Enhanced Classification and Prediction of Atrial Fibrillation Episodes Using ECG Signals: A Regression-Based Approach

Hodaya Rabinovich supervised by Daniel Langa

Abstract

Atrial fibrillation (AF) is a common cardiac arrhythmia associated with significant morbidity and mortality worldwide. This study investigates the classification and prediction of AF episodes using electrocardiogram (ECG) signals. We propose a novel approach that combines regression-based prediction of AF termination time with classification of AF episodes into terminating and non-terminating categories. Our model achieves promising results, with an accuracy of 88.3% and precision, recall, and specificity values of 76.7%, 93.3%, and 91%, To access the code and resources used in this research project, visit the GitHub repository

Introduction

Atrial fibrillation (AF) is a common heart rhythm disorder that affects millions of people worldwide. It poses a significant health burden and is associated with various adverse outcomes, including stroke and heart failure. According to the Centers for Disease Control and Prevention (CDC) [1], projections indicate that by 2030, approximately 12.1 million individuals in the United States will experience AF. In 2019 alone, AF was cited on 183,321 death certificates, with 26,535 of these deaths attributed to AF as the underlying cause.

AF is classified into terminating AF, also known as paroxysmal AF, and non-terminating , also known as sustained AF [2]. Each sub-type necessitates tailored treatment strategies to effectively manage the condition and minimised the chances of heart attack and sudden cardiac death (SCD) [3].

Given that the Electrocardiogram (ECG) test is the primary diagnostic procedure for detecting AF, this study aims to classify AF types using ECG data. This study builds on previous research [4] that detected AF using the chirplet transform and a deep convolutional bidirectional long short-term memory network with ECG signals

This study aims to introduce a new approach to AF classification by framing it as a regression problem, diverging from traditional classification methodologies. Rather than categorizing AF into discrete types, we seek to predict the duration until AF termination

ECG database

The dataset utilized for this research is the AF Termination Challenge Database available on Physio-Net [2]. Consisting of 30 one-minute recordings excerpted from longer ECG recordings, each instance contains two simultaneously recorded ECG signals. All recordings exhibit atrial fibrillation as the cardiac rhythm. The dataset is categorized into three groups: the first represents non-terminating AF, the second denotes AF terminating one minute after the end of the record, and the last signifies AF terminating immediately after the end of the record. The database is categorized into three groups. However, for the purpose of this study, all terminating AF signals were combined into a single class, irrespective of their duration to termination, to simplify the classification task.

Method

To address the classification problem of atrial fibrillation (AF), our method leverages additional information extracted from the ECG records. Given that each record spans one minute, assumptions regarding signal similarity between the beginning and end of the record may not hold true. Furthermore, the end of the signal in cases where AF terminates within one minute may resemble the beginning of records where AF immediately terminates after the recording ends. Consequently, traditional classification approaches may yield sub-optimal results. Therefore, our methodology aims for greater accuracy by addressing the challenge through a complete regression approach, predicting the exact time until AF termination for each segment of the record. the block diagram of the approach for classification of AF is shown in 1.

Signal Preprocessing and Labeling The ECG signals were preprocessed by filtering to remove noise and artifacts 4. Subsequently, each recording was segmented into non-overlapping 4-second frames. For records where AF terminated, the exact time until termination was labeled. For records where AF did not terminate within the recording duration, a time duration greater than 120 seconds (the longest known termination time) was assigned. An illustration comparing filtered and unfiltered signals is provided 2.

Regression Network Architecture The regression network employed in this study utilized architecture, which was previously applied in ECG automatic classification tasks [5] [6]. The model architecture is characterized by multiple residual blocks with skip connections, enabling effective

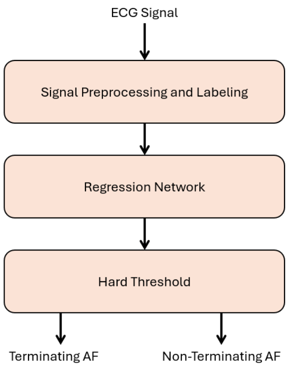

Figure 1: Block diagram showing the regression approach to classify Terminating AF vs. Non Terminating AF

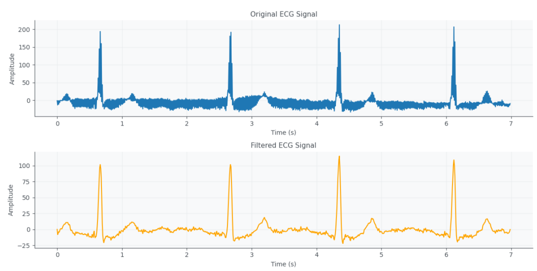

Figure 2: example of ECG signals before and after filtering [2]

learning of complex features from ECG signals 3. In this study, slight modifications were made to adapt the architecture for regression modeling. Specifically, the sigmoid activation function was removed, and an additional fully connected layer was added to the output, reducing the dimension from 32 to 1. figure

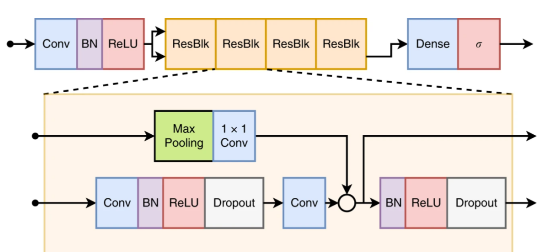

Figure 3: Regression Model as used in [5]. In this study the sigmoid activation function was removed, and an additional fully connected layer was added to the output

Hard Threshold for Classification Following regression, a hard threshold was applied to convert the predicted time durations into binary classifications. Instances where the predicted time exceeded a predefined threshold (e.g., 120 seconds) were classified as non-terminating AF, while instances below the threshold were classified as terminating AF

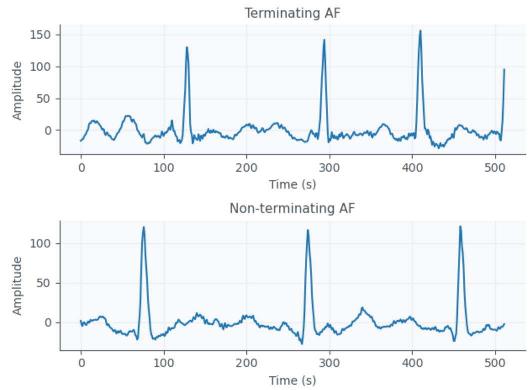

Figure 4: examples of ECG signals for terminating AF and non-terminating from the AF database [2]

Empirical Evaluation

In this section, we present the results of our study on the binary classification AF episodes using ECG signals (Terminating AF vs. Non-Terminating AF).

Training

The regression model was trained using the Adam optimizer with a learning rate of 0.001 and a reduce LR on Plateau scheduler. A batch size of 512 was utilized, with 20% of the dataset reserved for testing and 10% for validation. The training process spanned 600 epochs.

Figure 5 illustrates the training progress, showing the training and validation loss over epochs, as well as the learning rate adjustments made by the scheduler. Our model achieved convergence with decreasing training and validation losses over the epochs, demonstrating effective learning.

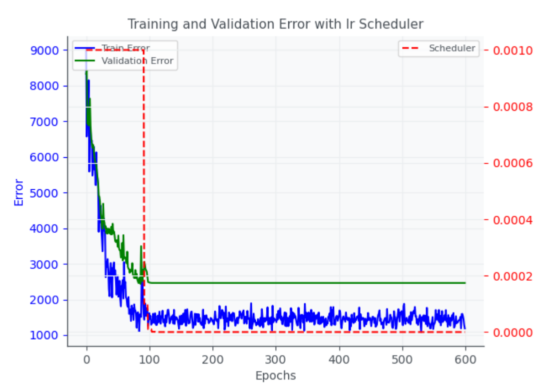

Figure 5: Training progress showing training and validation loss over epochs, and learning rate adjustments

Results

The regression model aimed to predict the time to termination of AF episodes based on segmented ECG signals. However, it was observed that the model did not identify a coherent link between each segment and its time to termination.

On the other hand, the classification results yielded promising performance metrics. The model achieved an overall accuracy of 88.3%, with a precision of 76.7% and a recall of 93.3%. These results indicate that the model was effective in distinguishing between terminating and non-terminating AF episodes, with a high level of recall indicating the model's ability to correctly identify terminating AF episodes.

The results shown in 6. The x-axis represents the true labels, indicating whether an AF episode actually terminated or not, while the y-axis represents the model's predictions. The green background highlights the area where the model correctly classified terminating AF episodes. Conversely, the red background indicates the region where the model correctly classified non-terminating AF episodes. Samples falling on the white background represent instances where the model made incorrect classifications. Furthermore, a

Figure 6: Regression plot illustrating predicted versus actual time until AF termination for the test set. the green and red regions represent correctly classification (Terminating, Non-Terminating respectively

confusion matrix (Figure 7) provides insight into the classification performance of the model, showcasing the distribution of true positive, true negative, false positive, and false negative predictions.

Our study demonstrates improved performance 1 in classifying atrial fibrillation (AF) episodes compared to prior research. With an accuracy of 88.3%, our model surpasses the 71.86% accuracy reported in previous work [4].

Discussion

The classification results yielded promising performance metrics, with our model achieving an accuracy of 88.3%, precision of 76.7%, recall of 93.3%, and specificity of 93%. These results indicate the effectiveness of our approach in distinguishing between terminating and non-terminating AF episodes.

In addition to the regression approach employed in this study, a notable difference compared to previous methods lies in the complexity level of the model. While previous

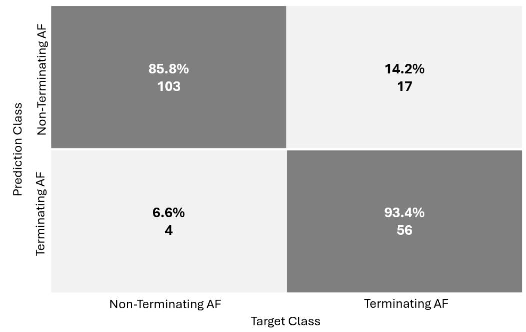

Figure 7: Confusion matrix depicting the classification performance of the model.

models may have been highly sophisticated and intricate in their design, our approach offers a more streamlined and simplified solution. This simplicity is advantageous, particularly in the context of small datasets, as it reduces the risk of over-fitting and enhances generalizability.

One avenue for future research is to augment the ECG signals with additional features extracted using classic methods. While our study focused primarily on utilizing raw ECG signals for AF classification and prediction, integrating complementary features derived from traditional signal processing techniques could provide valuable insights and improve model performance.

By incorporating features such as heart rate variability and QRS complex morphology, a more comprehensive representation of the underlying cardiac activity can be obtained. These features capture various aspects of cardiac dynamics and may offer complementary information to enhance the discriminative power of the model.

Conclusion

In summary, this study introduces a novel approach to classifying and predicting atrial fibrillation (AF) episodes using electrocardiogram (ECG) signals. By combining regression-based prediction of AF termination time with classification of AF episodes, our model demonstrates promising performance metrics, effectively distinguishing between terminating and non-terminating AF episodes. However, it is important to acknowledge the limitations of this work. Our study is based on a small dataset, and further validation on a larger and more diverse dataset is warranted to assess the robustness and generalizability of the findings. Additionally, the results of the regression model may be improved by exploring alternative architectures or incorporating classic feature extraction methods. These limitations provide opportunities for future research to refine and enhance the proposed approach, ultimately advancing our understanding and management of AF in clinical practice.

References

[1] Centers for Disease Control and Prevention (CDC). Atrial fibrillation, 2024.

Table 1: Comparison of Classification Metrics

Model	Accuracy [%]	F1 Score [%]	Sensitivity [%]	Specificity [%]
Our Model	88.33	82.21	76.71	93.33
Previous Work	71.86 ± 3.4	66.02 ± 4.21	76.36 ± 4.18	75.48 ± 3.57

[2] George E Moody. Spontaneous termination of atrial fibrillation: a challenge from physionet and computers in cardiology 2004. In *Computers in Cardiology, 2004*, pages 101–104. IEEE, 2004.

[3] Raúl Alcaraz, Frida Sandberg, Leif Sörnmo, and José Joaquín Rieta. Classification of paroxysmal and persistent atrial fibrillation in ambulatory ecg recordings. *IEEE Transactions on Biomedical Engineering*, 58(2):1441–1449, 2011.

[4] Tejas Radhakrishnan, Jay Karhade, Samit Kumar Ghosh, Priya Ranjan Muduli, RK Tripathy, and U Rajendra Acharya. Afcnnet: Automated detection of af using chirplet transform and deep convolutional bidirectional long short term memory network with ecg signals. *Computers in Biology and Medicine*, 137:104783, 2021.

[5] Emilly M. Lima, Antônio H. Ribeiro, Gabriela MM Paixão, Manoel Horta Ribeiro, Marcelo M. Pinto Filho, Paulo R. Gomes, Derick M. Oliveira, Ester C. Sabino, Bruce B. Duncan, Luana Giatti, Sandhi M. Barreto, Wagner Meira, Thomas B. Schön, and Antonio Luiz P. Ribeiro. Deep neural network estimated electrocardiographic-age as a mortality predictor. *Nature Communications*, 12, 2021.

[6] Sean Shensheng Xu, Man-Wai Mak, and Chi-Chung Cheung. Towards end-to-end ecg classification with raw signal extraction and deep neural networks. *IEEE journal of biomedical and health informatics*, 23(4):1574–1584, 2018.