

# Comparison of the Effects of Cross-validation Methods on Determining Performances of Classifiers Used in Diagnosing Congestive Heart Failure

Yalcin Isler<sup>1</sup>, Ali Narin<sup>2</sup>, Mahmut Ozer<sup>2</sup>

<sup>1</sup>Department of Biomedical Engineering, Izmir Katip Celebi University, Balatcik Campus, 35620, Izmir, Turkey

<sup>2</sup>Department of Electrical and Electronics Engineering, Bulent Ecevit University, Incivez, 67100, Zonguldak, Turkey

**Congestive heart failure (CHF) occurs when the heart is unable to provide sufficient pump action to maintain blood flow to meet the needs of the body. Early diagnosis is important since the mortality rate of the patients with CHF is very high. There are different validation methods to measure performances of classifier algorithms designed for this purpose. In this study, k-fold and leave-one-out cross-validation methods were tested for performance measures of five distinct classifiers in the diagnosis of the patients with CHF. Each algorithm was run 100 times and the average and the standard deviation of classifier performances were recorded. As a result, it was observed that average performance was enhanced and the variability of performances was decreased when the number of data sections used in the cross-validation method was increased.**

**Keywords:** Heart rate variability, heart failure, cross-validation, classification.

## 1. INTRODUCTION

THE MOST IMPORTANT FUNCTION of the heart is to pump enough oxygen and nutrients which tissue and cells require [1]. Heart failure emerges when the heart is not able to pump enough blood efficiently [2]. Due to the inadequacy in the blood circulation, the blood piles up in some tissues and veins, which is called congestion. Because the congestion occurs among the patients with heart failure in general, the disease is commonly called Congestive Heart Failure (CHF) [3].

Hospital discharges for Heart Failure increased 174% between 1979 and 2003 (National Hospital Discharge Survey, CDC/NCHS). In addition, based on the 44-year follow-up of the NHLBI's FHS, 80% of men and 70% of women under age 65 who have Heart Failure will die within 8 years. In people diagnosed with HF, sudden cardiac death occurs at 6-9 times the rate of the general population [4]. The death rate will be decreased and the life quality of the patients will be increased if the patients with CHF can be diagnosed early. Analysis of heart rate variability (HRV) is used for the diagnosis of CHF in addition to most clinical examination tests such as electrocardiography, echocardiography and chest x-ray which are commonly used [3, 4].

Recently HRV analysis, which is derived from intervals between two adjacent peaks of ECG records, has been used in the patients with CHF [5, 6, 7, 8, 9, 10, 11]. For example, Isler and Kuntalp studied the CHF discrimination task using classical time- and frequency-domain measures by combining Wavelet Entropy measures and the Nearest Neighbor classifier resulting in the sensitivity rate of 79.3% and the specificity rate of 94.4% [5]. The same authors also studied the same task by adding heart rate normalization and the k-Nearest Neighbors (k=3) classifier resulting in the sensitivity rate of 82.76% and the specificity rate of 100% [6]. Yu and Lee used standard

time- and frequency-domain features with bi-spectral analysis based features and Support Vector Machines (SVM). They achieved the sensitivity rate of 95.55% and the specificity rate of 100% for RBF kernel with the selected features using Genetic Algorithm (GA) [7] and the sensitivity rate of 93.10% and the specificity rate of 98.14% for linear kernel [8]. Jovic and Bogunovic used standard time-domain features and non-linear measures with SVM, MLP, C4.5, and Bayesian classifiers resulting in sensitivity rates of 77.2%, 96.6%, 99.2%, and 98.4%, respectively and specificity rates of 87.4%, 97.8%, 98.4%, and 99.2%, respectively [9]. Pecchia et al used non-standard features and CART classifier resulting in the sensitivity rate of 89.7% and the specificity rate of 100% [10]. In another previous study, Narin et al. achieved the sensitivity of 82.75%, the specificity of 96.29%, and the accuracy of 91.56% using various classifiers and the selected features by the backward elimination method [11]. The performances of these classification algorithms vary according to the selection of acquired features [5, 11, 12], normalization procedures [6] and other pre-processing applications [13]. There have been various cross-validation methods for the evaluation of performances for these algorithms in the literature [14]. The aim of this study is to examine the effects of cross-validation methods of k-fold (k=2,3,5,10) and leave-one-out on the performances of various classifiers such as LDA, KNN, MLP, RBF and SVM. These methods and classifiers are commonly used in the classification of the biomedical signals in the literature.

## 2. METHODS

### 2.1. Data

The HRV data used in this study were obtained from the normal sinus rhythm and congestive heart failure RR interval databases from the MIT/BIH database in PhysioNET

(<http://www.physionet.org>) [15]. The HRV data used are as follows:

- **Congestive Heart Failure RR Interval Database (chf2db)** has data from 29 patients with CHF (8 men, 2 women and 19 unrecorded) with an age range of 34-79 years, and
- **Normal Sinus Rhythm RR Interval Database (nsr2db)** has data from 54 normal subjects (30 men, 24 women) with an age range of 24-76 years.

Although databases are composed of 24(h) HRV data, only 5 minutes (300 seconds) of HRV data is used to achieve results faster and to disturb the patients at least in a clinical environment [16].

## 2.2. Heart rate variability measures

HRV data is collected after determining QRS wave structure in the ECG signals because QRS wave structure is the component that has the most distinctive amplitude values in the ECG signals. After the determination of QRS, the time difference of two consecutive R peaks is commonly called RR-interval ( $T_n = t_n - t_{n-1}$ ) [16]. Patient information (age), analysis of time domain (mean, standard derivation and other statistical measures), analysis of frequency domain (spectral power values through different frequencies) and evaluations obtained by non-linear methods are used in the HRV studies. For further details, the studies given in the references could be examined [2, 11].

Welch periodogram method, which makes use of Fast Fourier Transform (FFT) method, is commonly used for the evaluation of frequency domain [5]. By using this method, power spectrum density (PSD) over the data that is sampled at equal intervals through the time can be computed. In this study, HRV data is re-sampled by using the cubic interpolation method at 4Hz before using FFT and detrending of data is eliminated for providing stability analysis [13]. In this study, evaluation of frequency domain obtained by Lomb periodogram, which is developed as an alternative to the classical PSD method, is also used. Owing to this method, PSD can be computed directly from the HRV data without needing to re-sample on time domain [17]. HRV evaluation of conventional frequency domain (Table 1) for 5-minute period is examined in detail by following recommended standard [16].

There are three frequency bands commonly used in HRV analysis: VLF (0 - 0.033 Hz), LF (0.033 - 0.15 Hz) and HF (0.15 - 0.4 Hz) [16]. In the evaluation of frequency domain, peak frequencies and power of these frequency bands from calculated PSD are examined. Therefore, the following frequency domain measures were computed separately for both methods of Welch periodogram and Lomb periodogram.

On the other hand, wavelet analysis not only makes possible the examination of one signal for both the time and scale domains but also eliminates polynomial non-stability [18]. Wavelets are reported very useful for analysis of the RR intervals due to this capability. Thus, Daubechies-4 main wavelet

Table 1: Standard frequency domain HRV measurements

Measure	Description
VLF	The total power of the frequency band of VLF
LF	The total power of the frequency band of LF
HF	The total power of the frequency band of HF
LFHF	The ratio of the frequency band of LF/HF
NLF	Normalized LF power of LF/(LF+HF)
NHF	Normalized HF power of HF/(LF+HF)

over HRV data re-sampled at 4 Hz with the 7-level [5] wavelet transform method was also used in this analysis to calculate the standard frequency domain features of HRV analysis (Table 1).

Furthermore, non-linear methods including Poincare map, approximate entropy (AppEn), sample entropy (SampEn), detrended fluctuation analysis (DFA) and symbolic dynamics were used in this study. *Poincare Plot*, which is a technique taken from non-linear dynamics, is a graph of each RR interval plotted against the next interval. The plot provides summary information as well as detailed beat-to-beat information on the behavior of the heart [19, 20]. The Poincare Plot is a popular technique thanks to its simple visual interpretation and its proved clinical ability as a predictor of disease and cardiac dysfunction [21]. Fitting an ellipse to the Poincare Plot's shape and calculating standard deviations of the distance of the points on the plot is the most popular method [22]. *Detrended Fluctuation Analysis* (DFA) is used to quantify the fractal scaling properties of short interval RR interval signals. This technique is a modification of root-mean-square analysis of random walks applied to non-stationary signals [23]. The root mean-square fluctuation of an integrated and detrended time series is measured at different observation windows and plotted against the size of the observation window on a log-log scale.

*Symbolic Dynamics*, as an approach to investigate complex systems, facilitates the analysis of dynamic aspects of the signal of interest. The concept of symbolic dynamics is based on a coarse-graining of the dynamics [24]. That is, the range of original observations or the range of some transform of the original observations, such as the first difference between the consecutive values, is partitioned into a finite number of regions and each region is associated with a specific symbolic value so that each observation or the difference between successive values is uniquely mapped to a particular symbol depending on the region into which it falls. For instance, the heart rate time series ( $RR_1, RR_2, \dots, RR_n$ ) can be transformed into a symbol sequence ( $s_1, s_2, \dots, s_n$ ) with symbols from the alphabet 0, 1, 2, 3 using

$$\begin{aligned}
 0, & \quad RR_i \leq \mu - a \\
 1, & \quad \mu - a < RR_i \leq \mu \\
 2, & \quad \mu < RR_i \leq \mu + a \\
 3, & \quad \mu + a < RR_i
 \end{aligned}$$

where  $\mu$  denotes the mean and the  $a$  is a constant equal to the standard deviation of all RR intervals [25]. In this way

some detailed information is lost, but a more general dynamic behavior can be analyzed [26, 25]. Then, Shannon entropies from probability values of successive L-length symbols are calculated. In this study, values of L were tested using all integer values between 1 and 10. SampEn is a new family of statistics measuring complexity and regularity of clinical and experimental time-series data [27]. SampEn statistics provide an improved evaluation of time-series regularity and is a useful tool in the studies of the dynamics of human cardiovascular physiology. In SampEn, the comparison between the template vector and the rest of the vectors excludes the comparison with itself. This causes that probabilities may be zero. SampEn is the modified and corrected form of *AppEn*.

As a result, 59 features were used. Those were obtained from patient information, time domain evaluation, frequency spectrum evaluation and non-linear methods with the numbers of 1, 4, 38 and 16, respectively.

### 2.3. Normalization

The used features must be equalized on the same scale because units and amplitudes of features used in classification and clustering algorithms differ highly compared to each other. This process is called normalization [14]. In this study, Min-Max method in which all features vary through the scale [0, 1] is used.

### 2.4. Classification algorithms

#### 2.4.1. Linear discriminant analysis (LDA)

Linear discriminant analysis (LDA) is a statistical analysis method which can differentiate the groups according to two or more features. Discriminant functions are used mathematically. LDA is a traditional method for both the size reduction and the classification [14].

#### 2.4.2. K-nearest neighbors (KNN) classifier

The main idea of the K-Nearest Neighbors (KNN) classifier is to predict the class of a test case based on its k-nearest neighbors. It consists of two steps: the first step aims to find the k-nearest neighbors, whereas the second one predicts the class of the test sample based on these neighbors. Although several distances could be used to estimate the k-nearest neighbors, the Euclidean distance and a weighted voting strategy were used in this paper. The subject of interest is classified by the majority classes; then, the subject is being assigned to the class most common amongst its k nearest neighbors [14].

#### 2.4.3. Multi-layer perceptron (MLP)

Multi-Layer Perceptron (MLP) is a widely used artificial neural network structure [14]. MLP can be used in both linear and non-linear applications. MLP network consists of three layers in general: an input layer, one or more hidden layers, and an output layer. The number of neurons in the input layer is equal to the number of features (or inputs) and the number

of neurons in the output layer is equal to the number of outputs. The input-layer neurons are connected to all the hidden-layer neurons and the hidden-layer neurons are connected to all the output-layer neurons with weights. The weight values are changed until the minimum error value is found, which is called training of the network. The general output function and the error function are defined as follows:

$$y_i = f\left(\sum_{i=1}^N w_{ji}x_i\right) \quad (1)$$

$$E = \frac{1}{2} \sum_i (d_i - y_i)^2 \quad (2)$$

where  $x_i$  is the input data,  $w_{ji}$  is weight values,  $f(\cdot)$  is an activation function,  $y_i$  is i-th network output,  $d_i$  is i-th expected output.

#### 2.4.4. Radial basis functions (RBF)

Radial basis function (RBF), another artificial neural network structure, tends to learn faster than the multi-layer perceptron [14]. It is also composed of three layers similar to MLP. Activation function in the hidden layer is Gaussian function. The output for Gaussian activation function in general is defined as follows:

$$y_i(x) = \sum_{i=1}^k w_{ji} e^{-\frac{\|x - \mu_j\|^2}{2\sigma_j^2}} \quad (3)$$

where  $w_{ji}$  are weights,  $k$  is the number of hidden layer neurons,  $x$  is input data,  $\mu_j$  is j-th neuron centers,  $\sigma_j$  is j-th propagation parameter and  $y_i$  is the i-th output of the system. The output of the system is obtained with hidden layer output and weights of the system. Thus, it is aimed to determine the optimum weight values.

#### 2.4.5. Support vector machines (SVM)

Support vector machines (SVM) is based on statistical learning rule. SVM is widely used for many applications, both linear and non-linear data points. The main purpose of support vector machines is to find a hyper-plane that separates the two groups. Two boundary lines, which are close and parallel to each other in two groups, are defined. These boundary lines are called support vectors. This classifier is based on determining an ordinary discrimination plane from the boundary lines [14].

Some data cannot be separated linearly in SVM. In such cases, data is moved to another dimension using the kernel function which may be a non-linear transformation.

### 2.5. Cross-validation methods

In cross-validation method, database is divided into two groups. When the first group (training database) is used for determination of model parameters of a classifier, the other group (test database) is used for testing the performance of the trained classifier [14].

### 2.5.1. k-fold cross-validation method

In the k-fold cross-validation method, the database that has total  $n$  samples is divided to  $k$  groups that have  $\frac{n}{k}$  samples in each group. Each time, when a different data group is set aside for the test, the other  $k - 1$  groups are used for training. The classifier is trained  $k$  times by switching the test cluster in every cycle. In this way, the performance of a classifier is determined by calculating the mean of the error of  $k$  times.

### 2.5.2. Leave-one-out cross-validation method

In the leave-one-out cross-validation method, the database with the  $n$  samples is divided to  $n$  groups. The classifier is trained with the  $n - 1$  group by leaving one group for the test each time. In this manner, the performance of classifier is determined by calculation of the mean value for error of 83 times. In other words, this method can be considered as k-fold cross-validation method for the  $k = n$ .

### 2.6. Performance measures

The performance of the classifier is given as follows [6]:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$SEN = \frac{TP}{TP + FN} \quad (5)$$

$$SPE = \frac{TN}{TN + FP} \quad (6)$$

If the instance is positive and it is classified as positive, it is counted as true positive (TP); if it is classified as negative, it is counted as false negative (FN). If the instance is negative and it is classified as negative, it is counted as true negative (TN); if it is classified as positive, it is counted as false positive (FP).

## 3. RESULTS

In this study, all the HRV measures, all the algorithms and cross-validation methods were calculated by MATLAB 2013a. The HRV measures were investigated to distinguish 29 patients with CHF from 54 healthy subjects in the control group. Each algorithm was run 100 times and the average and the standard deviation was recorded. The classifier parameters used in this study were chosen as follows:

- **LDA:** Both linear and quadratic functions,
- **KNN:** Odd values between 1 and 19,
- **MLP:** The number of neurons in hidden layer from 1 to 50,
- **RBF:** The distribution parameter of the Gaussian function from 0.1 to 3.0 by the increment of 0.1,
- **SVM:** Linear kernel function.

Table 2: Classifier Performance by cross-validation method where L.O.O. stands for the leave-one-out method

	Validation	SEN	SPE	ACC
LDA	2-fold	76.3±3.9	79.1±5.0	78.2±3.7
	3-fold	78.2±3.6	78.7±4.2	78.5±3.1
	5-fold	67.7±3.9	80.2±4.2	75.8±3.3
	10-fold	79.6±2.0	79.2±2.3	79.3±1.7
	L.O.O.	79.3±1.2	79.3±0.9	79.4±0.6
KNN	2-fold	50.6±6.7	90.7±3.3	76.7±3.1
	3-fold	49.5±4.7	95.8±1.2	79.6±1.7
	5-fold	49.8±3.8	96.1±1.0	79.9±1.5
	10-fold	50.8±4.2	94.8±1.1	79.4±1.7
	L.O.O.	52.1±1.1	96.3±0.5	80.9±0.4
MLP	2-fold	53.9±17.3	81.4±14.7	71.8±8.8
	3-fold	57.4±13.9	83.1±10.7	74.1±6.3
	5-fold	59.1±10.5	85.5±8.7	76.3±5.3
	10-fold	56.7±9.5	87.4±5.3	76.7±4.4
	L.O.O.	60.9±6.7	85.4±4.0	76.9±3.6
RBF	2-fold	60.2±11.5	78.4±9.1	72.1±5.4
	3-fold	59.3±9.1	78.1±7.4	71.6±4.5
	5-fold	53.8±8.1	84.9±4.8	74.0±3.2
	10-fold	54.4±5.0	87.3±3.3	75.8±2.8
	L.O.O.	60.3±1.0	85.5±2.7	76.7±1.1
SVM	2-fold	64.2±8.3	80.4±6.0	74.7±4.4
	3-fold	66.8±6.0	80.4±4.8	75.6±3.4
	5-fold	69.4±5.5	80.4±4.2	76.6±3.4
	10-fold	71.3±3.5	80.1±3.4	77.0±2.5
	L.O.O.	75.1±1.6	81.6±1.2	79.3±0.9

The performances of each classifier were determined by using 2-fold, 3-fold, 5-fold, 10-fold and leave-one-out cross-validation methods that are commonly used in the literature. The obtained performances of SEN, SPE and ACC were arranged as mean ± standard deviation (Table 2). According to the results, leave-one-out cross-validation methods have both the highest values of the average performance of the classifier and the lowest standard deviation values. Also, when the number of parts increases in the cross-validation method, standard deviation of the performance of the classifier has been decreasing gradually. The maximum performance was achieved by using KNN classifier with leave-one-out cross-validation methods by means of SPE and ACC in this study. On the other hand, LDA classifier gave the maximum SEN rate during the study.

## 4. DISCUSSION

The results, as seen from Table 2, show that the choice of the cross-validation method to determine the classifier performances affects the results of the study. According to the table, if the number of folds is increased, the average classifier performance is increased while the variance is decreased. If the reproducibility of the results is preferred, the leave-one-out method may be the best choice. The number of folds (data parts) is the number of data samples in the leave-one-out method. The data for train and test purposes are the same

in every selection in this method, which eliminates the randomness of data parts. Thus, the leave-one-out method gives the minimum variability in performance evaluation of classifiers. Similar evidences have been achieved in the literature [29].

As a result, the maximum performances of the constructed systems were achieved by using KNN classifier with leave-one-out cross-validation methods as 52.10%, 96.29%, and 80.85% for the sensitivity, specificity, and accuracy values, respectively. Although this gave the highest accuracy in this study, it had a poor sensitivity. Due to this fact, this classifier can be used as a negative predictor of CHF, which means that the subject is normal if this classifier reports negative but the subject may not be a patient with CHF even if this classifier reports positive result. On the other hand, LDA gave 79.31%, 79.27%, and 79.35% for the sensitivity, specificity, and accuracy values, respectively. If the positive prediction is also important, then LDA can be chosen as a classifier.

Although these were not the highest values in the literature, some studies have reported higher performances than this study by either failing to include the gender information [5, 7, 8] or using different datasets [9]. Because the gender information is not recorded for most of the patients in the chf2db database, including it in evaluating the performance misleads the performance of designed classifiers [6, 11].

Because the discrimination power of the methods was not the main purpose of the study, no feature selection method was used in the study. It is possible to achieve higher performances by applying one or more feature selection methods.

In addition, in probability theory, the central limit theorem (CLT) states that, given certain conditions, the arithmetic mean of a sufficiently large number of iterates of independent random variables, each with a well-defined expected value and well-defined variance, will be approximately normally distributed, regardless of the underlying distribution [?]. That is, suppose that a sample is obtained containing a large number of observations, each observation being randomly generated in a way that does not depend on the values of the other observations, and that the arithmetic average of the observed values is computed. Although the number of patients with CHF and normal subjects is supposed to meet the condition, the higher number of records will give better comparable results.

## 5. ACKNOWLEDGEMENT

Authors want to express their special thanks to Ms Deniz Angun for her support on improving the paper.

## REFERENCES

- [1] U.S. National Library of Medicine. *Heart failure (Medical Encyclopedia)*. <http://www.nlm.nih.gov>.
- [2] Flavell, C., Stevenson, L.W. (2001). Take Heart with Heart Failure. *Circulation*, 104, 89.
- [3] Wilbur, J., James, P. (2005). Diagnosis and management of heart failure in the outpatient setting. *Primary Care*, 32, 1115–1129.
- [4] American Heart Association (2006). Heart Disease and Stroke Statistics-2006 Update: A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*, 113, 85–151.
- [5] Isler, Y., Kuntalp, M. (2007). Combining Classical HRV Indices with Wavelet Entropy Measures Improves to Performance in Diagnosing Congestive Heart Failure. *Computers in Biology and Medicine*, 37(10), 1502–1510.
- [6] Isler, Y., Kuntalp, M. (2010). Heart Rate Normalization in the Analysis of Heart Rate Variability in Congestive Heart Failure. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 224(3), 453–463.
- [7] Yu, S.N., Lee, M.Y. (2012). Bispectral analysis and genetic algorithm for congestive heart failure recognition based on heart rate variability. *Computers in Biology and Medicine*, 42, 816–825.
- [8] Yu, S.N., Lee, M.Y. (2012). Conditional mutual information-based feature selection for congestive heart failure recognition using heart rate variability. *Computer Methods and Programs in Biomedicine*, 108, 299–309.
- [9] Jovic, A., Bogunovic, N. (2011). Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. *Artificial Intelligence in Medicine*, 51, 175–186.
- [10] Pecchia, L., Melillo, P., Sansone, M., Bracale, M. (2011). Discrimination power of short-term heart rate variability measures for CHF assessment. *IEEE Transactions on Information Technology in Biomedicine*, 15(1), 40–46.
- [11] Narin, A., Isler, Y., Ozer, M. (2014). Investigating the performance improvement of HRV Indices in CHF using feature selection methods based on backward elimination and statistical significance. *Computers in Biology and Medicine*, 45, 72–79.
- [12] Narin, A., Isler, Y. (2012). Effect of Principal Component Analysis on Diagnosing Congestive Heart Failure Patients using Heart Rate Records. In *IEEE 20th Signal Processing and Communications Applications Conference (SIU2012)*, 18-20 April 2012, Fethiye / Mugla.
- [13] Isler, Y., Selver, M.A., Kuntalp, M. (2005). Effects of Detrending in Heart Rate Variability Analysis. In *II. Muhendislik Bilimleri Genç Arastirmacilar Kongresi MBGAK'2005*, 17-19 October 2005, Istanbul, 213–219.
- [14] Duda, R.O., Hart, P.E., Stork, D.G. (2000). *Pattern Classification*, 2nd edition, New York. Wiley.
- [15] Goldberger, A.L., et.al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23), e215–e220.

- [16] Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Circulation*, 93, 1043–1065.
- [17] Lomb, N.R. (1976). Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39, 447–462.
- [18] Quiraga, R.Q., Rosso, O.A., Basar, E., Schurmann, M. (2001). Wavelet entropy in event-related potentials: A new method shows ordering of EEG oscillations. *Biological Cybernetics*, 84(4), 291–299.
- [19] Woo, M.A., Stevenson, W.G., Moser, D.K., Trelease, R.B., Harper, R.H. (1992). Patterns of beat-to-beat heart rate variability in advanced heart failure. *American Heart Journal*, 123, 704–710.
- [20] Kamen, P.W., Krum, H., Tonkin, A.M. (1996). Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity. *Clinical Science*, 92, 201–208.
- [21] Kamen, P.W. (1996). Heart rate variability. *Australian Family Physician*, 25, 1087–1094.
- [22] Brennan, M., Palaniswami, M., Kamen, P. (2001). Do existing measures of poincare plot geometry reflect non-linear features of heart rate variability? *IEEE Transactions on Biomedical Engineering*, 48(11), 1342–1347.
- [23] Huikuri, H.V., Makikallio, T.H., Peng, C.K., Goldberger, A.L., Hintze, U., Moller, M. (2000). Fractal correlation properties of R-R interval dynamics and mortality in patients with depressed left ventricular function after an acute myocardial infarction. *Circulation*, 101, 47–53.
- [24] Acharya, U.R., Kannathal, N., Seng, O.W., Ping, L.Y., Chua, T. (2004). Heart rate analysis in normal subjects of various age groups. *BioMedical Engineering OnLine*, 3(24).
- [25] Caminal, P., Vallverdu, M., Giraldo, B., Benito, S., Vazquez, G., Voss, A. (2005). Optimized symbolic dynamics approach for the analysis of the respiratory pattern. *IEEE Transactions on Biomedical Engineering*, 52(11), 1832–1839.
- [26] Xu, J.H., Liu, Z.R., Liu, R. (1994). The measures of sequence complexity for EEG studies. *Chaos*, 4(11), 2111–2119.
- [27] Richman, J.S., Randall, M.J. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology - Heart and Circulatory Physiology*, 278, H2039–H2049.
- [28] Akgul, A. (2003). Tibbi Arastirmalarda Istatistiksel Analiz Teknikleri: SPSS Uygulamalari (Statistical Analysis Techniques in Medical Researches: SPSS Experiments). 2<sup>nd</sup> edition, Ankara, Turkey, Seckin Yayincilik.
- [29] Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *14th International Joint Conference on Artificial Intelligence (IJCAI)*, 20-25 August 1995, Montreal, Quebec, Canada.

Received August 18, 2014.

Accepted August 5, 2015.