Towards Prediction of Heart Arrhythmia Onset Using Machine Learning

Agnieszka Kitlas Golińska[1][0000−0001−8737−765X], Wojciech Lesiński[1][0000−0001−6628−9466], Andrzej Przybylski[2][0000−0002−8136−7589], and Witold R. Rudnicki[1][0000−0002−7928−4944]

[1] Institute of Informatics, University of Białystok, ul. Konstantego Ciołkowskiego 1M, 15-245 Białystok, Poland `akitlas@ii.uwb.edu.pl`
[2] Faculty of Medicine, University of Rzeszów, Poland

Abstract. Current study aims at prediction of the onset of malignant cardiac arrhythmia in patients with Implantable Cardioverter-Defibrillators (ICDs) using Machine Learning algorithms. The input data consisted of 184 signals of RR-intervals from 29 patients with ICD, recorded both during normal heartbeat and arrhythmia. For every signal we generated 47 descriptors with different signal analysis methods. Then, we performed feature selection using several methods and used selected feature for building predictive models with the help of Random Forest algorithm. Entire modelling procedure was performed within 5-fold cross-validation procedure that was repeated 10 times. Results were stable and repeatable. The results obtained (AUC=0.82, MCC=0.45) are statistically significant and show that RR intervals carry information about arrhythmia onset. The sample size used in this study was too small to build useful medical predictive models, hence large data sets should be explored to construct models of sufficient quality to be of direct utility in medical practice.

Keywords: Arrhythmia · Implantable Cardioverter-Defibrillators · Artificial Intelligence· Machine Learning · Random Forest.

1 Introduction

Some types of cardiac arrhythmia, such as VF (ventricular fibrillation) or VT (ventricular tachycardia), are life-threatening. Therefore, prediction, detection, and classification of arrhythmia are very important issues in clinical cardiology, both for diagnosis and treatment. Recently research has concentrated on the two latter problems, namely detection and classification of arrhythmia which is a mature field [1]. These algorithms are implemented in Implantable Cardioverter-Defibrillators (ICD) [2], which are used routinely to treat cardiac arrhythmia [3]. However, the related problem of prediction of arrhythmia events still remains challenging.

In recent years we have observed an increased interest in application of Machine Learning (ML) and artificial intelligence methods in analysis of biomedical

data in hope of introducing new diagnostic or predictive tools. Recently an article by Shakibfar et al. [4] describes prediction results regarding electrical storm (i.e. arrhythmic syndrome) with the help of Random Forest using daily summaries of ICDs monitoring. Authors then generated 37 predictive variables using daily ICD summaries from 19935 patients and applied ML algorithms, for construction of predictive models. They concluded that the use of Machine Learning methods can predict the short-term risk of electrical storm, but the models should be combined with clinical data to improve their accuracy.

In the current study ML algorithms are used for prediction of the onset of malignant cardiac arrhythmia using RR intervals. This is important problem, since the standard methods of prediction aim at stratification of patients into high- and low-risk groups using various sources of clinical data [3]. Then, the patients from the high-risk group undergo surgical implantation of ICD [3], which monitors the heart rate. The algorithms for identification of arrhythmia events implemented in these devices recognise the event and apply the electric signal that restarts proper functioning of the heart. Despite technological progress, inappropriate ICD interventions are still a very serious side-effect of this kind of therapy. About 10–30% of therapies delivered by ICD have been estimated as inappropriate [2]. These are usually caused by supraventricular tachyarrhythmias, T-wave oversensing, noise or non-sustained ventricular arrhythmias.

The goal of the current study is to examine whether one may predict an incoming arrhythmia event using only the signal available for these devices. If such predictions are possible with high enough accuracy, they might be communicated by ICD's to warn patients of incoming event, helping to minimise adverse effects or even possibly avoid them completely. One of the first studies considering this problem was a Master of Science thesis by P. Iranitalab [5]. In that study the author used time and frequency domain analysis of QRS-complex as well as R-R interval variability analysis for only 18 patients, but he concluded that none of these methods proved to be an effective predictor that could be applied to a large patient population successfully. This analysis was performed on normal (sinus) and pre-arrhythmia EGM (ventricular electrogram) data. The newest article considering prediction of ventricular tachycardia (VT) and ventricular fibrillation (VF) was published in September 2019 by Taye et al. [6]. Authors extracted features from HRV and ECG signals and used artificial neural network (ANN) classifiers to predict the VF onset 30 s before its occurrence. The prediction accuracy estimated using HRV features was 72% and using QRS complex shape features from ECG signals – 98.6%, but only 27 recordings were used for this study.

Other studies, which seem to be related [7–9] in fact consider different issues. In [7] authors investigate a high risk patients of an ICD and evaluate QT dispersion, which may be a significant predictor of cardiovascular mortality. They claim that QT dispersion at rest didn't predict the occurrence and/or reoccurrence of ventricular arrhythmias. In [8] authors proposed a new atrial fibrillation (AF) prediction algorithm to explore the prelude of AF by classifying ECG before AF into normal and abnormal states. ECG was transformed into spectrogram using

short-time Fourier transform and then trained. In paper [9], it seems like it's more about detection or classification than prediction of onset of arrhythmia. Authors used a clustering approach and regression methodology to predict type of cardiac arrhythmia.

Machine Learning algorithms are powerful tools, but should be used with caution. Loring et al. mention in their paper [10] the possible difficulties in application of methods of this kind (e.g. critical evaluation of methodology, errors in methodology difficult to detect, challenging clinical interpretation). We have planned our research taking this into account.

2   Materials and Methods

The data used in the study, in the form of RR intervals, comes from patients with implanted ICD's; the details are described below. The raw RR intervals were transformed into descriptive variables using several alternative methods. Then, the informative variables were identified with the help of several alternative feature selection methods. Finally, predictive models for arrhythmia events were constructed using Machine Learning algorithms (Fig 1). The details of this protocol are described in the following sections.
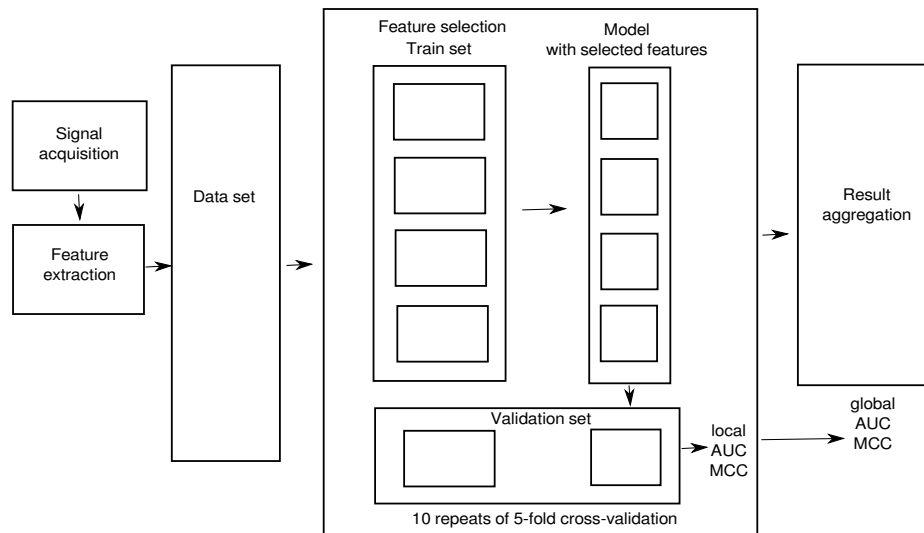
Fig. 1. Block diagram of data processing: data acquisition, feature extraction and modelling in cross-validation loop. AUC – area under receiver operator curve, MCC – Matthews correlation coefficient.

2.1   Data set

The input data consisted of 184 tachograms (signals of RR-intervals i.e. beat-to-beat intervals, observed in ECG) from 29 patients with single chamber ICD implanted in the years 1995—2000 due to previous myocardial infarction. Only data from patients with devices compatible with the PDM 2000 (Biotronik) and STDWIN (Medtronic) programs were analysed in the study. Patients who had a predominantly paced rhythm were excluded from the study. The VF zone was active in all patients with the lower threshold from 277 ms to 300 ms. The VT zone was switched on in all patients. Antitachycardia pacing (ATP) was the first therapy in the VT zone. Ventricular pacing rate was 40–60 beats/min (bpm).

Samples were recorded both during normal heartbeat (121 events) and onset of arrhythmia – ventricular fibrillation (VF – 12 events) or ventricular tachycardia (VT – 51 events). Both types of arrhythmia were considered as a single class. The length of these signals varied from 1000 to 9000 RR intervals. The signals have been collected from patients from The Cardinal Wyszyński Institute of Cardiology in Warsaw [11]. Patient characteristics are presented in Table 1.

Table 1. Patients clinical characteristics ($n = 29$), ACEI – Angiotensin Converting Enzyme Inhibitors, ARB – Angiotensin Converting Enzyme Inhibitor, CABG – Coronary Artery Bypass Grafting, PCI – Percutaneous Transluminal Intervention, SCD – Sudden Cardiac Death.

Age (years)		56.1 ± 9.8
Male gender (%)		26 (90)
Left ventricular ejection fraction %		32 ± 12.4
PCI n (%)		8 (27.6)
CABG n (%)		3 (10.3)
Indications for ICD implantation n (%)	Primary prophylaxis of SCD	9 (31)
	Secondary prophylaxis of SCD	20 (69)
Pharmacological treatment n (%)	Amiodarone	18 (62.5)
	Sotalol	2 (6.8)
	β-blockers	28 (96.5)
	ACEI and (or) ARB	29 (100)
	Statins	23 (82.4)
	Antiplatelet drugs	27 (93.1)
	Diuretics	10 (34.5)
	Aldosterone blockers	12 (41.3)
ICD manufacturer n (%)	Biotronik	13 (44.8)
	Medtronic	16 (54.2)

Typical signals of RR intervals from a patient with ICD during normal rhythm and during arrhythmia (VF) are shown in Fig. 2.
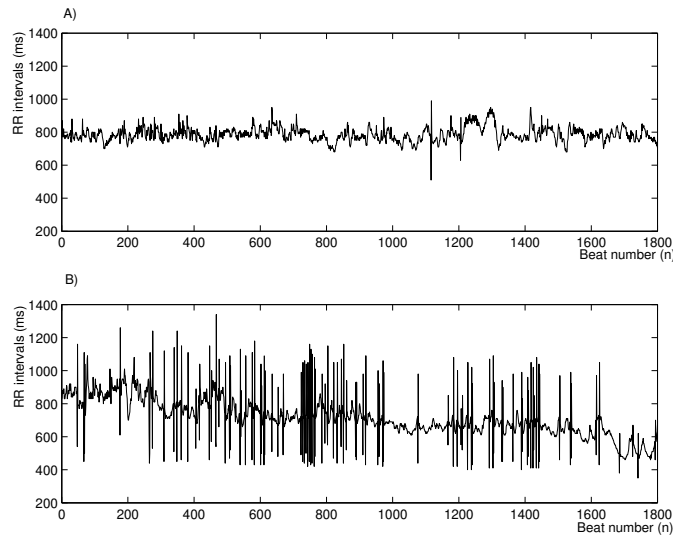
Fig. 2. RR intervals from patient with ICD: A) during normal rhythm, B) during arrhythmia (VF).

2.2   Data preprocessing

Data preprocessing was performed with the help of the RHRV package for analysis of heart rate variability of ECG records [12] implemented in R [13]. We followed the basic procedure proposed by the authors of this package. First, the heart beat positions were used to build an instantaneous heart rate series. Then, the basic filter was applied in order to eliminate spurious data points. Finally, the interpolated version of data series with equally spaced values was generated and used in frequency analysis. The default parameters were used for the analysis, with the exception of the width of the window for further analysis, as described later. For every signal we generated descriptors – performed basic analysis in time domain, frequency domain and also we calculated parameters related to selected nonlinear methods.

2.3   Descriptors

The preprocessed data series was then used to generate 47 descriptors using following approaches: statistical analysis in time domain, analysis in frequency (Fourier analysis) and time-frequency (wavelet analysis) domains, nonlinear analysis (Poincaré maps, the detrended fluctuation analysis, and the recurrence quantification analysis). The detailed description of the parameters is presented below.

Statistical parameters in time domain. Statistical parameters [12] calculated in time domain are:

– SDNN — standard deviation of the RR interval,
– SDANN — standard deviation of the average RR intervals calculated over short periods (50s),
– SDNNIDX -– mean of the standard deviation calculated over the windowed RR intervals,
– pNN50 — proportion of successive RR intervals greater than 50ms,
– SDSD — standard deviation of successive differences,
– r-MSSD — root mean square of successive differences,
– IRRR — length of the interval determined by the first and third quantile of the ΔRR time series,
– MADRR — median of the absolute values of the ΔRR time series,
– TINN — triangular interpolation of RR interval histogram,
– HRV index — St. George's index.

Parameters in frequency domain and time-frequency domain. In frequency domain and time-frequency domain we performed Fourier transform and wavelet transform, obtaining a power spectrum for frequency bands.

Spectral analysis is based on the application of Fourier transform in order to decompose signals into sinusoidal components with fixed frequencies [14]. The power spectrum yields the information about frequencies occurring in signals. In particular we used RHRV package and we applied STFT (short time Fourier transform) with Hamming window (in our computations with parameters size=50 and shift=5, which, after interpolation, gives 262–376 windows, depending on the signal).

Wavelet analysis allows to simultaneously analyse time and frequency contents of signals [15]. It is achieved by fixing a function called mother wavelet and decomposing the signal into shifted and scaled versions of this function. It allows to precisely distinguish local characteristics of signals. By computing wavelet power spectrum one can obtain the information about frequencies occurring in the signal as well as when these frequencies occur. In this study we used Daubechies wavelets.

We obtained mean values and standard deviations for power spectrum (using Fourier and wavelet transform) for 4 frequency bands: ULF — ultra low frequency component 0–0.003 Hz, VLF — very low frequency component 0.003–0.03 Hz, LF — low frequency component 0.03–0.15 Hz, HF — high frequency component 0.15–0.4 Hz. We have also computed mean values and standard deviations of LF/HF ratio, using Fourier and wavelet transform.

Parameters from nonlinear methods

Poincaré maps. We used standard parameters derived from Poincaré maps, They are return maps, in which each result of measurement is plotted as a function of

the previous one. A shape of the plot describes the evolution of the system and allows us to visualise the variability of time series (here RR-intervals). There are standard descriptors used in quantifying Poincaré plot geometry, namely SD1 and SD2 [16, 17], that are obtained by fitting an ellipse to the Poincaré map. We also computed $SD1/SD2$ ratio.

DFA method. Detrended Fluctuation Analysis (DFA) quantifies fractal-like autocorrelation properties of the signals [18, 19]. This method is a modified RMS (root mean square) for the random walk. Mean square distance of the signal from the local trend line is analysed as a function of scale parameter. There is usually a power-law dependence and an interesting parameter is the exponent. We obtained 2 parameters: short-range scaling exponent (fast parameter f.DFA) and long-range scaling exponent (slow parameter s.DFA) for time scales.

RQA method. We computed several parameters from Recurrence Quantification Analysis (RQA) which allow to quantify the number and duration of the recurrences in the phase space [20]. Parameters obtained by RQA method [12] are:

- REC – recurrence, percentage of recurrence points in a recurrence plot,
- DET – determinism, percentage of recurrence points that form diagonal lines,
- RATIO – ratio between DET and REC, the density of recurrence points in a recurrence plot,
- Lmax – length of the longest diagonal line,
- DIV – inverse of Lmax,
- Lmean – mean length of the diagonal lines; Lmean takes into account the main diagonal,
- LmeanWithoutMain – mean length of the diagonal lines; the main diagonal is not taken into account,
- ENTR – Shannon entropy of the diagonal line lengths distribution,
- TREND – trend of the number of recurrent points depending on the distance to the main diagonal,
- LAM – percentage of recurrent points that form vertical lines,
- Vmax – longest vertical line,
- Vmean – average length of the vertical lines.

2.4 Identification of informative variables

We have used several methods to identify the descriptors generated from the signal that are related to the occurrence of arrhythmia, namely the straightforward t-test, importance measure from the Random Forest [21], relevant variables returned by Boruta algorithm for all-relevant feature selection [22], as well as relevant variables returned by the MDFS (Multi-Dimensional Feature Selection) algorithm [23, 24]. Boruta is a wrapper on the Random Forest algorithm, whereas MDFS is a filter that relies on the multi-dimensional information entropy and therefore can take into account non-linear relationships and synergistic interactions between multiple descriptors and decision variable. We have applied MDFS

in one and two-dimensional mode, using default parameters. All computations were performed in R [13], using R packages.

2.5   Predictive models

Predictive models were built using Random Forest algorithm [21] and SVM (Support Vector Machine) [25]. The Random Forest model achieved better accuracy than the SVM model, which is consistent with the results presented by Fernández-Delgado et al. [26]. Hence, we focused on the Random Forest model exclusively, a method that can deal with complex, nonlinear relationships between descriptors and decision variable. It is routinely used as "out of the box" classifier in very diverse application areas. In a recent comprehensive test of 179 classification algorithms from 17 families, Random Forest was ranked as best algorithm overall [26]. It is an ensemble of decision tree classifiers, where each tree in the forest has been trained using a bootstrap sample of individuals from the data, and each split attribute in the tree is chosen from among a random subset of attributes. Classification of individuals is based upon aggregate voting over all trees in the forest. While there are numerous variants of Random Forest general scheme, we chose to use the classic algorithm proposed by Breiman implemented in the randomForest package in R [27]. Each tree in the Random Forest is built as follows:

- let the number of training objects be N, and the number of features in features vector be M,
- training set for each tree is built by choosing N times with replacement from all N available training objects,
- number $m << M$ is an amount of features on which to base the decision at that node. These features are randomly chosen for each node,
- each tree is built to the largest extent possible. There is no pruning.

Repetition of this algorithm yields a forest of trees, which all have been trained on bootstrap samples from training set. Thus, for a given tree, certain elements of training set will have been left out during training. The randomForest function was called with default parameters, with one modification – 1000 trees were used instead of 500.

Measuring quality of models and validation of modelling procedure
Three metrics were used to assess the quality of models: AUC (area under ROC curve) and MCC (Matthews Correlation Coefficient) [28] in addition to ordinary error level. Two former functions are more robust, in particular for imbalanced data sets.

It is well-known that variable selection can introduce significant over-fitting, especially when parameters selected within cross-validation are not highly informative [29]. To deal with the problem and to estimate the robustness of the models we applied the entire modelling was performed in five-fold cross-validation scheme. Then the procedure was repeated ten times and results are

averaged to remove dependence on the particular split of data set into folds. This protocol is very demanding computationally, since entire modelling procedure is performed 50 times. In particular also the most time-consuming part of protocol, namely identification of informative variables, is performed 50 times. Nevertheless, these computations are essential for robust estimate of performance of the machine learning models.

3   Results and Discussion

3.1   Feature selection

Feature selection was performed with the help of five algorithms using t-test, MDFS in 1 dimension (MDFS 1d) and 2 dimensions (MDFS 2d), Random Forest (RF) feature importance, and Boruta algorithm. Table 2 displays the number of times when variable was deemed relevant in fifty runs of each algorithm. The best results are presented according to the results obtained by Boruta.

Table 2. Number of occurrences of parameters in cross-validation loop for different feature selection methods.

Variable	Feature selection method				
	t-test	MDFS 1d	MDFS 2d	RF the best 10	Boruta
SD1/SD2	50	50	50	50	50
SD2	50	50	43	50	50
s.DFA	50	40	46	21	50
HRVi	47	49	49	49	45
mean.fULF	10	9	9	-	45
r-MSSD	50	49	34	49	31
SDNNIDX	19	29	16	44	31
IRRR	4	10	19	-	31
TINN	3	49	49	50	24
MADRR	2	11	13	46	23
SDSD	-	48	40	47	16
sd.fHF	41	17	-	-	13
sd.wHF	-	18	-	1	11
pNN50	50	47	34	-	9

For Random Forest feature importance one can see results for the best 10 features. The most frequently appearing parameters SD1/SD2 and SD2 are obtained from the Poincaré maps. The s.DFA arises from the Detrended Fluctuation Analysis. The HRVi, IRRR, r-MSSD, SDNNIDX, TINN, MADRR, SDSD and pNN50 variables are the statistical parameters in the time domain. The mean.fULF, sd.fHF and sd.wHF arise in the wavelet analysis. Interestingly, all methods agree on that variables arising from nonlinear analysis are most important. Then the relative importance of variables diverges among methods. Most methods agree that statistical variables in time domain are important, but there are significant differences between methods with respect to which of them are

most relevant. The largest disagreement concerns variables arising from spectral analysis, which are generally considered irrelevant by most methods, but some variables are considered very important by some methods.

3.2 Predictive models

First, we tested whether predicting arrhythmia is even possible. The results of five point summary (Minimum, Maximum, Median, 1st Quartile, 3rd Quartile) statistics on a set of observations are presented in Table 3.

Table 3. Testing of arrhythmia prediction's possibility (*true* labels versus *random* labels).

Labels	Min	1st Quartile	Median	Mean	3rd Quartile	Max
true	0.14	0.25	0.31	0.30	0.33	0.58
random	0.22	0.42	0.50	0.50	0.58	0.83

We focused on the Random Forest model. The evaluation of prediction was done by 5-fold cross validation. Tests were carried out in two ways. First we performed 1000 iterations with true labels (Table 3 row labelled *true*). The result was poor: error median and mean were about 0.3. Nevertheless, it shows that it is possible to perform prediction. Then, we did the same procedure, but with random labels. Before each iteration a new set of labels was randomised. The next step was to perform the prediction using Random Forest in 5-fold cross validation. The results are in Table 3 (row labelled *random*). Mean and median of prediction error were 0.5. The comparison of the results described in Table 3 shows that there is a significant difference in prediction based on real and random labels.

The prediction results in cross-validation loop for different feature selection algorithms measured by AUC and MCC are presented in Table 4. The best

Table 4. Results of prediction on different feature sets with selected parameters (mean value±standard deviation of the mean).

Parameters set	Error	MCC	AUC
t-test	0.264 ± 0.008	0.432 ± 0.022	0.813 ± 0.009
MDFS 1d	0.274 ± 0.009	0.422 ± 0.022	0.807 ± 0.009
MDFS 2d	0.289 ± 0.010	0.404 ± 0.025	0.796 ± 0.011
RF the best 10	0.253 ± 0.008	0.450 ± 0.023	0.823 ± 0.010
Boruta	0.287 ± 0.008	0.396 ± 0.019	0.796 ± 0.010

results were obtained for classifier that used 10 most relevant variables from the Random Forest. Results are stable and repeatable.

Each model was built using all variables that were deemed relevant by a feature selection algorithm in a given iteration of the cross-validation. Usually the number of relevant variables was close to 10 — depending on applied feature selection method.

The prediction results in cross-validation loop for different feature selection algorithms measured by AUC and MCC are presented in Fig. 3. One can observe outlier points in MCC results of RF the best 10 features. Results are stable and repeatable.
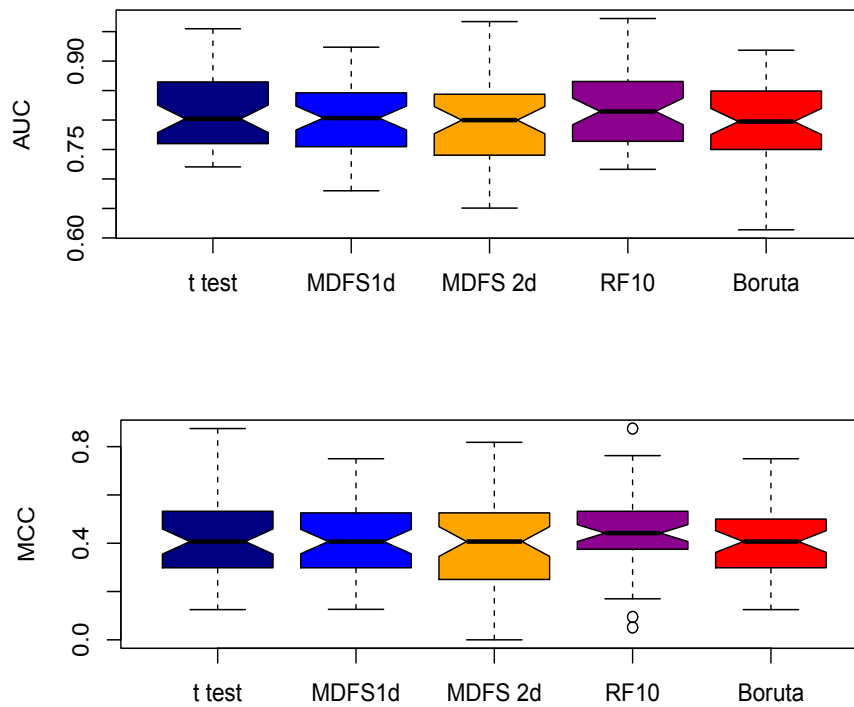
Fig. 3. Prediction results in cross-validation loop for different feature selection algorithms measured by AUC (top) and MCC (bottom).

In Fig. 4 we present AUC (area under ROC curve) for different feature selection algorithms.
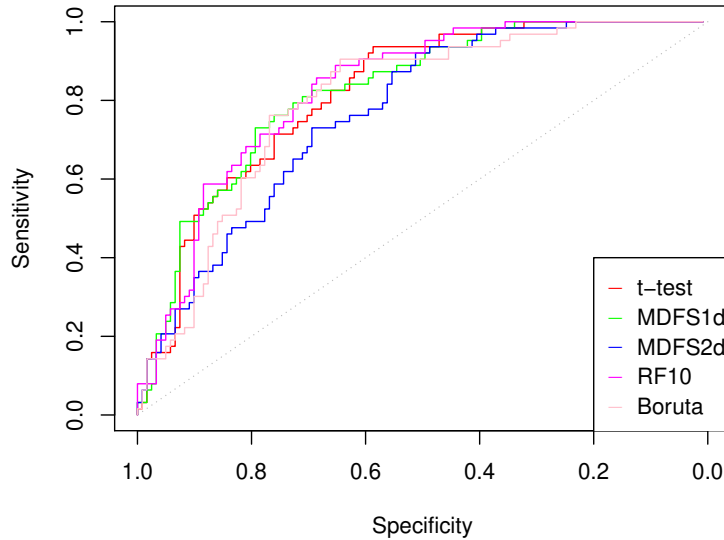
Fig. 4. AUC for different feature selection algorithms.

4   Conclusion

Based on obtained results we concluded that it's possible to find information about arrhythmia in RR intervals, but it's too weak to build useful medical predictive models using currently available methods. The subject requires further research to find algorithms better suited to the problem. In particular, a substantial increase of the size of the experimental sample, for instance by two or three orders of magnitude, should improve the quality of the models, as has been shown in numerous cases in applications of Machine Learning tools to different problems [30]. Additionally, it is likely that building individual models for each patient could yield better results.

Conflict of Interests: A. P. was a consultant for Biotronik and receives lectures fees from Medtronic, Biotronik and Abbott. He receives also a proctoring contract from Medtronic.

References

1. Luz, E.J.S, Schwartz, W.R., et al.: ECG-based heartbeat classification for arrhythmia detection: A survey. Computer Methods and Programs in Biomedicine. **127**, 144—164 (2016). https://doi.org/:10.1016/j.cmpb.2015.12.008

2. Wilkoff, B.L., Fauchier, L., et al.: 2015 HRS/EHRA/APHRS/SOLAECE expert consensus statement on optimal implantable cardioverter-defibrillator programming and testing. EP Europace **18**(2), 159–183 (2015). https://doi.org/:10.1093/europace/euv411

3. Al-Khatib, S.M., Stevenson, W.G., et al.: 2017 AHA/ACC/HRS Guideline for Management of Patients With Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death: Executive Summary. Circulation **138**(13), e210–e271 (2018). https://doi.org/:10.1161/CIR.0000000000000548

4. Shakibfar, S., Krause, O., et al.: Predicting electrical storms by remote monitoring of implantable cardioverter-defibrillator patients using machine learning. EP Europace. **21**, 268—274 (2019). https://doi.org/:10.1093/europace/euy257

5. Iranitalab, I.: Prediction of arrythmia through analysis of the ventricular electrogram. A thesis presented to The Faculty of the Department of Chemical and Materials Engineering. San Jose State University (2009)

6. Taye, G. T., Shim, E. B., Hwang, H.-J., et al.: Machine Learning Approach to Predict Ventricular Fibrillation Based on QRS Complex Shape. Frontiers in Physiology **10**, 1193 (2019)

7. Blužaitė, I., Rickli, H., et al.: Assessment of QT Dispersion in Prediction of Life-threatening Ventricular Arrythmias in Recipients of Implantable Cardioverter Defibrillator. Elektronika Ir Elektrotechnika **75**(3), 73-76 (2007)

8. Cho, J., Kim, Y., Lee, M.: Prediction to Atrial Fibrillation Using Deep Convolutional Neural Networks. In: Rekik, I., et al. (eds.) PRIME 2018, LNCS vol. 11121, pp.164–171. Springer Nature Switzerland AG (2018)

9. Cp, P., Suresh, A., Suresh, G.: Prediction of cardiac arrhythmia type using clustering and regression approach (P-CA-CRA). In: 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp.51-54. IEEE (2017)

10. Loring, Z., Mehrotra, S., Piccini, J.P.: Machine learning in 'big data': handle with care. EP Europace. **21**(9), 1284–1285 (2019). https://doi.org/:10.1093/europace/euz130

11. Przybylski, A., Baranowski, R., et al.: Verification of implantable cardioverter defibrillator (ICD) interventions by nonlinear analysis of heart rate variability – preliminary results. Eur. Eur. Pacing Arrhythm. Card. Electrophysiol. J. Work. Groups Card. Pacing Arrhythm. Card. Cell. Electrophysiol. Eur. Soc. Cardiol. **6**, 617—624 (2004). https://doi.org/:10.1016/j.eupc.2004.08.001

12. Martínez, C.A.G., Quintana, A.O., et al.: Heart Rate Variability Analysis with the R package RHRV, Springer International Publishing, 2017 www.springer.com/gp/book/9783319653549 (accessed September 6, 2019)

13. R Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2008)

14. Challis, R.E., Kitney, R.I.: Biomedical signal processing (in four parts). Part 2. The frequency transforms and their inter-relationships. Medical & Biological Engineering & Computing **29**, 1—17 (1991)

15. Mallat, S.: A Wavelet Tour of Signal Processing: The Sparse Way. Academic Press (2008)

16. Brennan, M., Palaniswami, M., Kamen, P., et al.: Do existing measures of Poincaré plot geometry reflect nonlinear features of heart rate variability? IEEE Transactions on Biomedical Engineering **48**, 1342—1347 (2001). https://doi.org/:10.1109/10.959330

17. Tulppo, M.P., Mäkikallio, T.H., et al.: Quantitative beat-to-beat analysis of heart rate dynamics during exercise. American Journal of Physiology **271**, H244–252 (1996). https://doi.org/:10.1152/ajpheart.1996.271.1.H244

18. Rodriguez, E., Echeverria, J.C., Alvarez-Ramirez, J.: Detrended fluctuation analysis of heart intrabeat dynamics. Physica A: Statistical Mechanics and its Applications **384**, 429–438 (2007). https://doi.org/:10.1016/j.physa.2007.05.022

19. Peng, C.K., Havlin, S., et al.: Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos **5**(1), 82–87 (1995). https://doi.org/:10.1063/1.166141

20. Zbilut, J.P., Thomasson, N., Webber, C.L.: Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals. Medical Engineering & Physics **24**, 53–60 (2002). https://doi.org/:10.1016/S1350-4533(01)00112-6

21. Breiman, L.: Random Forests. Machine Learning **45**, 5–32 (2001). https://doi.org/:10.1023/A:1010933404324

22. Kursa, M.B., Jankowski, A., Rudnicki, W.R.: Boruta – A System for Feature Selection. Fundamenta Informaticae **101**, 271–285 (2010)

23. Piliszek, R., Mnich, K., et al.: MDFS – MultiDimensional Feature Selection in R. R Journal **11**, 198–210 (2019)

24. Mnich, K., Rudnicki, W.R.: All-relevant feature selection using multidimensional filters with exhaustive search. Information Sciences (2020). In Press. https://doi.org/:10.1016/j.ins.2020.03.024

25. Cortes, C., Vapnik, V.: Support-Vector Networks. Machine Learning **20**, 273-–297 (1995). https://doi.org/:10.1023/A:1022627411411

26. Fernández-Delgado, M., Cernadas E., et al: Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? Journal of Machine Learning Research **15**, 3133–3181 (2014)

27. Liaw, A., Wiener, M.: Classification and Regression by randomForest. R News. **2**, 18–22 (2002)

28. Matthews, B.W.: Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA) – Protein Structure **405**, 442–451 (1975). https://doi.org/:10.1016/0005-2795(75)90109-9

29. Cawley, G.C., Talbot, N.L.C.: On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. Journal of Machine Learning Research **11**, 2079–2107 (2010)

30. Halevy, A., Norvig, P., Pereira, F.: The Unreasonable Effectiveness of Data. IEEE Intelligent Systems **24**, 8–12 (2009). https://doi.org/:10.1109/MIS.2009.36