# Detection of Atrial Fibrillation Using Decision Tree Ensemble

Guangyu Bin[1], Minggang Shao[1], Guanghong Bin[1], Jiao Huang[1], Dingchang Zheng[2], Shuicai Wu[1]

[1]College of Life Science and Bioengineering, Beijing University of Technology, Beijing, China
[2]Faculty of Medical Science, Anglia Ruskin University, Chelmsford, United Kingdom

## Abstract

*2017 PhysioNet/CinC Challenge proposed a global competition for classifying a short single ECG lead recording into normal sinus rhythm, atrial fibrillation (AF), alternative rhythm, and unclassified rhythm. This study developed and evaluated a pragmatic approach to solve the challenge, which was based on a decision tree ensemble with 30 features from ECG recording. The model was trained using the AdaBoost.M2 algorithm. The results reported here were obtained using 100-fold cross-validation, and the lowest MSE was 0.12 with the maximum number of splits of 55, and the number of trees of 20. The entry was tested and scored in the second phase of the challenge. The achieved scores for "Normal", "AF", "Other", were 0.93, 0.86, and 0.79, respectively, while the F1 measure was 0.86, and the official overall score was 0.82.*

## 1. Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia. The estimated age-standardized prevalence of AF in Chinese population is 0.65%, which is positively correlated with age [1]. AF is also associated with increased rates of death, stroke, heart failure etc. [2].

AF detection is therefore clinically important. AF detectors are normally designed from two aspects: 1) analysis of P waves by time domain method or frequency domain method [3, 4], 2) analysis of RR interval [5]. Recently, multivariate approaches based on machine learning have achieved excellent performance [6]. In practice, as in present 2017 PhysioNet/CinC challenge, many factors made accurate AF detection very challenging, including: short ECG recording between 9 s and 60 s; High level of noise; Different types of abnormalities. In particular, many non-AF rhythms exhibit irregular RR intervals that are similar to AF features [7].

Besides, because there are a very larger number of records in the dataset of the challenge, multivariate approach based on machine learning has advantages of processing this large database. Therefore, this study aimed to develop and evaluate a pragmatic approach based on decision tree ensemble for AF detection.

## 2. Method

### 2.1. Feature extraction

Feature selection is a fundamental process in machine learning. To achieve this, a point in the QRS complex was detected (QRS point), using an improved Hamilton and Tompkins algorithm [8]. Thirty features were then derived from an ECG recording.

Table 1. Lists the features used in this study.

| Group Label | | Feature | ID |
|---|---|---|---|
| AF Feature | ● | AFEv | F10 |
| | ● | Shannon Entropy | F11 |
| | ● | Radius | F12 |
| | ● | K-S test value | F13 |
| Morphology Feature(A,B) | ● | QRS Duration (offset −onset) | F20A,F20B |
| | ● | PR interval | F21A,F21B |
| | ● | QT interval | F22A,F22B |
| | ● | QS interval | F23A,F23B |
| | ● | ST amplitude | F24A,F24B |
| | ● | P amplitude and flag | F25A,F25B |
| | ● | Q Amplitude | F26A,F26B |
| | ● | R amplitude | F27A,F27B |
| | ● | S amplitude | F28A,F28B |
| | ● | T amplitude | F29A,F29B |
| RR intervals | ● | Median RR interval | F30 |
| | ● | Index for arrhythmia | F31 |
| Similarity index between beats | ● | Similarity index of QRS | F40 |
| | ● | Similarity index of R amplitude | F41 |
| | ● | Ratio of high similarity beats | F42 |
| | ● | Signal Qualify index | F43 |

*1) AF features:*
Four AF features were extracted based on the RR intervals. They were AFEv(F10), Shannon Entropy (F11), Radius (F12), and Kolmogorov-Smirnov (K-S) test value (F13).

AFEv is an excellent AF detector, which uses the

Lorenz distribution of a time series of RR intervals [9]. Shannon Entropy measures the irregularity of heartbeats, which is also widely used in AF detection area [10]. K-S test value has shown exceptional accuracy for classifying AF and sinus rhythm [11]. Moreover, a simple feature, Radius was proposed here, which was consider to be effective for short ECG recordings, Specifically, the Radius was defined as the radius of the smallest circle which takes the 60% area in the normalized Lorenz plot of dRR intervals.

*2) Morphology features:*

The absence of P waves is an important feature for AF detection. Long PR interval, width QRS, ST value are also related to abnormal beats. Because of the high noise level, two improved methods (ECGPUWAVE Osealib) were applied separately for extracting ten morphology features in this study (F20A-F29A for ECGPUWAVE, F20B-F29B for Osealib), as listed in Table 1. The ECGPUWAVE, an open source QRS detector and waveform limit locator [12], has shown excellent performance for P wave and QRS detection. Osealib is another open source tool for detecting beats and classifying arrhythmias [13].

*3) RR interval features:*

It is known that average heart rate during a recording could provide meaningful information for Tachycardia and bradycardia. In the present study, Median RR interval (F30) over whole recording was used.

Since large proportion of "Other" records have arrhythmia, an index of arrhythmia (F31) from Tsipouras's method was derived, which was based on knowledge-based rules derived from three continues RR intervals [14].Those rules included:

Rule 1: $1.2 * RR2 < RR1$ and $1.3 * RR2 < RR3$
Rule 2: $|RR1 - RR2| < 0.3 * MRR$
$\quad and (RR1 < 0.8 * MRR \ or \ RR2 < 0.8 * MRR)$
$\qquad and \ RR3 > 0.6 * (RR1 + RR2)$
Rule 3: $|RR3 - RR2| < 0.3 * MRR$
$\quad and (RR2 < 0.8 * MRR \ or \ RR3 < 0.8 * MRR)$
$\qquad and \ RR1 > 0.6 * (RR2 + RR3)$
Rule 4: $RR2 > 1.5 * MRR$ and $1.5 * RR2 < 3 * MRR$

where $RR1$, $RR2$ and $RR3$ represent three continues where RR intervals, and $MRR$ were from the average RR interval of five nearest beats.

*4) Similarity index between beats:*

The similarity of ECG beats in a recording was computed from the correlation of QRS wave and R amplitude, from which three features (F40-F42), were obtained. Furthermore, the fluctuation of isoelectric level (from the T wave offset of previous beat to P wave onset of current beat) was quantified to evaluate the signal quality of each beat, and a statistic value, Signal Quality index (F43) was also obtained to evaluate the quality of a recording .

## 2.2. Classification

Decision trees based classifier models were utilized throughout this study. Decision trees are one of the most widely used classification models due to their interpretability and the availability of efficient and scalable learning algorithms [15]. The function "fitctree" in MATLAB was used for fitting binary classification decision tree. The depth of the trees can be controlled by the parameters of "maximal number of decision splits" and "minimum number of leaf node observations". Cross-validation was used for optimizing the depth of the trees. Furthermore, a decision tree ensemble was trained using the adaBoost.M2 algorithm [16]. The function "fitensemble" in MATLAB was used for fitting a decision tree ensemble. The number of trees, the key parameters, was determined from 100 fold cross-validation.

## 3. Results

## 3.1 Scores with a binary classification

An ECG expert was asked to re-label some of the recordings, and 207 labels were updated. After re-labeling, there were 8528 recordings in the training dataset, 284 recordings for "Noise", 4949 recordings for "Normal", 736 recordings for "AF", and 2059 recordings for "Other".

A binary classification decision tree model was trained with the 8528×30 data. Fig.1 represents the trend of cross validation errors as a function of the maximal number of decision splits. A lowest error 0.13 was achieved with the maximal number of decision splits of 55. We submitted an entry using the obtained model, and the scores were 0.91, 0.83, and 0.76 for "Normal", "AF", "Other", respectively. The final F1 measure was 0.84.
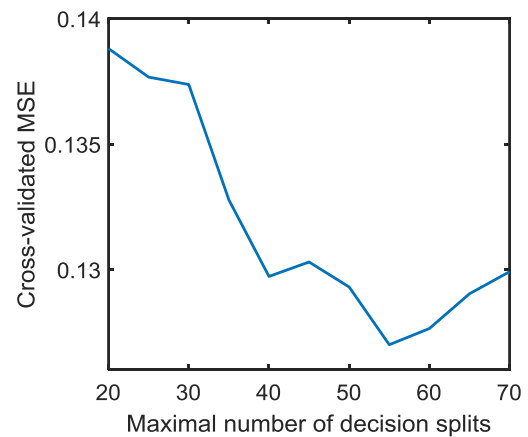


Figure 1. Cross validation classification errors against the maximal number of decision splits.

## 3.2 Identification of key features

This study employed thirty features for AF detection. . In order to identify which features contributed more to our classifier, an ensemble of bagged decision trees was applied to investigate the importance of individual features.

As shows in Fig2, Index for arrhythmia (F31), Median RR interval (F30) are the most important ones. AFEv (F10) contributes the most for AF classification, followed by K-S test value (F13) and Radius (F12). The four features (F42, F43, F40, F41) of Similarity index of beats can a may have contribution to noise classification. QRS Duration (F20B) which is imported for diagnosis abnormal beats also be selected. P wave amplitude (F25A, F25B) and PR interval (F21B) is imported for diagnosis the disease such as AF, AV block etc. Those selected features are consistent with the doctor experience.
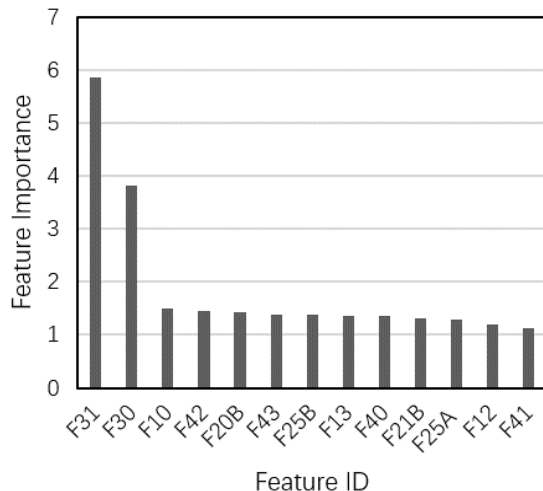


Figure 2. Illustration of the importance of different features. Only the most important 13 features are shown.

## 3.3 Scores with decision tree ensemble

A decision tree ensemble was trained using the adaBoost.M2 algorithm. Fig3 represents the cross validation classification errors as a function of the number of decision trees. A lowest error 0.12 was achieved with number of decision trees of 20. We submitted an entry using the obtained model, and the scores for "Normal", "AF", "Other", were 0.93, 0.86, and 0.79, respectively. The final F1 measure was 0.86.

## 4. Discussion

This study proposed machine learning based multivariate approach for AF detection. Serval carefully selected features were extracted from short ECG
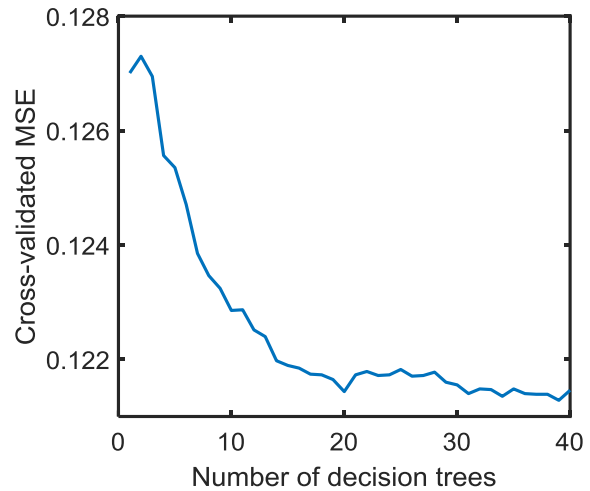


Figure 3. Cross validation classification errors with different number of decision trees.

recording, and a decision tree ensemble was trained to classify the recordings. The present approach showed excellent AF detection performance with the achieved scores of 0.93, 0.86, and 0.79 for "Normal", "AF", "Other", respectively, and the final F1 measure of 0.86.

In order to further improve the AF detection performance, the following aspects could be considered in the future study:

1) Improve the detection accuracy of P wave;
2) More accurate labels may help improve the accuracy;
3) Deep learning may be a good selection.

Nevertheless, our study has demonstrated that the algorithm using a decision tree ensemble for AF detection could achieve acceptable detection accuracy, providing evidence for its clinical application.

## Acknowledgements

## References

[1] Zhou Z, Hu D. An epidemiological study on the prevalence of atrial fibrillation in the Chinese population of mainland China. Journal of epidemiology, 2008, 18(5): 209-216.

[2] Camm A J, Kirchhof P, Lip G Y H, et al. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). European heart journal, 2010, 31(19): 2369-429.

[3] Dilaveris P E, Gialafos E J, Sideris S K, et al. Simple electrocardiographic markers for the prediction of

paroxysmal idiopathic atrial fibrillation. American heart journal, 1998, 135(5): 733-738.

[4] Aytemir K, ÖZER N, Atalar E, et al. P Wave Dispersion on 12-Lead Electrocardiography in Patients with Paroxysmal Atrial Fibrillation. Pacing and Clinical Electrophysiology, 2000, 23(7): 1109-1112.

[5] Tateno K, Glass L. Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and ΔRR intervals. Medical and Biological Engineering and Computing, 2001, 39(6): 664-671.

[6] Colloca R, Johnson A, Mainardi L, et al. A Support Vector Machine approach for reliable detection of atrial fibrillation events. Computing in Cardiology Conference (CinC), 2013. IEEE, 2013: 1047-1050.

[7] Clifford G, Liu C, Moody B, Silva I, et al. AF Classification from a Short Single Lead ECG Recording: the PhysioNet Computing in Cardiology Challenge 2017. Computing in Cardiology, Vol 44, 2017 (In Press).

[8] Hamilton P S, Tompkins W J. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Transactions on Biomedical Engineering, 1986 (12): 1157-1165.

[9] Sarkar S, Ritscher D, Mehra R. A detector for a chronic implantable atrial tachyarrhythmia monitor. IEEE Transactions on Biomedical Engineering, 2008, 55(3): 1219-1224.

[10] Dash S, Chon K H, Lu S, et al. Automatic real time detection of atrial fibrillation. Annals of Biomedical Engineering, 2009, 37(9): 1701-1709.

[11] Huang C, Ye S, Chen H, et al. A novel method for detection of the transition between atrial fibrillation and sinus rhythm. IEEE Transactions on Biomedical Engineering, 2011, 58(4): 1113-1119.

[12] QRS detection and waveform boundary recognition using ecgpuwave. https://physionet.org/physiotools/ecgpuwave/

[13] Hamilton PS, Tompkins WJ, Open Source ECG Analysis Software Documentation. http://www.eplimited.com

[14] Tsipouras M, Fotiadis D, Sideris D. An arrhythmia classification system based on the RR-interval signal. Artificial intelligence in medicine, 2005, 33(3): 237-250.

[15] Breiman L, Friedman J, Olshen R, et al. Classification and Regression Trees. Boca Raton, FL: CRC Press, 1984

[16] Freund Y, and Schapire R. A Decision Theoretic Generalization of On-Line Learning and an Application to Boosting. J. of Computer and System Sciences, Vol. 55, pp. 119–139, 1997.

Address for correspondence:

Guangyu Bin
Room B210, College of Life Science and Bioengineering, Beijing University of Technology, No.100 Pingleyuan, Chaoyang District, Beijing 100124 P. R. China.
guangyubin@bjut.edu.cn