between samples. The algorithm is based upon measures taken directly over the signal and for this reason it is clear and easy to use. Despite there are other algorithms (*cf.* [3, 14]) which are more precise than the Kantz methodology to compute the LE, this robust algorithm is widely used, its success has been largely probed and it can be applied in the presence of noisy signals and in the case of short data series.

Fractal dimension. The characteristic roughness of a one dimensional signal can be computed using the Fourier frequency spectrum. The outputs of this methodology can provide noisy fluctuations of the results, which is why an average of the power spectrum should be taken over a long interval of the signal to obtain stable values. Due to the statistical constituents of many signals, that generally vary over small time intervals, it would not be appropriate to use this technique in all the cases. An alternative to solve this problem, applied usually in nonlinear analysis, is to

compute the fractal dimension (FD) denoted by D . There are several proposals to estimate the FD, such as the algorithm introduced by Higuchi [7] which is the one adopted in this work.

In order to calculate the FD of a time series $Y = \{y_t\}_{t=1}^N$ a subsequence of length k is constructed as follows:

$$Y_k^h = \{y_h, y_{h+k}, \dots, y_{h+\eta}\},$$

for all $h = 1, 2, \dots, k$, where $\eta = [(N - h)/k]k$ and $[\cdot]$ indicates the integer part of a number. Then, on every one of these subsequences the following length is calculated:

$$L^h(k) = \frac{N - 1}{\eta} \sum_{i=1}^{\eta/k} |y_{h+ik} - y_{h+(i-1)k}|,$$

where $(N - 1)/\eta$ is a normalization factor of the length of the signal. Finally, to calculate the FD a regression must be performed on the averages of the lengths of the subsequences because these will follow a law of the type $L(k) \propto k^{-D}$. The only parameter should be set up to compute of the Higuchi FD is the subsequence length k , in such a way that the whole algorithm runs over the signal data points.

Rényi entropy. Let $\mathcal{P} = \{p_i\}_{i=1}^n$ be a discrete probability density function (pdf). The Rényi entropy [12] of order α , which is an extension of the well-known Shannon entropy is defined as:

$$S_\alpha^R[\mathcal{P}] = \frac{1}{1 - \alpha} \ln \sum_{i=1}^n p_i^\alpha,$$

where $\alpha \in \mathbb{R}^+$ and $\alpha \neq 1$. The normalized Rényi entropy is obtained by:

$$H_\alpha^R[\mathcal{P}] = \frac{S_\alpha^R[\mathcal{P}]}{\ln n}, \tag{3.1}$$

In order to compute the entropy given by equation (3.1) the selected pdf is the one proposed by Bandt and Pompe [1], which is based on the ordinal dynamism of the observations in a chronological sequence. In this approach, two parameters are necessary: the embedding dimension m which determines the length of the subsequence to be associated with the permutations, and the embedding time delay τ which measures the distance between two consecutive observations in the subsequence. Explicitly, for a time series $\{y_t\}_{t=1}^N$, overlapping partitions of length m ($N \gg m!$) are formed as follows:

$$s \rightarrow \{y_s, y_{s+\tau}, \dots, y_{s+(m-1)\tau}\},$$

with $s = 1, \dots, N - m + 1$. If π is one of the $m!$ permutations of the set $\{0, 1, \dots, m - 1\}$, s is said of type π if $y_{\pi(s)} \leq y_{\pi(s+\tau)} \leq \dots \leq y_{\pi(s+(m-1)\tau)}$. Thus, the pdf of the permutations is defined as:

$$p_i = P(\pi_i) = \frac{\#\{s \text{ is of type } \pi_i\}}{N - m + 1}. \tag{3.2}$$

Statistical complexity. Given n events, $\mathcal{P}_e = (1/n, \dots, 1/n)$ is the uniform distribution which maximizes the Rényi entropy. In [10] the authors develop a way to measure statistical complexity in terms of the concept of distance to \mathcal{P}_e . A disequilibrium [11] is defined as:

$$Q[\mathcal{P}] = Q_0 \mathcal{D}[\mathcal{P}, \mathcal{P}_e],$$

where the distance is given by:

$$\mathcal{D}[\mathcal{P}, \mathcal{P}_e] = \frac{1}{2(\alpha-1)} \left[\ln \sum_{i=1}^n p_i^\alpha \left(\frac{p_i + n^{-1}}{2} \right)^{1-\alpha} + \ln \sum_{i=1}^n \frac{1}{n^\alpha} \left(\frac{p_i + n^{-1}}{2} \right)^{1-\alpha} \right],$$

and the normalization constant is

$$Q_0 = \frac{1}{2(\alpha-1)} \ln \left[\frac{(n+1)^{1-\alpha} + n-1}{n} \left(\frac{n+1}{4n} \right)^{1-\alpha} \right].$$

Thus, the complexity using the normalized Rényi entropy as the amount of disorder, can be expressed as:

$$C_\alpha^R[\mathcal{P}] = Q[\mathcal{P}] H_\alpha^R[\mathcal{P}].$$

Fisher information. The entropy and the complexity being both global measures, cannot be sensible to abrupt changes in a small portion of a pdf. On the contrary, The Fisher information [5] is a local measure defined by:

$$F[f] = \int \frac{|\nabla f(x)|^2}{f(x)} dx,$$

where f is a continuous pdf and $\nabla f(x)$ denotes its gradient.

In [13], the authors propose ten different expressions to obtain the Fisher information in the discrete case. The mathematical expression to be adopted in the present work is

$$I_F[\mathcal{P}] = \sum_{i=2}^{n-1} (\sqrt{p_{i+1}} - \sqrt{p_{i-1}})^2. \quad (3.3)$$

Decision tree is a non-parametric supervised learning technique that can be applied either in classification or regression tasks. Basically, a decision tree formulates a statement and then makes a decision based on whether this statement is true or false. In other words, it learns simple decision rules inferred by the feature variables, which can be numerical, nominal, or a mixture of both types.

Among the strengths of decision trees it can be mentioned that they are simple to interpret, they can be visualized, they require little data preparation, their cost to prediction is logarithmic in the number of training data, they can easily explain an observable condition in a model and in general they perform well even under relax assumptions on the true model. On the opposite side, the generated tree may be extremely complex causing overfitting which can be avoided with a pruning mechanism, such as imposing the minimum number of elements in a leaf node or the maximum depth of the tree. Decision trees can be also unstable when small variations are introduced in the data. For more details on this algorithm it is referred to [6, 8].

Applying the decision tree technique, seven classification models are proposed in terms of different combinations of variables to setup the feature space (see Table 1).

Throughout the rest of this work, the presence of fibrillation is considered as the positive class. The classical notation to be used is TP for true positive, TN for true negative, FP for false positive

Table 1: Proposed classifier models.

Model	Feature space
LDH	(λ, D, H_α^R)
LDC	(λ, D, C_α^R)
LDF	(λ, D, I_F)
LDHC	$(\lambda, D, H_\alpha^R, C_\alpha^R)$
LDHF	$(\lambda, D, H_\alpha^R, I_F)$
LD CF	$(\lambda, D, C_\alpha^R, I_F)$
LDHCF	$(\lambda, D, H_\alpha^R, C_\alpha^R, I_F)$

and FN for false negative cases. With the aim of studying the contribution of the variables in the performance of the classifier, the following well-known measures are computed.

1. Balanced Accuracy, $BA = (TP/P + TN/N)/2$, is the mean of the TP and TN rates.
2. F1-score, $F1 = 2TP/(2TP + FP + FN)$ is the harmonic mean of recall and precision.
3. Area Under the Curve (AUC) for the Receiver Operating Characteristic (ROC) curve that relates the rates of TP and FP.
4. Sensitivity is the rate of TP.
5. Specificity is the rate of TN.
6. Positive Predicted Value, $PPV = TP/(TP + FP)$, is the proportion of TP among the positive prediction.
7. Negative Predicted Value, $NPV = TN/(TN + FN)$, is the proportion of TN among the negative prediction.

All the previous indicators ranges in value from 0 when all the predictions are wrong, to 1 when the model predicts with total accuracy.

4 RESULTS AND DISCUSSION

Throughout what follows, $m = 5$ and $\tau = 5$ are fixed for the involved calculus in the probability of permutations given by (3.2). Moreover, the order for normalized Rényi entropy (RE) and for the statistical complexity (RC) is $\alpha = 4$. To unify notation, FI indicates the Fisher information defined in (3.3). In addition, $k = 20$ was used for the computation of the FD.

Figure 1 graphically represents the scatter plot defined by the selection of two variables in the feature space. As it can be noticed in all the possible combinations, the clouds composed by fibrillation and normal ECGs show a considerable overlapping. The same behaviour can be also understood in the histograms of Figure 2 where some variables, such as FD, RE and FI, show a multimodal frequency distribution and a notorious overlapping of their respective plots.

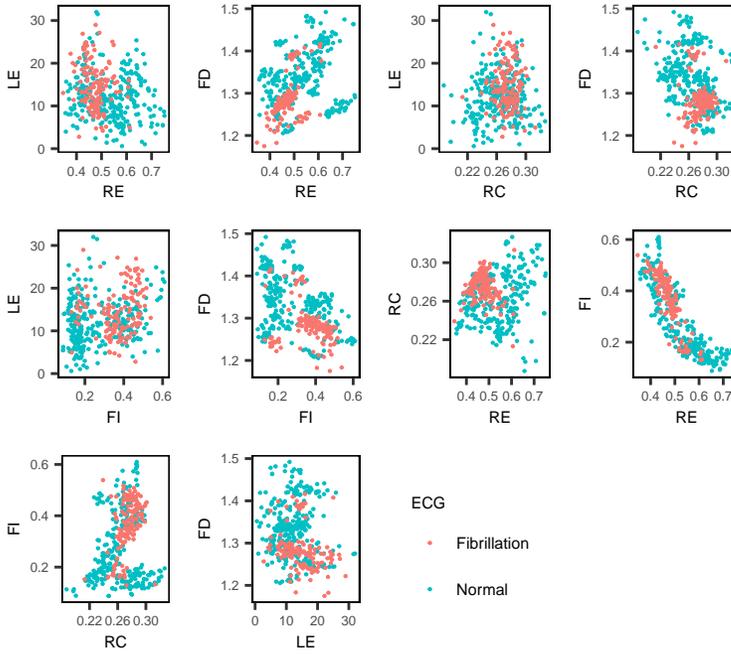


Figure 1: Scatter plots defined by two features.

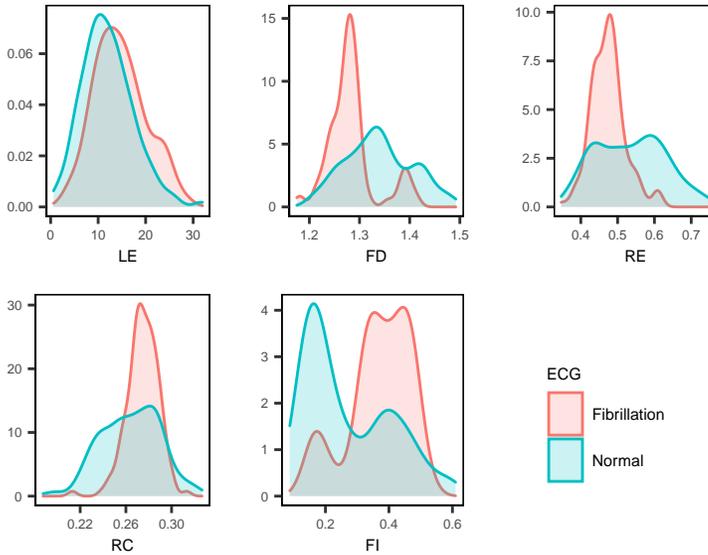


Figure 2: Histograms of every feature.

The classifying models are training with the set formed by the random selection of the 70% of normal ECG records plus the 70% of ECG with fibrillation, in order to get a dataset with more balanced classes. Meanwhile, the remaining ECGs from each type are considered to test the models. Due to the random context, a simulation of Montecarlo with 1000 trials is performed. Table 2 exhibits the mean (μ) and standard deviation (σ) of the quality classification measures introduced in the previous section per model. All the indicators achieve a desirable result, specially the specificity (91.3%), the BA (84.6%) and the AUC (84.6%) for the model LDHCF. It is worth to recall the context of study in which the classes of ECG do not show a remarkable condition to be separable groups.

It can be observed that the model that best learns is the one that incorporate altogether the five features. However, the ability to detect fibrillation is not as good as in normal ECG identification; *i.e.* the sensitivity (78%) is lower that the specificity (91.3%). This fact is also reflected in the value of F1 (79.5%). This remarks are also shown in the boxplots on Figures 3, 4 and 5.

Table 2: Mean (μ) and standard deviation (σ) of the quality classification measures after 1000 Montecarlo trials of each model. Best values per measure are highlighted in green.

Model	Sensitivity		Specificity		PPV		NPV		BA		F1-score		AUC	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
LDH	0.723	0.088	0.877	0.042	0.745	0.059	0.870	0.034	0.800	0.039	0.729	0.051	0.800	0.039
LDC	0.657	0.095	0.868	0.050	0.715	0.072	0.842	0.035	0.762	0.043	0.678	0.060	0.762	0.043
LDF	0.766	0.081	0.905	0.049	0.804	0.079	0.891	0.032	0.835	0.040	0.780	0.054	0.835	0.040
LDHC	0.771	0.081	0.904	0.040	0.802	0.064	0.893	0.033	0.838	0.038	0.782	0.049	0.838	0.038
LDHF	0.752	0.086	0.894	0.049	0.783	0.074	0.884	0.033	0.823	0.039	0.762	0.052	0.823	0.039
LDCF	0.789	0.078	0.897	0.041	0.792	0.063	0.899	0.032	0.843	0.037	0.787	0.048	0.843	0.037
LDHCF	0.780	0.072	0.913	0.041	0.818	0.065	0.897	0.028	0.846	0.033	0.795	0.044	0.846	0.033

Finally, an example of decision tree for the model LDHCF is illustrated in Figure 6. The corresponding values for the quality measures are: sensitivity 0.780, specificity 0.976, PPV 0.941, NPV 0.902, BA 0.878, F1 0.853 and AUC 0.878.

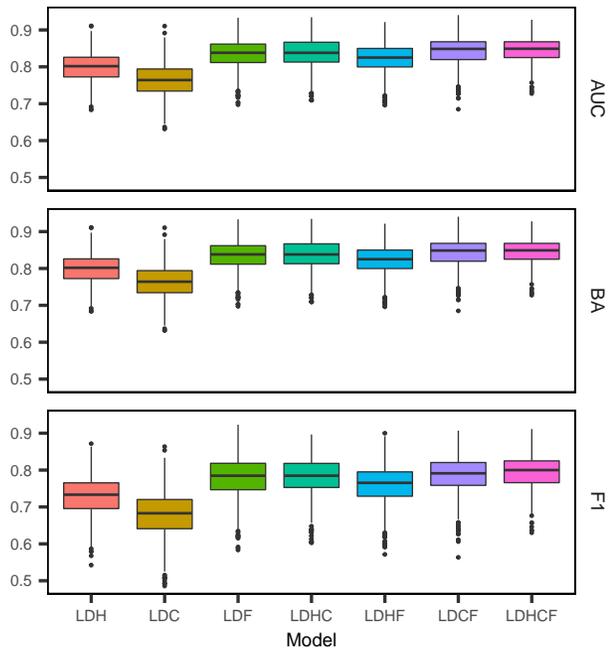


Figure 3: Distribution of AUC, BA and F1 values per model after a simulation of Montecarlo with 1000 trials.

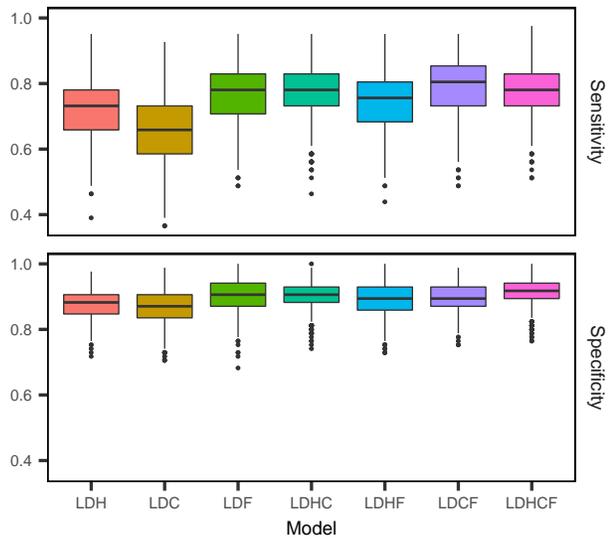


Figure 4: Distribution of sensitivity and specificity values per model after a simulation of Montecarlo with 1000 trials.

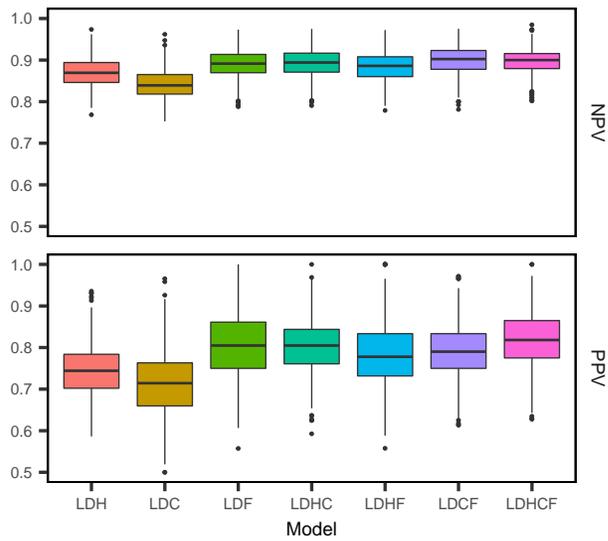


Figure 5: Distribution of NPV and PPV per model after a simulation of Montecarlo with 1000 trials.

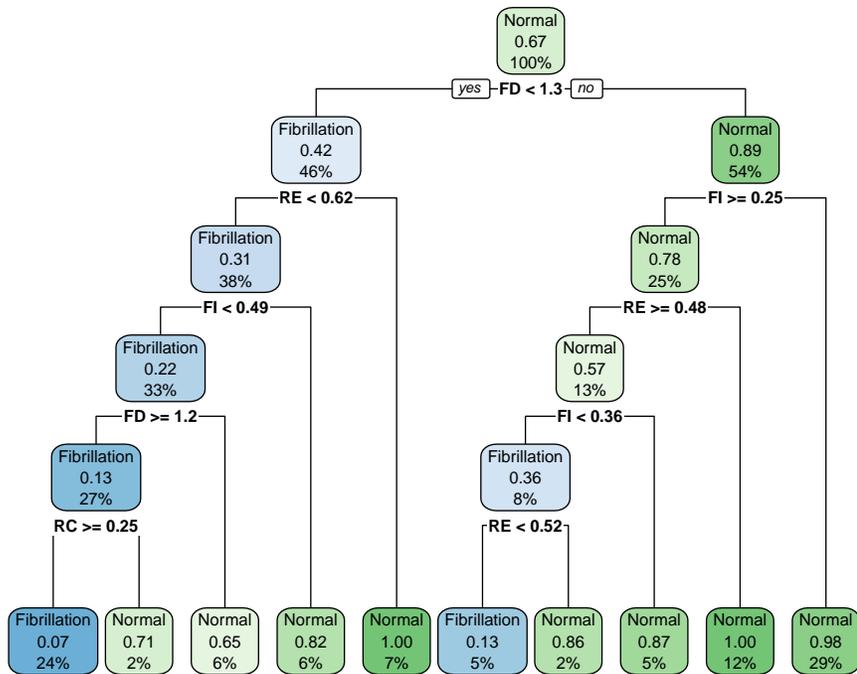


Figure 6: Decision tree for the model LDHCF.

5 CONCLUSIONS

The aim of the present work is to analyze the performance of a classifier in order to detect atrial fibrillation in a sample of ECG signals provided by the PhysioNet platform under the management of the MIT Laboratory for Computational Physiology, in Cambridge, USA. Thus, not only the selection of a good model but a suitable set of variables has been considered.

Decision tree algorithm has shown to be a proper choice in the classification procedure applied to the database under study due to its simplicity and low computational cost in the case of dichotomous clustering. In previous analysis [16], the considered variables belong to the entropic measures family, explicitly permutation and Rényi entropy were used. The incorporation of additional information measures of global and local character, such as complexity and Fisher information, respectively, in addition to measures that take into account the non-linear nature of the signal, as the case of Lyapunov exponent (global property) and fractal dimension (local property), showed a wide improvement in the performance of the classifier. In this way, the combination of entropic measures, joint with non-linear signal invariants, emerge as adequate variables to construct the feature space from which the classifier works. This can be seen since the model LDHCF, which includes the five variables, has achieved the best performance.

Despite the notorious overlapping of the selected variables that define the data set used in the classification, the computed quality measures give very acceptable results in the application of the proposed algorithm to the involved database.

From a biomedical point of view, the present proposal has thrown less false negative cases than false positive cases. This effect suggests the continuity of a deeper study in this line of research which may include the consideration of other kinds of variables to be part of the feature space. Moreover, the proposed approach could be tested on larger population samples or with the purpose to distinguish among different types of cardiac conditions in a in multi-class context.

Acknowledgments

The authors wish to express their gratitude to the student scholarship program of the Universidad Tecnológica Nacional Facultad Regional Buenos Aires and its Secretary of University Affairs and to the Secretary of Science and Technology and Productive Innovation. This work was developed within the framework of the projects PID UTN 4729 and PID UTN 8120.

REFERENCES

- [1] C. Bandt & B. Pompe. Permutation entropy: a natural complexity measure for time series. *Physical review letters*, **88**(17) (2002), 174102.
- [2] A. Bayés de Luna. “Electrocardiografía básica Patentes ECG normales y anormales”. Blackwell, Oxford (2007).
- [3] J.P. Eckmann, S.O. Kamphorst, D. Ruelle & S. Ciliberto. Liapunov exponents from time series. *Physical Review A*, **34**(6) (1986), 4971.

- [4] J. Evans. “Lo esencial en sistema cardiovascular”. Elsevier Health Sciences (2013).
- [5] R.A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, **222**(594-604) (1922), 309–368.
- [6] J. Friedman, T. Hastie, R. Tibshirani *et al.* “The elements of statistical learning”, volume 1 of *10*. Springer series in statistics New York (2001).
- [7] T. Higuchi. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, **31**(2) (1988), 277–283.
- [8] G. James, D. Witten, T. Hastie & R. Tibshirani. “An introduction to statistical learning”, volume 112. Springer (2013).
- [9] H. Kantz. A robust method to estimate the maximal Lyapunov exponent of a time series. *Physics letters A*, **185**(1) (1994), 77–87.
- [10] R. López-Ruiz, H.L. Mancini & X. Calbet. A statistical measure of complexity. *Physics letters A*, **209**(5-6) (1995), 321–326.
- [11] M. Martin, A. Plastino & O. Rosso. Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A: Statistical Mechanics and its Applications*, **369**(2) (2006), 439–462.
- [12] A. Rényi. On measures of entropy and information. In “Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics”. University of California Press (1961), p. 547–561.
- [13] P. Sánchez-Moreno, R. Yáñez & J. Dehesa. Discrete densities and Fisher information. In “Proceedings of the 14th International Conference on Difference Equations and Applications. Difference Equations and Applications. Istanbul, Turkey: Bahçesehir University Press” (2009), p. 291–298.
- [14] M. Sano & Y. Sawada. Measurement of the Lyapunov spectrum from a chaotic time series. *Physical review letters*, **55**(10) (1985), 1082.
- [15] Ö. Yıldırım, P. Pławiak, R.S. Tan & U.R. Acharya. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Computers in biology and medicine*, **102** (2018), 411–420.
- [16] I. Ziccardi, P. Martinez & W. Legnani. Detection of atrial fibrillation by entropic calculation of ECG signals. In “Proceedings of VIII MACI 2021, Volume 8”. Asociación Argentina de Matemática Aplicada, Computacional e Industrial (2021), p. 637–640.

