

Predicting Ventricular Tachycardia Using LSTMs and ECG Signals

David Tinarwo
tinarwodave@gmail.com
(469)516-6938

Abstract

Ventricular tachycardia (VT) is one of the main causes of sudden cardiac death (SCD) in the world. An algorithm that is capable of reliably predicting an imminent episode of VT from a continuously recorded electrocardiogram (ECG) signal could potentially be incorporated into an implantable device capable of delivering preventative treatments and therefore reduce the frequency of SCD. In this article, we present and discuss one such algorithm that predicts VT by sequentially feeding consecutive heartbeats (in the form of consecutive chunks of ECG signals) into a trained Long Short Term Memory (LSTM) neural network. To evaluate this algorithm, we extracted many sequences of consecutive heartbeats each of which spans a time duration of t_d seconds and ends bce seconds prior to the beginning of one of a number of cardiac events, including VT, from physician-annotated ECG signals. When these sequences of consecutive heartbeats were fed into the trained LSTM neural network, no clear trend between the values of bce and t_d and the sensitivity and specificity of the LSTMs ability to predict episodes of VT could be deduced. When $t_d = 2$ seconds and $bce = 0$ seconds, a sensitivity of 95.2% and a specificity of 85.5% in VT prediction were produced, a promising but not groundbreaking result. This sensitivity and specificity suggest that the use of LSTMs in

the prediction of VT has potential and therefore encourages additional optimizations of LSTMs for this and similar purposes.

Introduction

The electrocardiogram (ECG) is a safe, noninvasive method commonly used to quickly diagnose cardiac diseases and to detect and predict cardiac arrhythmias, including ventricular tachycardia (VT) [1]. VT is a type of rapid cardiac arrhythmia in which the heart beat ectopically originates in the ventricles. Studies have found that about 80% of instances of sudden cardiac death (SCD) are caused by spontaneous ventricular tachyarrhythmia (VTA), a category that includes VT [2]. As a result, an algorithm capable of predicting VT from ECG signals can potentially contribute to the decreased prevalence of SCD and the decreased mortality of patients with heart conditions.

There are already several such algorithms reported and discussed in the literature for predicting VT from ECG signals of patients. Some of these algorithms used particular morphological features, such as Heart Rate Variability [2] or QT interval variability [3]. In this article, we present a VT predictor that segments the ECG signal of an ICU patient into separate consecutive heart beats and then feeds them into a trained Long Short Term Memory (LSTM) network, a type of neural network specifically well-suited to learn from experience when there are very long time lags of unknown lengths between important events [4]. There are a few additional advantages of this

approach to the design of the prediction algorithm including the fact that it does not require the prior identification of relevant features of the ECG signal and therefore does not rely on the particular cardiac event that the algorithm seeks to predict, which in this case is VT.

Data

Sources of the ECG Signals

All of the ECG signals used in this article were downloaded from the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) II databases, which contain records of physiological signals and vital time series collected from Intensive Care Unit (ICU) patients [5]. The particular MIMIC II database used in this article is the European ST-T Database , which contains 90 two-hour excerpts of 1 two-channel ambulatory ECG readings obtained from 79 subjects and which was digitized at 250 samples per second per channel with 12-bit resolution over a 20 mV range [6]. This database is available on the Physionet website [5]. This database contains annotations that indicate the beginning and ending of each heartbeat and the beginning of each occurrence of the following types of cardiac events: (1) atrial bigeminy (AB), (2) atrial fibrillation (AFIB), (3) ventricular bigeminy (B), (4) normal sinus rhythm (N), (5) supraventricular tachyarrhythmia (SVTA), (6) ventricular trigeminy (T), (7) sinus bradycardia (SBR) and (7) ventricular tachycardia (VT).

Extraction of Relevant Segments of the ECG Signal

The segments of ECG signals that were used to train and test the LSTM network were extracted from the European ST-T Database. To complete this extraction, each annotated occurrence of the eight cardiac events listed above is considered and the sequence of consecutive heart beats that spans a specific time duration, td seconds, and that ends a specific amount of time, bce seconds, before the beginning of the cardiac event is extracted, if possible, as shown in Figure (1). More specifically, the sequence of heartbeats that begins with the first beat that starts less than a specific amount of time, $bce + td$, before the beginning of the annotated cardiac event and that ends with the last beat that ends more than a specific amount of time, bce , before the annotated cardiac event, is extracted, if possible.

td and bce are both variables that take on a variety of values, as shown in Section IV. These values were chosen by reviewing other papers on the prediction of VT and investigating their chosen values of td and bce [8]. By choosing to use a variety of values instead of choosing a single pair of values, the data produced in this paper potentially can demonstrate the relationship between the values of td and bce and the sensitivity and specificity of the LSTM's ability to predict episodes of VT.

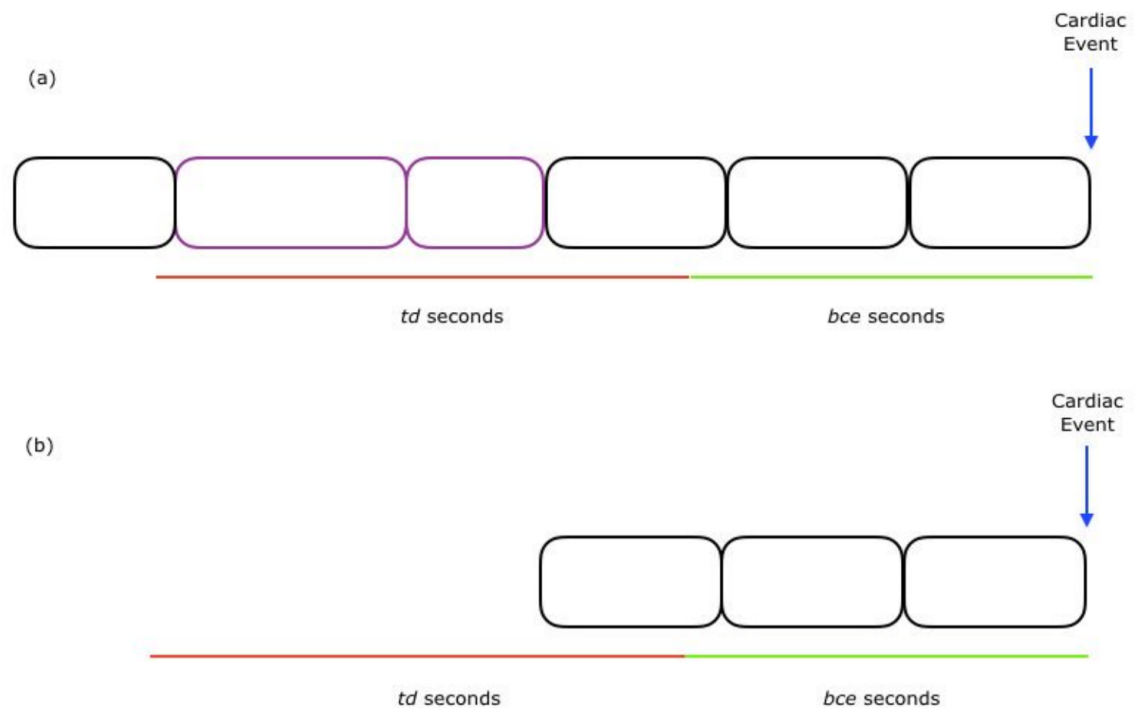


Figure 1:

Two example ECG signal excerpts are schematically shown above, each of which contains an annotated occurrence of a cardiac event that is designated by a blue arrow. The length of each of the two green horizontal lines is *bce* seconds and the length of each of the two red horizontal lines is *td* seconds. Each unextracted heart beat is designated by a black box and each extracted heartbeat is designated by a purple box. Note that the excerpt displayed in (a) contains enough ECG data prior to the beginning of the cardiac event for a relevant segment of ECG signal to be extracted. However, the excerpt displayed in (b) does not contain enough ECG data prior to the beginning of the cardiac event for a relevant segment of ECG signal to be extracted. These extracted sequences of beats can then be used to predict the cardiac events that they proceed.

Separation into Training and Testing Data

The sequences of heart beats extracted from the European ST-T database preceding a particular cardiac event is partitioned 80%-20% respectively into training data (i.e. data that will be used to train the LSTM) and testing data (i.e. data that will be used to evaluate the LSTM's ability to predict episodes of VT) as shown in Table (1).

	Number of Beat Sequences Preceding a Particular Cardiac Event	
	Training Data	Testing Data
Atrial Bigeminy (AB)	35	9
Atrial Fibrillation (AFIB)	2	0
Ventricular Bigeminy (B)	59	15
Normal Sinus Rhythm (N)	541	135
Supraventricular Tachyarrhythmia (SVTA)	35	9
Ventricular Trigeminy (T)	62	16
Sinus Bradycardia (SBR)	13	3
Ventricular Tachycardia (VT)	306	76

Table 1:

The number of training sequences of heart beats (i.e. training ECG readings) and the number of testing sequences of heart beats (i.e. testing ECG readings) preceding each type of cardiac event when $td = 2$ seconds and $bce = 0$ seconds and when the partition of data between the training and testing sets is 80%-20% respectively.

Proposed LSTM

As discussed in Section I, a Long Short Term Memory (LSTM) network is a type of neural network specifically well-suited to learn from experience when there are very long time lags of unknown lengths between important events [4]. As a result, LSTMs are well-suited to solving this problem of predicting episodes of VT from segments of ECG signals. Therefore, an LSTM was designed, trained, and evaluated as discussed in this section.

A Input and Output of the LSTM

Each input to the LSTM corresponds to a single sequence of heartbeats (i.e. a single extracted ECG reading) that together spans a duration of t_d seconds and precedes a cardiac event by t_{ce} seconds. This input takes the form of a two-dimensional array, in which each row contains all of the ECG samples associated with a single heartbeat in the sequence of heartbeats. As a result, the number of columns in each input array is the maximum number of ECG samples in any single extracted heartbeat and the number of rows in each input array is the maximum number of heartbeats in any extracted sequence of heartbeats. Since Python, the programming language used in the production of this paper, does not allow any ragged arrays, the input array must be padded. Left- and top-padding is used. Since portions of the input array are fed into the LSTM top to bottom, left to right, the last numbers to be fed into the LSTM will be the bottom right elements of the input array, which are the non-zero elements. And because the last elements

fed into the LSTM has the greatest impact on the LSTM's output, this padding theoretically should be optimal.

Note that one other input array structure was considered and this alternative option is discussed in Appendix A.2.

The output of the LSTM is an eight element vector, each element of which corresponds to the Bayesian probability that the sequence of heartbeats represented by the input array precedes one of the eight cardiac events listed in Section II.A. For instance, in the code used to produce the data presented in this paper, the seventh element of each output vector corresponds to the Bayesian probability that the sequence of heartbeats represented by the input array precedes an episode of ventricular trigeminy (T).

A training output vector contains seven 0s and one 1 and the position of the 1 within the vector indicates which of the eight cardiac events the sequence of heartbeats actually precedes. A testing output vector, on the other hand, contains eight non-negative floating point numbers, which together add up to 1. The position of the highest floating point number within the testing output vector indicates which of the eight cardiac events the LSTM most strongly believes the sequence of heartbeats precedes. Two example output vectors are shown in Table (2).

Index of Output Vector	Cardiac Event	Training Output Vector	Testing Output Vector
0	Normal Sinus Rhythm (N)	0	0.109
1	Supraventricular Tachyarrhythmia (SVTA)	0	0.012
2	Ventricular Bigeminy (B)	0	0.335
3	Atrial Bigeminy (AB)	0	0.012
4	Ventricular Tachycardia (VT)	0	0.058
5	Atrial Fibrillation (AFIB)	0	0.015
6	Ventricular Trigeminy (T)	1	0.413
7	Sinus Bradycardia (SBR)	0	0.045

Table 2: An example of a training output vector and an example of a testing output vector. The sequence of heartbeats associated with the training output vector precedes an episode of Ventricular Trigeminy (T) and the LSTM believes that the sequence of heartbeats associated with the testing output vector precedes an episode of T.

As is indicated by the structure of the output vectors, the LSTM is trained to be able to predict whether a sequence of heartbeats precedes any one of eight cardiac events and is not trained to only be able to predict whether or not a sequence of heartbeats precedes an episode of VT. This design choice was made in order to prevent the LSTM from overfitting during the training process and should paradoxically improve the LSTM's ability to predict VT.

Basic Structure of the LSTM

The LSTM was built using Keras [9], a Python package often used to quickly implement and test various deep learning algorithms. The structure of the LSTM is 3 depicted in Figure (2).

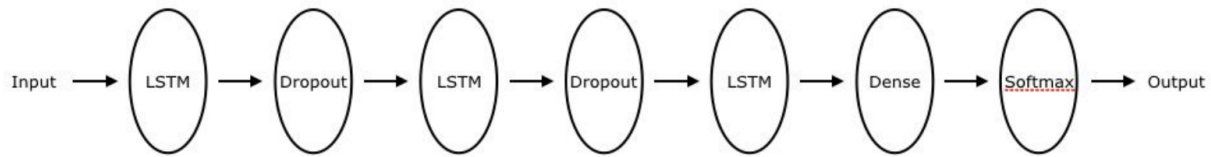


Figure 2: The structure of the LSTM. Note that each of the two Dropout layers, which were added to prevent overfitting of the LSTM model to the trained data, removes 20% of the data that is passed to it. Each of the LSTM layers contain 32 weights. And the Dense layer and the Softmax layer work together to ensure that the output vector has the structure described in Section III.A.

The objective function and the optimizer used to train the LSTM were the categorical cross entropy function and the adam optimizer [10], respectively, both of which had Keras's default parameters.

Results

The implemented LSTM was trained and tested on extracted sequences of heartbeats which were of varying lengths (i.e. varying values of td) and which precede cardiac events by varying amounts of time (i.e. varying values of bce). The sensitivity and the specificity of the trained LSTM's ability to predict episodes of VT for varying values of td and of bce are shown in Tables (3) and (4).

Sensitivity						
		Length of Sequence of Heartbeats (<i>td</i>) in seconds				
		2	7	12	17	22
Length of Time Before Cardiac Event (<i>bce</i>) in seconds	0	0.952	0.947	0.93	0.947	0.973
	5	0.746	0.701	0.989	1.000	0.645
	10	0.678	0.979	0.701	0.710	0.784
	15	0.798	0.909	0.973	0.654	0.838
	20	0.743	0.683	0.659	0.616	0.995

Table 3: The sensitivity of the LSTM's ability to predict episodes of VT for varying values of *td* and of *bce*.

Specificity						
		Length of Sequence of Heartbeats (<i>td</i>) in seconds				
		2	7	12	17	22
Length of Time Before Cardiac Event (<i>bce</i>) in seconds	0	0.855	0.882	0.829	0.855	0.816
	5	0.658	0.750	0.132	0.013	0.671
	10	0.763	0.013	0.816	0.658	0.526
	15	0.618	0.171	0.000	0.684	0.276
	20	0.658	0.803	0.842	0.855	0.000

Table 4: The specificity of the LSTM's ability to predict episodes of VT for varying values of *td* and of *bce*.

A comparison of the sensitivity and specificity achieved by the trained LSTM when *td* = 2 seconds and *bce* = 0 seconds with the sensitivity and specificity achieved by other published machine learning algorithms for predicting VT is shown in Table (5) .

Method	Sensitivity	Specificity	Year
Analysis of unpredictable component of QRS [12]	86.7%	88.1%	2010
Neural network using heart rate variability [13]	82.9%	71.4%	2010
Complex network theory [14]	74.3%	66.7%	2011
Various morphological features [11]	88%	100%	2011
<i>LSTM discussed in this paper</i>	<i>95.2%</i>	<i>85.5%</i>	<i>2016</i>

Table 5: A comparison of the sensitivity and specificity achieved by the trained LSTM when $td = 2$ seconds and $bce = 0$ seconds with the sensitivity and specificity achieved by other published machine learning algorithms for predicting VT.

Discussion and Conclusion

As can be deduced from Tables (3) and (4), the data produced by evaluating the trained LSTM unfortunately does not show any clear trends between the values of bce and td and the sensitivity and specificity of the LSTM's ability to predict episodes of VT. A future iteration of this article should use a wider range (i.e. larger than 20 seconds) and smaller increments (i.e. smaller than 5 seconds) for both bce and td .

Also, as can be seen in Table (5), the specificity and sensitivity achieved by the trained LSTM is higher, and equivalently better, than that achieved by some other recently published algorithms [12, 13, 14] but is lower, and equivalently worse, than that

achieved by at least one recently published algorithm [11]. As a result, the LSTM developed and implemented for this article to predict episodes of VT has potential but does not produce ground-breaking results. Therefore, additional optimizations of the LSTMs, primarily by training the LSTM using a larger set of extracted ECG signals, may prove promising.

Appendix

Rationale behind the Choice of the ECG Database

The European ST-T database was chosen out of all of the ECG databases in MIMIC II to be the source of all of the data used to train and test the LSTM for a variety of reasons. First of all, the European ST-T database contains a large number of ECG excerpts that include the six abnormal cardiac rhythms that are listed in section II.A. As a result, it provides a lot of data that can be used to train the LSTM to be able to predict these abnormal cardiac rhythms. A comparison of the data provided by the European ST-T Database with the data provided by the MIT-BIH Arrhythmia Database [7] is shown in Table (A.1). Notice that the European ST-T database contains over triple the number of beat sequences that can be used to train the LSTM to predict VT.

	Number of Beat Sequences Preceding a Particular Cardiac Event	
	European ST-T Database	MIT-BIH Arrhythmia Database
Atrial Bigeminy (AB)	44	6
Atrial Fibrillation (AFIB)	2	204
Ventricular Bigeminy (B)	74	438
Normal Sinus Rhythm (N)	676	988
Supraventricular Tachyarrhythmia (SVTA)	44	52
Ventricular Trigeminy (T)	78	166
Sinus Bradycardia (SBR)	16	0 ⁶
Ventricular Tachycardia (VT)	382	122

Table (A.1): The number of segments of ECG signals that are of length $t_d = 2$ seconds, that precede each type of cardiac event by $bce = 0$ seconds and that can be extracted from the data stored in the European ST-T database and from the data stored in the MIT-BIH Arrhythmia Database.

As a result, the specificity and sensitivity of the LSTM's ability to predict episodes of VT went up considerably when the source of the ECG signals used to train and test the LSTM was switched from the MIT-BIH Arrhythmia database to the European ST-T database, as shown in Table (A.2).

	European ST-T Database	MIT-BIH Arrhythmia Database
Specificity	0.989	0.992
Sensitivity	0.855	0.042

Table (A.2): The specificity and sensitivity of the LSTM's ability to predict episodes of VT using segments of ECG signals that are of length $t_d = 2$ seconds with data extracted from the MIT-BIH-Arrhythmia test conditions and that precede episodes of VT by $b_{ce} = 0$ seconds when trained and tested with data extracted from the European ST-T Database and when trained and tested with the MIT-BIH-Arrhythmia Database.

Second of all, each ECG signal excerpt in the European ST-T database is 90 minutes long, which is relatively long when compared with the ECG signal excerpts stored in the other ECG databases [5]. As a result, the amount of ECG signal preceding each annotated occurrence of the eight cardiac events is, on average, larger in the European ST-T database, increasing the likelihood of being able to extract a sequence of heartbeats of duration t_d seconds that precedes a given cardiac event by b_{ce} seconds. Therefore, by using an ECG database containing relatively long excerpts, we have effectively increased the amount of extracted ECG segments that can be used to test and train the LSTM and equivalently increased the LSTM's ability to predict episodes of VT.

In addition, the ECG signals used to train and test the LSTM in this article were extracted from only one database because different ECG databases in MIMIC II often contain signals that have been sampled at different frequencies and using ECG signals that have differing sampling frequencies to test and train an LSTM without any prior resampling and interpolation of the ECG signals is very deleterious to the LSTM's prediction abilities as shown in Table (A.3).

	European ST-T Database	European ST-T Database and MIT-BIH Arrhythmia Database
Specificity	0.989	0.986
Sensitivity	0.855	0.644

Table (A.3): The specificity and sensitivity of the LSTM's ability to predict episodes of VT using segments of ECG signals that are of length $td = 2$ seconds and that precede episodes of VT by $bce = 0$ seconds when trained and tested with data extracted from the European ST-T Database and when trained and tested with data extracted from both the MIT BIH-Arrhythmia Database and the European ST-T Database. No prior resampling or interpolation has been performed on any of the data. And the division of data between training and testing is 80% and 20%, respectively. The ECG readings stored in the European ST-T Database were sampled at a frequency of 250 samples per second [6] whereas the ECG readings stored in the MIT-BIH Arrhythmia Database were sampled at a frequency of 360 samples per second [7]

Therefore, in order to use ECG signals stored in multiple databases, the ECG signals need to be interpolated and resampled so that all of the ECG signals that are used to train and test the LSTM, regardless of their source database, artificially have the same sampling frequency. One way to accomplish this goal is to downsample all of the ECG signals to the greatest common divisor of all of the sampling frequencies. However, this downsampling reduces the amount of data that is used to train the LSTM and therefore negatively impacts the LSTM's ability to predict episodes of VT, as shown in Table (A.4).

	60% Training Data 40% Testing Data	80% Training Data 20% Testing Data
Specificity	0.962	0.989
Sensitivity	0.862	0.855

Table (A.5): The specificity and sensitivity of the LSTM's ability to predict episodes of VT using segments of ECG signals that are of length $td = 2$ seconds and that precede episodes of VT by $bce = 0$ seconds when trained and tested with data extracted from the European ST-T Database with varying divisions of the ECG data between training and testing data.

Rationale behind the Choice of the Structure of the Input to the LSTM

One other structure for the input array was considered. This alternative structure took the form of a vector containing all of the ECG signals associated with a single sequence of heartbeats. These input vectors can be fed into the LSTM in batches of size greater than one or of size one.

Since all of the input vectors in a single batch must have the same dimensions, using batches of size greater than one requires padding the input vectors. Following the same logic described in Section III.A, it is optimal to use left-padding and not right-padding. An LSTM produced using this structure of the input array, batches of size greater than 1 and left padding predicts episodes of VT with lower specificity and sensitivity than an LSTM produced using the structure of the input array described in Section III.A, as shown in Table (A.6).

	Input Array Structure Described in Section III.A	Input Vector Structure Described in Appendix A.2, Batches of Size Greater Than One And Left Padding
Specificity	0.989	0.697
Sensitivity	0.855	0.786

Table (A.6): The specificity and sensitivity of the LSTM's ability to predict episodes of VT using segments of ECG signals that are of length $td = 2$ seconds and that precede episodes of VT by $bce = 0$ seconds when trained and tested with different structures of the input array.

Using batches of size one would avoid the use of padding, which is optimal.

Unfortunately, due to technical difficulties, the author of this paper was unable to get this

method of using a batch of size one working. The data presented in Table (A.6) seems to indicate that this structure of the input array does not produce high performance LSTMs and, equivalent, is not promising.

References

1. K. Khunti, "Accurate interpretation of the 12-lead ECG electrode placement: A systematic review", Health Educ. J., vol. 73, no. 5, pp. 610-623, 2013
2. J. Segyeong, "A Predictor for Ventricular Tachycardia based on Heart Rate Variability Analysis" IEEE Conference, pp. 409-411, 2011.
3. G. Larisa, "Prediction of Ventricular Tachyarrhythmias by Intracardiac Repolarization Variability Analysis" Journal of the American Heart Association, pp.276-284, 2009
4. Pankaj Malhotra, Lovekesh Vig, Gautam Shroff and Puneet Agarwal, "Long Short Term Memory Networks for Anomaly Detection in Time Series", European Symposium on Artificial Neural Networks, vol 23, Computational Intelligence and Machine Learning, Belgium, 2015.
5. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet:

Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220 [Circulation Electronic Pages;

<http://circ.ahajournals.org/cgi/content/full/101/23/e215>]; 2000 (June 13).

6. Taddei A, Distant G, Emdin M, Pisani P, Moody GB, Zeelenberg C, Marchesi C. The European ST-T Database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *European Heart Journal* 13: 1164-1172 (1992).

7. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220 [Circulation Electronic Pages;

<http://circ.ahajournals.org/cgi/content/full/101/23/e215>]; 2000 (June 13).

8. Riasi, Atiye, and Maryam Mohebbi. "Prediction of ventricular tachycardia using morphological features of ECG signal." *Artificial Intelligence and Signal Processing (AISP)*, 2015 International Symposium on. IEEE, 2015.

9. Chollet, Francois, Keras, (2015), GitHub repository, <https://github.com/fchollet/keras>

10. Kingma, Diederik, and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).

11.Riasi, Atiye, and Maryam Mohebbi. "Prediction of ventricular tachycardia using morphological features of ECG signal." Artificial Intelligence and Signal Processing (AISP), 2015 International Symposium on. IEEE, 2015.

12.C. Lin., "Analysis of unpredictable components within QRS complex using a finite-impulse-response prediction model for the diagnosis of patients with ventricular tachycardia" in Computers in Biology and Medicine, vol. 40, pp. 643-649, 2010.

13.S.Joo, and K.Choi, "Prediction of ventricular tachycardia by a neural network using parameters of heart rate variability" Computer in cardiology, Vol. 37, pp. 585-588, 2010.