A Rule-Based Approach for the Prediction of Acute Hypotensive Episodes

MA Mneimneh, RJ Povinelli

Marquette University, Milwaukee, USA

Abstract

The 2009 Physionet Computers in Cardiology Challenge is to develop an automatic technique for the prediction of acute hypotensive episodes. Such episodes if left untreated can lead to organ failure and death. However, acute hypotensive episodes can be treated depending on its cause, and if diagnosed in time. This study uses a rule-based approach for the prediction of the acute hypotensive episodes. The physionet challenge is divided into two parts. The first is the ability to distinguish between patients who have experienced acute hypotensive episodes and patients who do not. The second part of the challenge is to predict acute hypotensive episodes. The technique uses the mean arterial blood pressure signal as an indicator for predicting AHE. The records used for this challenge are from the MIMIC II database. The challenge dataset is divided into 60 records for training, 10 records for testing the first part of the challenge, and 40 records for testing the second part of the challenge. The score is calculated as the fraction of correct classifications. The final results on the challenge are 1 and 0.875 for part 1 and part 2, respectively.

1. Introduction

Acute hypotensive episodes (AHE) are considered one of the most vital events. These events, if left untreated, can cause irreversible organ damage and death. However, quick intervention of such events is required to reduce the above risks. The goal of the Physionet /Computers in Cardiology challenge 2009 is to predict which patients will have an acute hypotensive episode within the forecast window [1]. The forecast window is defined as the one-hour subsequent to a certain time instance T_0.

The Physionet /Computers in Cardiology challenge 2009 is divided into two challenge events. The first event is to predict if patients who are receiving pressor medication will have an acute hypotensive episode in the forecast window. The second event of the challenge is to predict if patients will have an AHE within the forecast window.

By defining the mean arterial pressure (MAP) as the average of the blood pressure measured in the radial artery over the previous minute, AHE is defined for the purpose of this challenge as the period where ABP is below 60 mmHg for a period of 30 minutes and more [1].

In 2001, Bassale showed that there is a correlation between arterial blood pressure variability and predicting acute hupotensive episodes [2]. During the work of [2], the analysis showed that the arterial blood pressure (ABP) signal showed significant changes prior to an acute hypotension episode.

During the course of this challenge, several techniques were attempted to provide a solution for the challenge. The first approach was to use a reconstructed phase space neural network (RPS-NN) based approach for the prediction of AHE. The second approach uses a nearest neighbor approach to detect which MAP records show similar behavior. The third approach is a rule based approach. The third approach received the highest score among the other two. Therefore, this paper focuses on presenting the rule-based approach for the prediction of AHE. The method uses the mean arterial blood pressure as an indicator to predict AHE. The approach is based on determining the best threshold that can be used to identify which patients will have an AHE. Since, the main indicator used to identify acute hypotensive episode is the mean arterial blood pressure, the presented approach utilizes the MAP to predict AHE based on the MAP signal 20 minute prior to T_0.

2. Data set and preprocessing

The challenge dataset used in this work is a set of 110 records from the MIMIC II database. The MIMIC II database contains about 30,000 patient records from patients' entire stay in an intensive care unit [3].

The challenge dataset is divided into training and testing sets. The training set consists of 60 records divided equally into four groups of patients, acute hypotensive episodes treated with pressors (H_1), acute hypotensive episodes not treated with pressors (H_2), and no acute hypotensive episodes treated with pressors (C_1), and no acute hypotensive episodes and not treated with pressors (C_2) outside the forecast window. The testing set for part one of the challenge consists of 10 records. The number of records in the testing set of the second part of the challenge is 40. Each record proved the mean arterial blood pressure (MAP), which is the signal used in this work to predict AHE.

Since the whole duration of the training data is provided. The MAP signal had to be truncated until T_0. In order to determine the position of T_0 in the signal, it was compared to the start time of the record.

3. Methods

Three approaches were used to determine the method for predicting acute hypotensive episodes. The first utilizes an RPS-NN approach [4], the second is a nearest neighbor approach, and the third is a rule-based approach.

3.1. Reconstructed phase space-neural network approach (first pass)

A reconstructed phase space (RPS) is a time delayed embedding of a signal, which may be topologically equivalent to the state space of the system that generated the signal if certain assumptions are met [8]. We have shown in previous work that even when these assumptions are not met RPSs contain important information for classifying a signal [4]. Here the signal is the MAP signal, and the classes are patients who will have hypotensive episodes and patients who will not. The definition of each point in an RPS is determined as follows:

$$X_n = \left[x_{n-(d\ 1)\tau} \quad x_n \ \ldots \ x_n \right]_\tau \tag{1}$$

Following the determination of the set of features X_n, a neural network classifier is used.

For this approach, the truncated MAP signal of the training data is used to construct the vector X_n. The MAP signal of the test set was resampled from 1 sample per second to 1 sample per minute to match that of the training data.

3.2. K-nearest neighbor approach (second pass)

The k-nearest neighbor approach is a machine learning method that classifies an object based on the majority of neighbors closer to the object. The parameter k is a positive integer. In this work, the object is the average of the 20 minute MAP signal prior to T_0. The parameter k is selected to be 3.

3.3. Rule-based approach (third and fourth pass)

The motivation behind this approach came from observing the data, and from previous work that showed that MAP is an excellent indicator for predicting acute hypotensive episodes[1, 2]. The method uses the average of the 20 minutes of the MAP signal prior to forecast window. An optimization technique is applied to automatically determine the best threshold that can be used to classify patient who will have AHE and patients who will not. The training process can be seen in the figure below:

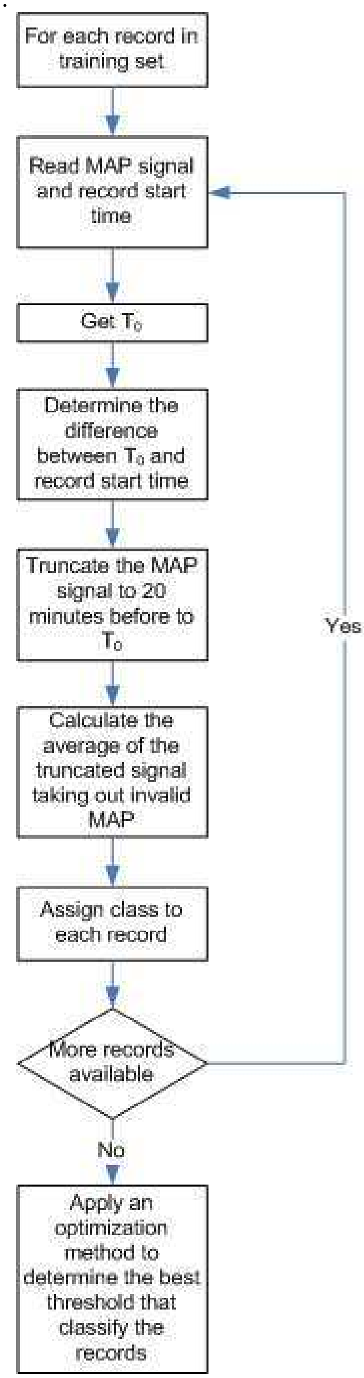

Figure 1: Flowchart describing training process.

From Figure 1, first the MAP signal and start time of each record is read. Second, the record is truncated from the start time until T_0. Third, the average of the truncated signal is calculated taking out all invalid entries in the signal. Fourth, assign a class to every record as given in the training set. After reading all records, an optimization technique is used to determine the threshold that best classifies the training set.

For the testing process, the test data records are read, and the average of the last 20 minutes of the MAP signal is determined. The average is then compared against the threshold. If the average is greater than the threshold, the record is classified as Class C, i.e. patients will not have AHE. Otherwise, the records are classified as Class H, i.e. patient will have AHE. The testing process can be seen in the flowchart below:

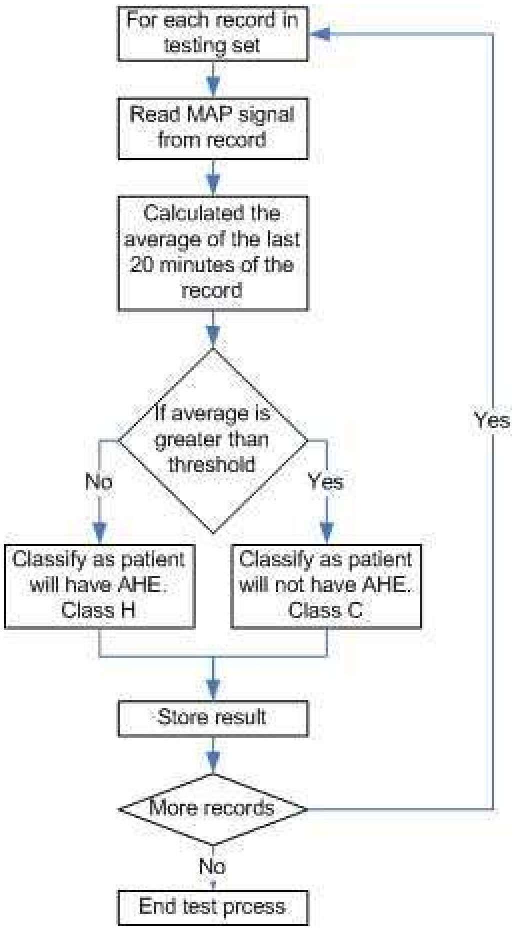

Figure 2: Flowchart describing the testing process.

To validate the threshold obtained from the training process, a ten-fold cross validation is applied to the training data. The ten fold cross validation is described as follows: Divide data into 10 set of size n/10, called folds, train on 9 sets and test on 1 set, repeat the process 10 times and store the diagnostic results, and finally combine the results and calculate the overall accuracy. The ten fold cross validation test ensures that the training and testing sets are patient independent.

4. Results

The method presented in this paper is applied to selected records from the MIMIC II database, which consist of the training and testing sets of the Physionet/Computers in Cardiology challenge 2009. The resulting threshold for the average of the t_0 – 20min signal of the MAP signal is 71.1mmHg.

The ten fold cross validation is applied to the training set. The resulting accuracy is 68.34% with sensitivity of 83.34%, and specificity of 53.34%. The confusion matrix for the proposed threshold applied to the training set of the Physionet/ Computers in Cardiology 2009 challenge can be seen as:

Table 1: Confusion matrix for the training data classification.

	Classified as	
	Class C	Class H
Class C	25	5
Class H	14	16

Figure 3 shows the average of 20min prior to T_0 of the training data separated by the threshold of 71.1mmHg.

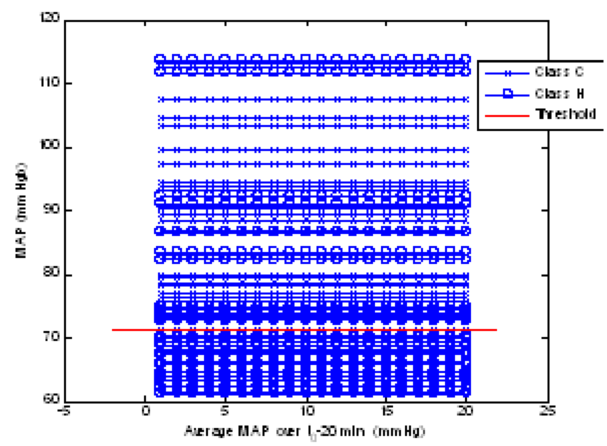

Figure 3: The training data separated by the threshold.

The challenge's score for each of the event is calculated as the fraction of correct classifications. Four attempts were made on the challenge. The first is using the RPS-NN approach, which yield a score of 0.2 for the first entry and 0.625 for the second entry. For the second attempt, the k-nearest neighbor approach was applied. The score for the second attempt was 1 for the first entry, and 0.825 for the second entry. For the third attempt, the rule-based approach's score is 1 for the first entry and 0.9 for the second. The sensitivity of the rule base approach for the second entry is 92.85% and specificity of 88.46%. The confusion matrix for the results is:

Table 2: Confusion matrix for the rule base approach applied to testing data classification.

	Classified as	
	Class C	Class H
Class C	23	3
Class H	1	13

For the fourth attempt, the rule based approach used training data from classes C_1 and H_1, and a threshold on 72mmHg was determined. Additionally, the rule based approach was trained on classes C_2 and H_2, and the resulting threshold is 74mmHg. A majority vote was made among applying the thresholds to the testing set. The result for entry 1 remained 1, and for entry 2 is 0.825.

5. Discussion and conclusions

As a summary, this paper presents a rule-based approach applied to the Physionet/Computers in Cardiology 2009 dataset. The rule based approach showed excellent result in predicting acute hypotensive episodes based on 20 minutes information prior to the forecast window using mean arterial blood pressure signal. The score of the challenge is 1 for the first entry and 0.9 for entry 2. The approach showed that the mean arterial blood pressure can be used to predict acute hypotensive episodes.

The rule based approach provides a guideline for predicting acute hypotensive episode. The approach can be used in real time patient monitoring systems in order to predict AHE. The training process for the rule base approach is performed offline. The prediction process is performed online.

References

[1] Moody G, Lehman L. Predicting Acute Hypotensive Episodes: The 10th Annual PhysioNet/Computers in Cardiology Challenge. Computers in Cardiology 2009; Park City, Utah. 2009; 36 (in press).

[2] Bassale J. Hypotension Preduction Arterial Blood Pressure Variability. Portland; 2001. Report No.: 01-002.

[3] Goldberger A, Amaral L, Glass L, Hausdorff J, Ivanov P, Mark R, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. AHA Journal Circulation 101. 2000;101(23):e215-e20.

[4] Mneimneh MA, Povinelli RJ. RPS/GMM Approach toward the Localization of Myocardial Infarction. Computers in Cardiology; 2007; North Carolina, USA. 2007; 34:185-8.

Address for correspondence
Mohamed A. Mneimneh
Mohamed.Mneimneh@ieee.org