Empirical Survey on Occult Atrial Fibrillation Prediction During Sinus Rhythm Through Heart Rate Variability and Premature Atrial Contractions

Daniele Padovano[1], Arturo Martinez-Rodrigo[1], Jose M Pastor[1], Jose J Rieta[2], Raul Alcaraz[1]

[1] Research Group in Electronic, Biomedical and Telecomm. Eng., Univ. of Castilla-La Mancha, Spain
[2] BioMIT.org, Electronic Engineering Department, Universitat Politecnica de Valencia, Spain

Abstract

The prevalence of atrial fibrillation (AF) has significantly increased over the past five decades, establishing it as the leading sustained heart arrhythmia, globally. An elevated risk of cardiovascular disease was observed in patients with AF, particularly in cases of silent paroxysmal AF (PAF). Previous research had relied on ECG monitoring, heart rate variability (HRV), and premature atrial contractions (PAC) assessment to predict the risk of PAF. This study was conducted to replicate and evaluate these methods using the publicly available Icentia-11k Dataset. The data preprocessing was performed following a typical machine learning (ML) pipeline, which included noise filtering and baseline wander removal, prior to feature extraction from HRV in time, frequency, and complexity domains. Various ML classifiers were trained and tested with different sets of features, revealing that the information provided by premature beats, such as PACs, notably influenced specificity at the expense of sensitivity.

1. Introduction

Paroxysmal Atrial Fibrillation (PAF) poses a significant challenge due to its sporadic nature and potential for adverse cardiovascular events. Predicting PAF risk before symptomatic episodes can lead to early interventions, reducing healthcare costs and improving the patient's quality of life [1]. In recent years, there has been considerable focus on employing heart rate variability (HRV) and premature atrial contractions (PACs) to predict AF onset through advanced machine learning (ML) techniques [2].

Early studies, such as those by Lynn et al. [3] and Zong et al. [4], demonstrated the feasibility of using HRV metrics and PAC counts to estimate AF risk through innovative ML techniques. These pioneering works laid the foundation for more advanced ML models by demonstrating the importance of detecting subtle changes in cardiac rhythm that precede AF episodes, thereby influencing subsequent research and development in the field.

Recent advancements have significantly improved prediction accuracy by incorporating modern ML algorithms. For instance, Ebrahimzadeh et al. employed a combination of feature vectors and a mixture of expert classifiers on HRV signals, achieving notable improvements in predictive performance [5]. Similarly, Boon et al. [6] and Narin et al.[7] utilized genetic algorithms and short-term HRV analysis, respectively, to refine predictive models, showing promising results in various clinical settings.

Additionally, the work of Castro et al. [8] and Parsi et al. [9] highlighted the potential of new HRV features and robust feature selection methodologies to enhance AF prediction accuracy. These studies integrated multiple physiological parameters, achieving robust performance with the same dataset. Despite these advancements, achieving consistently high prediction rates remains challenging, particularly in heterogeneous patient populations and real-world conditions. Future research directions may involve refining feature selection methods, integrating personalized patient data, and exploring hybrid models that combine ML with other physiological models. For instance, Bashar et al. introduced a novel density Poincaré plot-based ML method to detect AF from PACs, highlighting the potential of premature beats in this field [10].

Thus, the aim of this study is to replicate and evaluate some of the most relevant methods for predicting PAF using a meticulously curated subset of a publicly available dataset recorded in a real-life scenario. This empirical survey examines the application of various ML classifiers and HRV features in predicting AF, comparing their effectiveness and discussing the evolving landscape of predictive modeling techniques and their clinical implications.

2. Methodology

The pipeline for this study is summarized in Figure 1. Initially, the properties of the database are described, followed by the preprocessing of ECG signals. Features are then extracted from the HRV across different domains. Subsequently, various machine learning models are applied. Finally, the results are presented and discussed.

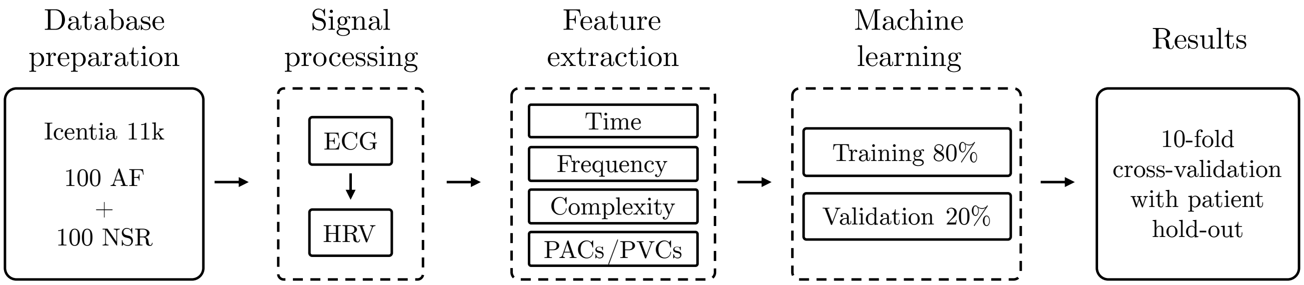

Figure 1: Methodology diagram overview.

2.1. Databases

The present study uses a controlled subset of the Icentia11k database [11], available on PhysioNet. This subset includes 30-minute ECG recordings from 200 patients (100 with occult AF and 100 with no cardiac issues). Annotations detail beat types and rhythm changes. The average patient age is 62.2 ± 17.4 years, with data primarily from Ontario hospitals (2017-2018). Technical specifications include a 250Hz sampling frequency, 16-bit resolution, and signal magnitude in mV.

2.2. Signal processing

The ECG recordings were processed using a second-order Chebyshev filter with cut-off frequencies ranging from 0.5 to 100 Hz, ensuring the preservation of the original signal's characteristics [12]. Following this, R-peak detection was executed utilizing the Pan-Tompkins algorithm [13], and a manual R-peak correction was applied to address potential delays. Finally, the HRV was calculated as the successive differences between R-peaks.

2.3. Feature extraction

The HRV features were extracted according to the established guidelines provided by the Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology [14]. Time domain features comprised various statistical measures of RR intervals, such as the mean, median, standard deviation, and other descriptive statistics detailed in Table 1. Frequency domain features were derived from power spectral density bands, including very low frequency, low frequency, and high frequency components, as outlined in Table 2. Complexity domain features incorporated measures of entropy and metrics from recurrence quantification analysis, including sample entropy, fuzzy entropy, and determinism, as presented in Table 3. In addition to these, PACs , premature ventricular contractions (PVCs), and atrioventricular block (AVB) episodes were identified and quantified, with detailed counts provided in Table 4.

Table 1: Time domain features.

Feature	Description
MAX	Maximum of RRi
MIN	Minimum of RRi
MEAN	Mean of RRi
MED	Median of RRi
SDNN	Standard deviation of normal RRi
SDSD	Standard deviation of the differences between adjacent RRi
RMSSD	Root mean square of differences between adjacent RRi
NN50	Pairs of adjacent NNi differing by more than 50 ms
pNN50	Determined by dividing NN50 by the total number of all NNi
IQR	Interquartile range

Table 2: Frequency domain features.

Feature	Description
VLF	FFT very low frequency component
LF	FFT low frequency component
HF	FFT high frequency component
LS-VLF	LSP very low frequency component
LS-LF	LSP low frequency component
LS-HF	LSP high frequency component

2.4. Machine Learning Classifiers

Various machine learning models were used to analyze HRV features, including Support Vector Machine (SVM), Linear Discriminant (LD), Decision Tree (DT), K-Nearest Neighbors (KNN), Adaptive Boosting (ADA), Gentle Boosting (GEN), and Bagging (BAG). The SVM was configured with a Gaussian kernel, scale of 6, and box constraint of 1, while LD used a gamma of 0. The DT model employed Gini's index with up to 4 splits, KNN was

Table 3: Complexity domain features.

Feature	Description
SampEn	Sample entropy
FuzzEn	Fuzzy Entropy
DistEn	Distribution Entropy
AttnEn	Attention Entropy
REC	RP Recurrence rate
DET	RP Determinism
L	RP average diagonal line length
LMAX	RP maximal diagonal line length
LAM	RP Laminarity

Table 4: Identified and counted episodes.

Feature	Description
PACs	Number of premature atrial contractions
PVCs	Number of premature ventricular contractions
AVBs	Number of atrioventricular block episodes

set with Euclidean distance, 10 neighbors, and squared inverse distance weighting. ADA and GEN were both configured with 30 learning cycles and a maximum of 8 splits in the template tree, and BAG utilized an ensemble of multiple decision trees. The above configurations where chosen based on state-of-the-art guidelines, and the models were assessed based on accuracy (Ac), sensitivity (Se), and specificity (Sp), with training (80:20 validation ratio) and testing carried out using 10-fold cross-validation with patient hold-out to avoid overlap between groups.

3. Results

The SVM model with all features demonstrated the highest overall performance, achieving an Ac of 69%, a Se of 66%, and a Sp of 72% (Table 5). In the time domain, the LD model ranked second with an Ac of 68% (Table 6). For the frequency domain, the DT model has shown high Se but low Sp (Table 7). In the complexity domain, the SVM model again achieved the highest performance with an Ac of 64% (Table 8).

4. Discussion

SVM consistently outperformed other classifiers across all feature sets, including time, frequency, and complexity domains. This superior performance is attributed to SVM's robust ability to handle high-dimensional data and its effective generalization capability across diverse features. Also, SVM achieved the highest overall Ac of 69.00% with all features combined, indicating its effectiveness in synthesizing information from multiple domains.

Table 5: Performance with All Features

Model	Ac	Se	Sp
ADA	65.50	67.00	64.00
BAG	65.00	69.00	61.00
DT	60.00	67.00	53.00
GEN	65.00	61.00	69.00
KNN	62.00	61.00	63.00
LD	65.50	61.00	70.00
SVM	69.00	66.00	72.00

Table 6: Performance with Time Domain Features

Model	Ac	Se	Sp
ADA	63.50	63.00	64.00
BAG	62.50	66.00	59.00
DT	57.00	62.00	52.00
GEN	60.50	63.00	58.00
KNN	59.00	52.00	66.00
LD	68.00	67.00	69.00
SVM	63.50	52.00	75.00

Table 7: Performance with Frequency Domain Features

Model	Ac	Se	Sp
ADA	52.50	55.00	50.00
BAG	59.00	71.00	47.00
DT	60.50	80.00	41.00
GEN	52.50	57.00	48.00
KNN	57.50	64.00	51.00
LD	63.50	71.00	56.00
SVM	62.00	72.00	52.00

Table 8: Performance with Complexity Domain Features

Model	Ac	Se	Sp
ADA	61.00	62.00	60.00
BAG	60.50	61.00	60.00
DT	51.00	52.00	50.00
GEN	51.50	50.00	53.00
KNN	59.50	58.00	61.00
LD	62.50	59.00	66.00
SVM	64.00	62.00	66.00

Table 9: Performance with Premature Beats Features

Model	Ac	Se	Sp
ADA	50.50	71.00	30.00
BAG	54.00	81.00	27.00
DT	49.50	75.00	24.00
GEN	54.50	82.00	27.00
KNN	48.50	9.00	88.00
LD	53.50	77.00	30.00
SVM	50.50	76.00	25.00

In contrast, DT showed strong Se of 80.00% in the frequency domain but exhibited lower specificity (Sp). This suggests that while DT is proficient at identifying true positive AF risk segments within the frequency domain, it struggles with minimizing false positives, which impacts overall predictive Ac. The DT's performance indicates that frequency domain features may be highly indicative of AF but require careful management to avoid false alarms.

Ensemble methods like GEN and BAG were notably effective when using only premature beats, achieving Se values greater than 80.00%, though at the cost of lower Sp. This performance demonstrate the strengths of ensemble methods in identifying patterns of premature beats to predict high AF risk segments accurately. However, the trade-off in Sp suggests a tendency towards higher false positive rates, necessitating a balance between Se and Sp.

Ultimately, variations in model effectiveness by feature type suggest a nuanced approach to feature extraction and model selection is necessary. Future work should refine feature extraction, especially for PACs, and optimize ensemble methods for better sensitivity and specificity.

5. Conclusions

In the context of machine learning algorithms for AF prediction based on heart rate variability features, support vector machine generally provided the best performance across all features, while premature beats improved sensitivity. Balancing sensitivity and specificity is crucial for effective AF prediction.

Acknowledgments

This research has received financial support from public grants PID2021-128525OB-I00, PID2021-123804OB-I00, and TED2021-130935B-I00 funded by MICIU/AEI/ 10.13039/501100011033, FEDER EU, and NextGenerationEU /PRTR., jointly with the European Regional Development Fund (EU), SBPLY/21/180501/000186, and SBPLY/21/180225/000080 from Junta de Comunidades de Castilla-La Mancha, and AICO/2021/286 from Generalitat Valenciana. Moreover, Daniele Padovano holds a predoctoral scholarship 2022-PRED-20642, which is co-financed by the operating program of European Social Fund (ESF) 2014-2020 of Castilla-La Mancha.

References

[1] Hindricks, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). European Heart Journal February 2021; 42(5):373–498.

[2] Pipilas D, Friedman SF, Khurshid S. The Use of Artificial Intelligence to Predict the Development of Atrial Fibrillation. Current Cardiology Reports May 2023;25(5):381–389.

[3] Lynn, et al. A two-stage solution algorithm for paroxysmal atrial fibrillation prediction. IEEE Transactions on Biomedical Engineering 2001;48:1023–1029.

[4] Zong, et al. Prediction of paroxysmal atrial fibrillation by analysis of atrial premature complexes. IEEE Transactions on Biomedical Engineering 2001;48:136–142.

[5] Ebrahimzadeh, et al. Prediction of paroxysmal atrial fibrillation: A machine learning based approach using combined feature vector and mixture of expert classification on hrv signal. Biomedical Signal Processing and Control 2018; 45:291–301.

[6] Boon, et al. Paroxysmal atrial fibrillation prediction method with shorter hrv sequences. Computers in Biology and Medicine 2016;75:1–8.

[7] Narin, et al. Early prediction of paroxysmal atrial fibrillation based on short-term heart rate variability. Journal of Electrocardiology 2018;51:483–489.

[8] Castro, et al. Methodology for the prediction of paroxysmal atrial fibrillation based on heart rate variability feature analysis. IEEE Transactions on Biomedical Engineering 2021; 68:567–576.

[9] Parsi, et al. Prediction of paroxysmal atrial fibrillation using new heart rate variability features. Computers in Biology and Medicine 2021;134:104456.

[10] Bashar, et al. Novel density poincaré plot based machine learning method to detect atrial fibrillation from premature atrial/ventricular contractions. IEEE Transactions on Biomedical Engineering 2021;68:123–134.

[11] Tan S, Ortiz-Gagné S, Beaudoin-Gagnon N, Fecteau P, Courville A, Bengio Y, Cohen JP. Icentia11k Single Lead Continuous Raw Electrocardiogram Dataset.

[12] Kaur M, Singh B, Seema. Comparison of different approaches for removal of baseline wander from ECG signal. In Proceedings of the International Conference & Workshop on Emerging Trends in Technology, ICWET '11. New York, NY, USA. ISBN 978-1-4503-0449-8, February 2011; 1290–1294.

[13] Pan J, Tompkins WJ. A Real-Time QRS Detection Algorithm. IEEE Transactions on Biomedical Engineering March 1985;BME-32(3):230–236.

[14] Task Force of the European Society of Cardiology the North American Society of Pacing. Heart Rate Variability. Circulation March 1996;93(5):1043–1065.

Address for correspondence:

Daniele Padovano
Campus Universitario S/N. 16071 – Cuenca (Spain)
daniele.padovano@uclm.es