Predicting Atrial Fibrillation from Intensive Care Unit Numeric Data

Sean McMillan[1], Ilan Rubinfeld[2], Zeeshan Syed[1]

[1]University of Michigan, Ann Arbor, United States of America
[2]Henry Ford Hospital, Detroit, United States of America

Abstract

Atrial fibrillation is a common occurrence in intensive care units (ICUs) and is associated with a significant increase in patient mortality and morbidity, healthcare costs, and length of hospital stay. This burden can be significantly reduced through clinical tools to identify patients at increased risk of developing atrial fibrillation during ICU admission and to match these patients to appropriate prophylaxis (e.g., amiodarone). Unfortunately, despite its prevalence, predicting atrial fibrillation remains a challenge. In this paper, we address the goal of developing an accurate approach to stratify patients for atrial fibrillation using information available in numerics data (e.g., vital signs, arterial blood pressures) commonly collected during ICU admission. We explore the use of a support vector machine (SVM) classifier optimized for multivariate non-linear performance using an area under the receiver operating characteristic curve (AUROC) loss function with summary features derived from ICU numerics collected during the first 8 hours of admission. When evaluated on a cohort of 1,531 ICU patients, this approach achieved an AUROC of 0.73 and a sensitivity of 71% in identifying patients who experienced atrial fibrillation during admission using data only from the start of ICU hospitalization.

1. Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia (affecting an estimated 2.7 million Americans in 2010 [1,2]) and is frequently observed in intensive care unit (ICU) patients. It is independently associated with a 3- to 5-fold increased risk of stroke, and an almost doubled risk of mortality [3,4]. It is also associated with a 73% increase in direct medical costs of patient care due to patients with AF requiring both longer hospital stays and more resources during treatment [5,6]. In the U.S., the direct annual cost for each inpatient who experienced AF was 19,575 USD in 2011 compared to 4,809 USD for inpatients who did not experience AF [7]. As populations continue to age across the world, the burden of AF and AF-related illnesses, which disproportionately affect patients in elderly cohorts, is expected to grow significantly.

Clinical tools to predict AF are a valuable resource in reducing this burden. In addition to guiding monitoring and treatment decisions (especially the use of prophylactic therapies such as anticoagulants and antiarrhythmics), having a long prediction horizon provides a way to better understand the physiological mechanisms underlying the arrhythmia. In this paper, we focus on realizing these opportunities. We address, in particular, the goal of predicting AF by using numerics data (e.g., vital signs and arterial blood pressure) that are commonly available in most ICU settings. We propose a support vector machine (SVM) classifier to stratify patients using summaries of these data from the first 8 hours of ICU admission for subsequent AF. In contrast to traditional SVM classification, which focuses only on reducing the error assessed at the level of individual training examples, we achieve this stratification by exploring an SVM classifier optimized for multivariate non-linear performance using an area under the receiver operating characteristic curve (AUROC) loss function [8]. This approach allows for a reduction of training error on individual examples to be augmented with the joint maximization of multivariate population-level performance metrics [8]. When evaluated in a preliminary study on over 1,500 ICU patients, our AUROC maximizing SVM-approach was able to use the first 8 hours of ICU numerics data to achieve moderate to high levels of discrimination for subsequent AF.

2. Background

Prior work on predicting AF falls into two categories. The first category differs from the problem statement considered here, and focuses on long-term prediction of AF (e.g., the likelihood of developing AF over the next decade). Schnabel et al. [9], for example, present a risk score for predicting long-term AF, using criteria such as age, body mass index, prior occurrences of AF, and other cardiac factors. While such approaches are useful from the perspective of information lifestyle and other

preventative medical choices, they are not appropriate for the short-term assessment of patients in an ICU setting. More closely related to our work is the second category of AF prediction studies, which focus on short-term prediction of AF. The majority of this research has been conducted exclusively through the use of electrocardiogram (ECG) data. For example, it has previously been shown that various aspects of the P-wave morphology can be used as a risk marker for AF [10]. In our study, we differ from these efforts and address the need for short-term prediction of AF by exploring the use of ICU numerics data to stratify patients for AF. Our decision to focus on ICU numerics data, rather than ECG waveforms, is based on two considerations. First, we believe that this approach offers potentially complementary information to information available through a study of ECG activity, and may reflect variations affecting physiological systems beyond electrophysiological phenomena that may be associated with AF onset. Second, the use of numerics data also allows for our efforts to be widely useful without the need for either highly accurate segmentation algorithms or hand-labeling to label P-waves, and without the need for resources to acquire high-resolution ECG from patients. These factors have led to many of the studies proposing ECG-based markers for AF to be limited to small patient cohorts [11-14]. We therefore position our research as exploring an alternate direction for short-term prediction relative to a rich and growing literature on ECG-based AF prediction.

3. Methodology

3.1. Population and data

The cohort for our study represents the entire set of ICU admissions at the Henry Ford Hospital in Detroit, MI from February 2011 until September 2011. Each patient had time-series for heart rate, respiration rate, oxygen saturation, and systolic, diastolic, and mean arterial blood pressures sampled at one minute intervals when the signal was present. Of the patients admitted to the ICU, 1,531 met the inclusion criteria of: (i) having measures for each of heart rate, respiration rate, oxygen saturation, and systolic, diastolic, and mean arterial blood pressures; and (ii) having these measures recorded for at least 24 hours prior to the first episode of AF (in patients with AF) and for at least 24 hours otherwise. The presence of AF was based on automated rhythm annotations archived within the MetaVision ICU System (IMD Soft, Tel Aviv, Israel) at the Henry Ford Hospital. Table 1 presents the characteristics of the patients who met these criteria. Also shown are the characteristics of patients with varying amounts of missing data within this group.

3.2. Experimental design

Data for the first 8 hours of ICU admission was used for each patient, with patients assigned labels (+1 if AF occurred and -1 if AF did not occur) depending on whether patients experienced AF subsequently. In addition to performing experiments on the entire set of 1,531 patients who met the inclusionary criteria described in Section 3.1, two subsets were also studied. These included patients who met the stricter criteria of having less than 50% and less than 10% of the data missing (Table 1).

The raw time series for each signal was passed through a median filter of length 5 minutes to smooth the data and make it more robust to noise. The data was then represented through summary features for the first 8 hours corresponding to the mean, standard deviation, range, maximum positive change, maximum negative change, and slope assessed using a least square regression

Table 1. Baseline characteristics of patients meeting the inclusionary criteria for the study. Characteristics are shown for all patients (<100% of the data missing) as well as for patients who had at least a certain fraction of one minute samples available for each of their signals (<50% or <10% of the data missing). Characteristics of patients with and without AF are shown separately. Given the inclusionary criteria of at least 24 hours of data prior to the first episode of AF all patients with AF did not experience the arrhythmia during the first 24 hours.

		Total	Age								Sex	
			10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99	M	F
<100%	AF	98	0.0000	0.0306	0.0102	0.0102	0.1531	0.3367	0.2857	0.1531	0.0204	0.6531	0.3469
	No AF	1433	0.0119	0.0321	0.0447	0.1298	0.2526	0.2729	0.1458	0.0963	0.0140	0.5645	0.4354
<50%	AF	62	0.0000	0.0323	0.0161	0.0000	0.1613	0.3226	0.2903	0.1774	0.0000	0.6774	0.3226
	No AF	893	0.0134	0.0336	0.0459	0.1310	0.2632	0.2576	0.1433	0.0974	0.0146	0.5722	0.4278
<10%	AF	47	0.0000	0.0426	0.0213	0.0000	0.1489	0.3191	0.2979	0.1702	0.0000	0.7021	0.2979
	No AF	460	0.0152	0.0370	0.0500	0.1304	0.2587	0.2696	0.1435	0.0826	0.0130	0.6130	0.3870

for each signal. Gaps present in the data were ignored in the calculations of these features and mean imputation was used to address missing values corresponding to patients with fewer than 5 samples for their signals within the first 8 hours of ICU admission.

Model training and testing was performed using a 10-fold cross-validation approach with performance assessed using the AUROC and sensitivity (recall)/positive predictive value (PPV or precision) performance metrics.

3.3. AUROC maximizing SVM

An AUROC maximizing SVM [8] was trained for each cross-validation fold in Section 3.2. The distances of test points from the separating hyperplanes were used to measure AUROC for assessing performance, while sensitivity and PPV were measured by choosing the cut point from the AUROC curve that maximized the minimum of sensitivity and PPV.

More formally, the training data for model development can be represented as n training examples of the form $(\mathbf{x}_1,y_1),...,(\mathbf{x}_n,y_n)$, where each $(\mathbf{x}_i,y_i)$ tuple denotes the risk variables $\mathbf{x}_i$ for patient i and the corresponding label $y_i \in \{+1,-1\}$ signifying whether the patient experienced an event (+1) or remained event free (-1). The AUROC for a model parameterized by the weight vector $\mathbf{w}$ (i.e., $\hat{y}_i = \mathbf{w}^T\mathbf{x}_i$ is the label predicted by the model) can then be computed from the number of swapped pairs Φ corresponding to the number of pairs of examples that are ranked in the incorrect order as:

$$\text{AUROC} = 1 - \Phi / n^+ n^-$$

where:

$$\Phi = |\{(i,j) : (y_i > y_j) \text{ and } (w^T x_i < w^T x_j)\}|$$

and n^+ and n^- represent the number of positive and negative examples respectively.

An alternate formulation of measuring the AUROC can be obtained by re-expressing the training data in terms of comparable positive-negative pairs. In this case, the data are represented as tuples of the form $(\mathbf{x}_{ij},y_{ij})$ where $y_{ij} = 1$ and $\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$ for a given pair of positive $(\mathbf{x}_i)$ and negative $(\mathbf{x}_j)$ training examples. The error between the predicted $\hat{y}_{ij} = \mathbf{w}^T\mathbf{x}_{ij} = (\mathbf{w}^T\mathbf{x}_i - \mathbf{w}^T\mathbf{x}_j)$ and y_{ij} is then proportional to 1 - AUROC. Defining this quantity as the AUROC loss function:

$$\Delta_{\text{AUROC}}(y,\hat{y}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{2}(1-\hat{y}_{ij}) = \Phi$$

where $\mathbf{y} = (1,..,1)^T$ and $\hat{\mathbf{y}}$ is a vector denoting the $\hat{y}_{ij}$ predicted by the model stacked together, SVM training can be framed as finding a solution to the following optimization problem:

$$\min_{\mathbf{w},\xi\geq 0}\frac{1}{2}\|\mathbf{w}\|^2 + C\xi$$

subject to $\forall_{\hat{\mathbf{y}}\in\mathbf{Y}\setminus\mathbf{y}}$:

$$\mathbf{w}^T[\sum_{i=1}^{n} y_i\mathbf{x}_i - \sum_{i=1}^{n}\hat{y}_i\mathbf{x}_i] \geq \Delta_{\text{AUROC}}(\mathbf{y},\hat{\mathbf{y}}) - \xi$$

where ξ corresponds to the slack variable, C represents the cost variable, and $\mathbf{Y}$ is the universe of possible vectors $\hat{\mathbf{y}}$. Due to the exponential size of $\mathbf{Y}$, the problem is solved efficiently using a sparse approximation-based approach [8].

4. Results

The results for the different experiments are presented in Table 2. While model performance was generally poor when imposing no restrictions on the amount of data that could be missing, the results improved substantially when looking at the subset of patients with at least 50% or 90% of the data present for each signal. This result can be attributed to the availability of more data to fully characterize patient state over the entire 8 hour period at the start of admission. We note that the amount of missing data as a feature by itself was not a significant predictor of AF (AUROC 0.53, sensitivity 0.22, PPV 0.19).

5. Discussion

This study explored the feasibility of an ICU numeric-based approach to stratify patients for AF using data that are available at the very start of ICU admission. In particular, we proposed the use of an SVM model optimized for non-linear multivariate performance using an AUROC loss function that was trained on summary features for the first 8 hours of ICU admission to predict subsequent AF. Our preliminary findings are promising, and suggest that an ICU numeric and AUROC SVM-based approach may have value in guiding monitoring and prophylaxis for AF.

Table 2. Performance for the experiments (results shown separately for different amounts of data missing).

	<100%	<50%	<10%
AUROC	0.55	0.67	0.73
Sensitivity	63%	88%	71%
PPV	9%	11%	18%

Acknowledgements

This research was supported, in part, by the National Institutes of Health (National Library of Medicine).

References

[1] Artucio H, et al. Cardiac arrhythmias in critically ill patients: epidemiologic study. Critical Care Medicine 1990;18:1383-8.

[2] Roger VL, et al. Heart diseases and stroke statistics—2012 update: a report from the american heart association. Circulation 2012;125:e2-e220.

[3] Wolf PA, et al. Atrial fibrillation as an independent risk factor for stroke: the framingham study. Stroke 1991;22:983-988.

[4] Benjamin EJ, et al. Impact of atrial fibrillation on the risk of death: the framingham study. Circulation 1998;98:946-952.

[5] Kim MH, et al. Estimation of total incremental health care costs in patients with atrial fibrillation in the united states. Circulation: Cardiovascular Quality and Outcomes 2011;4:313-320.

[6] Crawford TC, et al. Cardiac arrhythmias: management of atrial fibrillation in the critically ill patient. Critical Care Clinics 2007;23:855-872.

[7] Wu EQ, et al. Economic burden and co-morbidities of atrial fibrillation in a privately insured population. Current Medical Research and Opinion 2005;21:1693-1699.

[8] Joachims T. A support vector method for multivariate performance measures. Proceedings of the International Conference on Machine Learning (ICML) 2005.

[9] Schnabel RB, et al. Development of a risk score for atrial fibrillation (framingham heart study): a community-based cohort study. The Lancet 2009;373:739-745.

[10] Guidera SA, et al. The signal-averaged P-wave duration: a rapid and noninvasive marker of of risk of atrial fibrillation. Journal of the American College of Cardiology 1993;21:1645-1651.

[11] Steinberg JS, et al. Value of the P-wave signal-averaged ECG for predicting atrial fibrillation after cardiac surgery. Circulation 1993;88:2618-2622.

[12] Hayashida N, et al. P-wave signal-averaged electrocardiogram for predicting arrhythmia after cardiac surgery. The Annals of Thoracic Surgery 2005;79:859-864.

[13] Klein M, et al. Use of P-wave-triggered, P-wave signal-averaged electrocardiogram to predict atrial fibrillation after coronary artery bypass surgery. American Heart Journal 1995;129:895-901.

[14] Amar D, et al. Clinical prediction rule for atrial fibrillation after coronary artery bypass grafting. Journal of the American College of Cardiology 2004;44:1248-1253.

Address for correspondence.

Sean McMillan
2260 Hayward Street
Ann Arbor, MI 48109
spmcmill at umich.edu