Süleyman Demirel Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 22, Sayı 2, 380-387, 2018

Süleyman Demirel University
Journal of Natural and Applied Sciences
Volume 22, Issue 2, 380-387, 2018

DOI: 10.19113/sdufbed.82260

ECG Data Compression Using ε-insensitive Quadratic Loss Function

Omer KARAL[*1]

[1]Yıldırım Beyazıt University, Faculty of Engineering and Natural Sciences, Electrical and Electronics Engineering. 06010, Ankara

(Alınış / Received: 07.11.2017, Kabul / Accepted: 07.03.2018, Online Yayınlanma / Published Online: 18.04.2018)

Keywords
Data compression,
Electrocardiogram,
Modeling,
Quadratic loss function,
Support vector regression

Abstract: ElectroCardioGram (ECG) is a graphical representation of the electrical activity that occurred during the heartbeat. It plays a significant role in the diagnosis and analysis of heart diseases. ECG signals must be recorded continuously for the effective detection and diagnosis of heart diseases. However, such records as it produces large amounts of data at a level that makes it difficult storage and transmission can also be impaired due to the ambient noise. Thanks to the reasons mentioned above, an efficient ECG data compression algorithm is required even in a noisy environment. This study proposes ε-insensitive quadratic loss based Support Vector Regression (ε-quadratic SVR) technique for the compression of ECG signals. There is a well-known relationship between loss functions and noise distributions. The proposed ε-insensitive quadratic loss function provides the optimal solution against Gaussian noise. Computer simulation results show that the proposed loss function is an attractive candidate for ECG data compression in the presence of Gaussian noise.

ε-duyarsız Kuadratik Kayıp Fonksiyonu ile EKG Verilerinin Sıkıştırılması

Anahtar Kelimeler
Veri sıkıştırma,
Elektrokardiyogram,
Modelleme,
Kuadratik kayıp fonksiyonu,
Destek vektör regresyon

Özet: ElektroKardiyoGram (EKG), kalbin çalışması esnasında ortaya çıkan elektriksel aktivitenin grafiksel bir gösterim şeklidir. Kalp hastalıklarının teşhisinde ve analizinde önemli bir rol oynamaktadır. Kalp hastalıklarının önceden etkin bir şekilde tespiti ve teşhisi için, EKG sinyalleri sürekli kaydedilmesi gerekir. Bununla birlikte, depolama ve aktarımın zorlaştığı bir seviyede büyük miktarda veri üreten kayıtlar, ortam gürültüsünden dolayı da bozulabilir. Bu nedenlerden dolayı, gürültülü bir ortamda bile etkin sonuçlar verebilecek bir EKG veri sıkıştırma algoritmasına ihtiyaç vardır. Bu çalışma EKG sinyallerinin sıkıştırılması için ε-duyarsız kuadratik kayıplı Destek Vektör Regresyon (ε-kuadratik DVR) tekniğini önermektedir. Kayıp fonksiyonları ile gürültü dağılımları arasında iyi bilinen bir ilişki vardır. Önerilen ε-duyarsız kuadratik kayıp fonksiyonu ise Gauss gürültüsüne karşı en uygun çözümü sunar. Bilgisayar simülasyon sonuçları, önerilen kayıp fonksiyonunun Gauss gürültüsü ile bozulmuş EKG verilerinin sıkıştırılması için çekici bir aday olduğunu göstermektedir.

1. Introduction

Electrocardiogram (ECG) is a method widely used by various healthcare institutions to record the electrical activity that occurs in conduction and relaxation phases of the atria and ventricles of the heart. It plays a very important role in the diagnosis and analysis of heart diseases. For the early diagnosis of heart diseases, ECG signals are continuously recorded. However, long term monitoring generates large amount of data that will make storage and transmission difficult. Moreover, these records may be subject to noise resulting from the recording environment. These problems can be overcome by compressing the ECG signals containing large amounts of data efficiently (preserving important information) at appropriate compression rates. Many algorithms have recently been suggested for compressing, modeling and reconstructing ECG signals. Existing algorithms can be divided into three categories [1]: I) direct data compression techniques, II) the parameterized model-based techniques and III) the transform based techniques. All of the abovementioned techniques also can be lossy or losses. For a detailed review of these methods, see [2-4]. In this paper, a lossy compression technique belonging to the last category is presented and evaluated.

*Corresponding author: karal@ybu.edu.tr

Transform based methods are generally preferred because they have higher compression ratios and are more insensitive to noise. In this article, a new lossy compression technique for the transform-based compression category is proposed and evaluated. In transform-based techniques, the signal is represented as a linear weighted sum of basis functions. The parameters of these basis functions such as coefficients (weights), and variances are appropriately coded and transmitted instead of original signal. Therefore, the best transformation is a transformation that needs a minimum number of basis functions to reconstruct the original signal for the given error criterion. Among transform-based techniques, the Karhunen-Loeve transformation (KLT) is the optimal transformation in that it requires a minimum number of basis functions to represent a random signal according to the mean square error criterion [5]. However, even though it gives a high compression ratio, the calculation time required to compute the KLT basis functions is too long. For this reason, suboptimum transform techniques are used more often for ECG signal compression. Reddy and Murthy [6] used the Fourier transform (FT) technique for compressing ECG data and showed that FT gives good results even in noisy recordings. However, it fails to determine the exact location of frequency information over time. Benzid et al. [7] obtained a very high compression ratio (CR) with discrete cosine transform (DCT). However, Shinde and Kanjalkar [8] have shown that Percentage Root Mean Square Difference (PRMSD) value is directly proportionally to CR, and thus leading to deterioration of the signal in DCT technique.

In recent years, wavelet-based transform (WBT) techniques have been widely used in ECG signal compression, due to their characteristics such as energy compression, time-frequency localization, [2]. WBT techniques are mainly divided into two distinct categories: the continuous wavelet transform (CWT) and the discrete wavelet transform (DWT). These two classes were analyzed separately by the Addison in 2005 [9]. He emphasized that since the DWT yields fewer wavelet coefficients, it needs less computation time than the CWT. Moreover, he noted that the DWT has an interesting mathematics and is based on traditional signal filtering and encoding methodologies. However, WBT techniques have a proportional relationship between block size and compression ratio. If the block size increases, then the compression ratio also increases for a given error criterion; but the calculation time and storage requirements of the adaptive wavelet coding schemes will also increase in the same way. For these reasons, how to decide the block size in ECG data compression with wavelet transform is still a fundamental problem [2].

Recently, unlike current transform based compression techniques, Karal [10] has proposed a new method based on standard support vector regression (SVR) to compress ECG data according to the given error tolerance in an optimal and rapid manner. It has also been shown that the performance of the proposed standard SVR-based compression technique is higher than that of transform-based compression techniques such as FT, DCT, and DWT commonly used in the literature. In terms of robustness, there is an effective connection between error distribution and loss function [11]. This relation stems from the Bayesian approach. In other words, if the error distribution is known, the optimal loss function can be derived using the Bayesian approach. Standard SVR technique adopts Vapnik's ε-insensitive loss function which is optimal for Laplacian error distribution. However, most of the current regression (compression) algorithms consider Gaussian error distribution. For this reason, SVR with Vapnik loss function (standard SVR) is not suitable for applications with Gaussian error distribution.

In order to address the aforementioned problem, in this paper, it is proposed that ε-insensitive quadratic loss function is included in the standard SVR formulation. The proposed ε-insensitive quadratic loss function has a significant advantage in that it properly takes into account Gaussian noise distribution in terms of robustness in SVR based compression. It also exploits most of the advantages of Vapnik's loss function such as disregarding small noises and sparseness in solution representation.

The main contributions of this study can be listed as follows:

(1) SVR with ε-insensitive quadratic loss function is derived for the ECG data compression.
(2) Since the proposed ε-insensitive quadratic loss function is convex and continuously differentiable, it can be solved by a quadratic programming.
(3) The performance of the proposed algorithm is better than the standard SVR in the presence of Gaussian noise distribution.

The remainder of this study is outlined as follows. The concept of support vector regression with ε-insensitive quadratic loss function is given in Section 2. In the following part, experimental results are displayed. In Section 4, the differences between standard SVR and proposed SVR are discussed. The main conclusion and feature directions for further are summarized in Section 5.

2. Support Vector Regression Method with ε-insensitive Quadratic Loss Function

Support vector networks, a type of universal feed-forward networks, were developed by Vapnik and Cortes [12] at AT & T Bell laboratories to classify data. They are often referred to in the literature as support vector machines (SVMs).

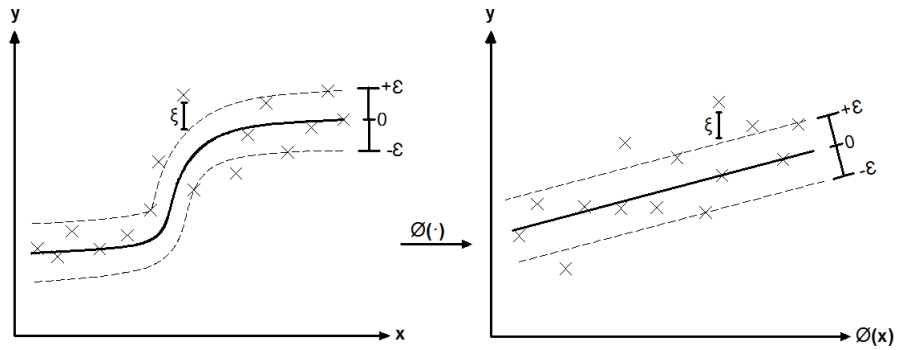


Figure 1. In the SVR method, the training samples represented by nonlinear regression in the input space are expressed by linear regression in the high dimensional feature space using nonlinear basis functions.

The SVM-based model is called SVR for regression [13]. When compared with traditional supervised learning methods of neural networks, SVR uses not only the empirical risk minimization but also the structural risk minimization principle, which aims to reduce the upper limit of the generalization error at the same time. Due to these important properties, SVR has been used in many areas such as biomedical [14-17], renewable energy [18-21], time series forecasting [22-24].

Considering N pairs of training samples $D=\{(\mathbf{x}_s, y_s)|s=1,\cdots N\}$, where $\mathbf{x}_s \in R^n$ is the sth input vector, and $y_s \in R$ is the desired output for the input $\mathbf{x}_s$. The goal in SVR is to find a function that has at most ε deviation from the desired outputs for all training data, and at the same time keep the relationship between $\mathbf{x}_s$ and y_s as flat as possible. Based on the SVR theory, training samples are transformed into high dimensional kernel feature space using nonlinear basis functions $\varphi(\cdot): R^n \to R^m$ and then a linear model is carried out (see Figure 1).

$$f(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x}) + b \qquad (1)$$

where, $\boldsymbol{w} \in R^m$ is the model parameter (weight) vector and b is a threshold to be determined in the function. Flatness in (1) means small $\boldsymbol{w}$. This can be ensured minimizing the Euclidean norm of $\boldsymbol{w}$ i.e. $\|\boldsymbol{w}\|_2^2$. Approximating all pairs with ε precision in (1) means reducing the error between the predicted and desired outputs. This corresponds to minimize the ε-insensitive error function i.e. $L(e_s)_\varepsilon = L(y_s - f(\mathbf{x}_s))_\varepsilon$. Therefore, the SVR optimization problem can be expressed as follows.

$$J(\boldsymbol{w}, b) = \frac{1}{2}\|\boldsymbol{w}\|_2^2 + C\sum_{s=1}^{N} L(y_s - f(\mathbf{x}_s))_\varepsilon \qquad (2)$$

where, $C \in R^+$ is a user defined parameter that determines the tradeoff between the flatness and empirical error.

Standard SVR theory adopts ε-insensitive Laplace (Vapnik) loss function (3), which ignores the errors lower than the predetermined ε value (see Figure 2). Otherwise, they are penalized by $|y_s - f(\mathbf{x}_s)|_\varepsilon$. In other words, small noisy training samples falling into the ε-insensitive area are not included in the solution presentation. Hereby, Vapnik loss based SVR yields a sparse (compressed) model in the obtained solution. The definition of Vapnik loss function is given as follows.

$$|y_s - f(\mathbf{x}_s)|_\varepsilon = \begin{cases} 0 & for\ |y_s - f(\mathbf{x}_s)| < \varepsilon \\ |y_s - f(\mathbf{x}_s)| - \varepsilon & otherwise \end{cases} \qquad (3)$$

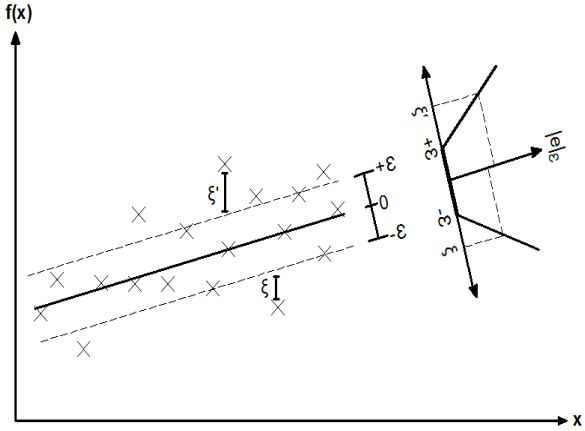


Figure 2. illustrative example of ε-insensitive Laplace (Vapnik) loss function.

From a statistical point of view, Vapnik's loss function is optimal for Laplacian noise distribution. However, many current regression (compression) algorithms consider Gaussian (normal) error distribution which corresponds to a quadratic loss function. In order to deal with Gaussian error distributions, the ε-insensitive quadratic loss function can be written as follows [11]:

$$L(e_s)_e = (e_s)_e^2 = \begin{cases} (e_s - e)^2 & if\ e_s \geq e \\ 0 & if\ -e \leq e_s \leq e \\ (e_s + e)^2 & if\ e_s \leq -e \end{cases} \qquad (4)$$

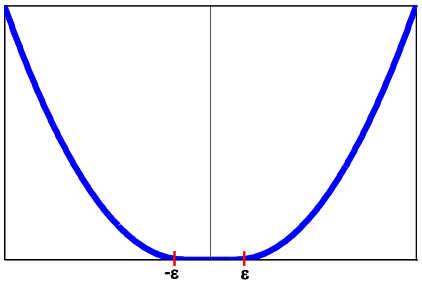


Figure 3. ε insensitive quadratic loss function $(e_s)_e^2$

One can see that the $(e_s)_e^2$ is a convex and continuous differentiable function. By including (4) into (2), the formulation of SVR with ε-insensitive quadratic loss function can be stated as follows.

$$\min_{\boldsymbol{w}\in R^m, b\in R} J(\boldsymbol{w},b)=\frac{1}{2}\|\boldsymbol{w}\|_2^2+C\sum_{s=1}^{N}[(y_s-f(\mathbf{x}_s))]_\varepsilon^2 \tag{5}$$

$$\text{subject to}\begin{cases} y_s-\boldsymbol{w}^T\varphi(\mathbf{x}_s)-b\le\varepsilon+\xi_s \\ -y_s+\boldsymbol{w}^T\varphi(\mathbf{x}_s)+b\le\varepsilon+\xi_s^{'} \\ \xi_s,\xi_s^{'}\ge 0, \qquad s\in\{1,\ldots,N\}. \end{cases} \tag{6}$$

where, ξ_s and $\xi_s^{'}$ are slack variables used to deal with positive and negative deviations outside the ε-insensitive region, respectively (see Figure 3).

In order to solve the primal objective function problem in (5), the linear constraints (6) are multiplied by means of nonnegative Lagrange multipliers for each sample and included in (5).

$$\begin{aligned} \min_{\substack{w\in R^m,b\in R \\ \alpha_s,\alpha_s^{'},\gamma_s,\gamma_s^{'},\xi_s,\xi_s^{'}\ge 0 \\ s\in\{1,\ldots,N\}}} & J(\boldsymbol{w},b,\alpha_s,\alpha_s^{'},\gamma_s,\gamma_s^{'},\xi_s,\xi_s^{'})= \\ & \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}+C\sum_{s=1}^{N}[(\xi_s)^2+(\xi_s^{'})^2] \\ & -\sum_{s=1}^{N}\alpha_s(\varepsilon+\xi_s-y_s+\boldsymbol{w}^T\varphi(\mathbf{x}_s)+b) \\ & -\sum_{s=1}^{N}\alpha_s^{'}(\varepsilon+\xi_s^{'}+y_s-\boldsymbol{w}^T\varphi(\mathbf{x}_s)-b) \\ & -\sum_{s=1}^{N}(\gamma_s\xi_s+\gamma_s^{'}\xi_s^{'}) \end{aligned} \tag{7}$$

where the variables α_s, $\alpha_s^{'}$ and γ_s, $\gamma_s^{'}$ are Lagrange multipliers (dual variables). For optimal solution to the Lagrangian problem (7), primal variables $(\boldsymbol{w},b,\xi_s,\xi_s^{'})$ must vanish. Therefore, the partial derivative of Lagrangian function $J(\boldsymbol{w},b,\alpha_s,\alpha_s^{'},\gamma_s,\gamma_s^{'},\xi_s,\xi_s^{'})$ is taken with respect to the input variables and the result is equal to zero.

$$\frac{\partial J}{\partial b}=\sum_{s=1}^{N}(\alpha_s^{'}-\alpha_s)=0 \tag{8}$$

$$\nabla_{\boldsymbol{w}}J=\boldsymbol{w}-\sum_{s=1}^{N}(\alpha_s-\alpha_s^{'})\varphi(\mathbf{x}_s)=0 \tag{9}$$

$$\frac{\partial J}{\partial \xi_s}=C(2\xi_s)-\alpha_s-\gamma_s=0 \tag{10}$$

$$\frac{\partial J}{\partial \xi_s^{'}}=C(2\xi_s^{'})-\alpha_s^{'}-\gamma_s^{'}=0 \tag{11}$$

If the equations (8) through (11) are substituted in the (7), the dual optimization problem is obtained.

$$\begin{aligned} \max_{\boldsymbol{\alpha}\in R^N} J(\alpha_s,\alpha_s^{'})=&-\frac{1}{2}\sum_{s=1}^{N}\sum_{r=1}^{N}(\alpha_s-\alpha_s^{'})\mathrm{K}(\alpha_r-\alpha_r^{'}) \\ &-\varepsilon\sum_{s=1}^{N}(\alpha_s+\alpha_s^{'})+\sum_{s=1}^{N}y_s(\alpha_s-\alpha_s^{'}) \\ &-\frac{1}{2C}\sum_{s=1}^{N}[(\alpha_s)^2+(\alpha_s^{'2})] \end{aligned} \tag{12}$$

$$\text{Subject to}\sum_{s=0}^{N}(\alpha_s^{'}-\alpha_s)=0 \text{ and } \alpha_s^{'},\alpha_s\in[0,\infty] \tag{13}$$

where, K represents the kernel matrix. The entries of kernel matrix are the kernel functions $\mathrm{K}(\mathbf{x}_s,\mathbf{x}_r)$ that are defined as the inner product of two samples $\varphi(\mathbf{x}_s)$ and $\varphi(\mathbf{x}_r)$.

$$\begin{aligned} \mathrm{K}=[\mathrm{K}(\mathbf{x}_s,\mathbf{x}_r)]_{s,r}&=[\varphi^{\mathrm{T}}(\mathbf{x}_s)\cdot\varphi(\mathbf{x}_r)]_{s,r} \\ &=\begin{bmatrix} \mathrm{K}(\mathbf{x}_1,\mathbf{x}_1) & \mathrm{K}(\mathbf{x}_1,\mathbf{x}_2) & \cdots & \mathrm{K}(\mathbf{x}_1,\mathbf{x}_N) \\ \mathrm{K}(\mathbf{x}_2,\mathbf{x}_1) & \mathrm{K}(\mathbf{x}_2,\mathbf{x}_2) & \cdots & \mathrm{K}(\mathbf{x}_2,\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{K}(\mathbf{x}_N,\mathbf{x}_1) & \mathrm{K}(\mathbf{x}_N,\mathbf{x}_2) & \cdots & \mathrm{K}(\mathbf{x}_N,\mathbf{x}_N) \end{bmatrix} \end{aligned} \tag{14}$$

The optimization problem in (12) constitutes a quadratic programming problem whose solution gives a unique minimum. After computation of Lagrange multipliers (support vectors) α_s and $\alpha_s^{'}$, the optimal model parameter $\boldsymbol{w}$ (9) can be written as follow.

$$\boldsymbol{w}=\sum_{\mathbf{x}_s\in DV}(\alpha_s-\alpha_s^{'})\varphi(\mathbf{x}_s) \tag{15}$$

The decision function (1) for the test sample **x** can be written as

$$f(\mathbf{x}) = \sum_{\mathbf{x}_s \in SV} (\alpha_s - \alpha_s^{'}) \mathrm{K}(\mathbf{x}_s, \mathbf{x}) + b \qquad (16)$$

where, SV are the training samples corresponding to $\alpha_s - \alpha_s^{'} \neq 0$ (called a support vector). Note that when evaluating $f(\mathbf{x})$, $\boldsymbol{w}$ does not need to be computed explicitly. From (16), the operations required for the SVR can be performed directly in the input space with the kernel function, without performing the process of moving the training samples from the input space to the high-dimensional space by means of non-linear functions. This is called the "kernel trick" in the literature [13] and reduces the computation time required to solve the optimization problem.

3. Experimental Results

This section presents various experimental results for the ε-insensitive quadratic loss function proposed in the SVR framework. Moreover, the performance of the proposed SVR model is compared with the standard SVR model. To provide same conditions for both models, all experiments were performed in Matlab 2016a environment installed on a personal computer with Intel Core I5 processor 3.1 GHz, 10 GB RAM and 64 bit Windows 10 operating system. Gauss function $\mathrm{K}(\mathbf{x}_s, \mathbf{x}) = \exp(-\|\mathbf{x} - \mathbf{x}_s\|_2^2 / 2\sigma^2)$ which is widely used in literature was chosen as kernel function.

The user-defined optimal parameters (C^*, σ^*) of the SVR technique were determined from the sets {0.1, 0.5, 1, 2, 4, 8, 16, 24, 48} and {0.001, 0.0011, 0.0012, 0.0013, 0.0014} respectively. In order to assess performance, the following accuracy evaluation parameters [2-4] are used in this study.

Compression ratio (CR):

$$CR = \frac{\#TSOS}{\#SVCS} \qquad (17)$$

where, *#TSOS* denotes the number of *training samples* required for the *original signal* and *#SVCS* specifies the number of support vectors needed for the *compressed signal*.

Percent Root Mean Square Difference (PRD):

$$PRD = \sqrt{\frac{\sum_{s=1}^{N} [(y_s - f(\mathbf{x}_s)]^2}{\sum_{s=1}^{N} [(y_s)]^2}} \times 100 \qquad (18)$$

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{s=1}^{N} (y_s - f(\mathbf{x}_s))^2} \qquad (19)$$

ECG records (360Hz with 11 bits/sample resolution) have been obtained from the MIT-BIH database [25]. A normal sinus rhythm (1 period ECG signal) was extracted from the ECG recordings for visualization and compression was performed on this signal (see Figure 4).

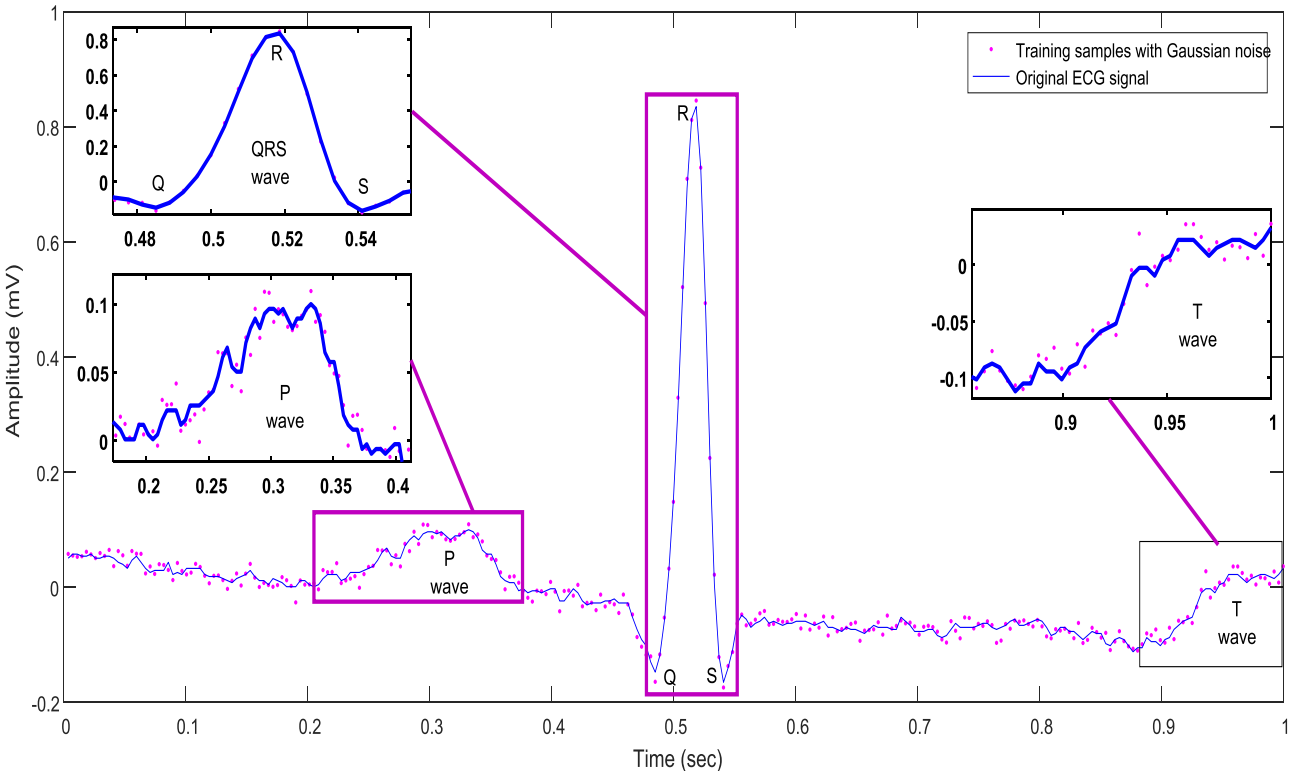


Figure 4. Original ECG (dark blue) signal, the Gaussian noise (magenta spots) added ECG signal and the P, QRS and T wave forms (dark violet rectangles) that make up the original ECG signal

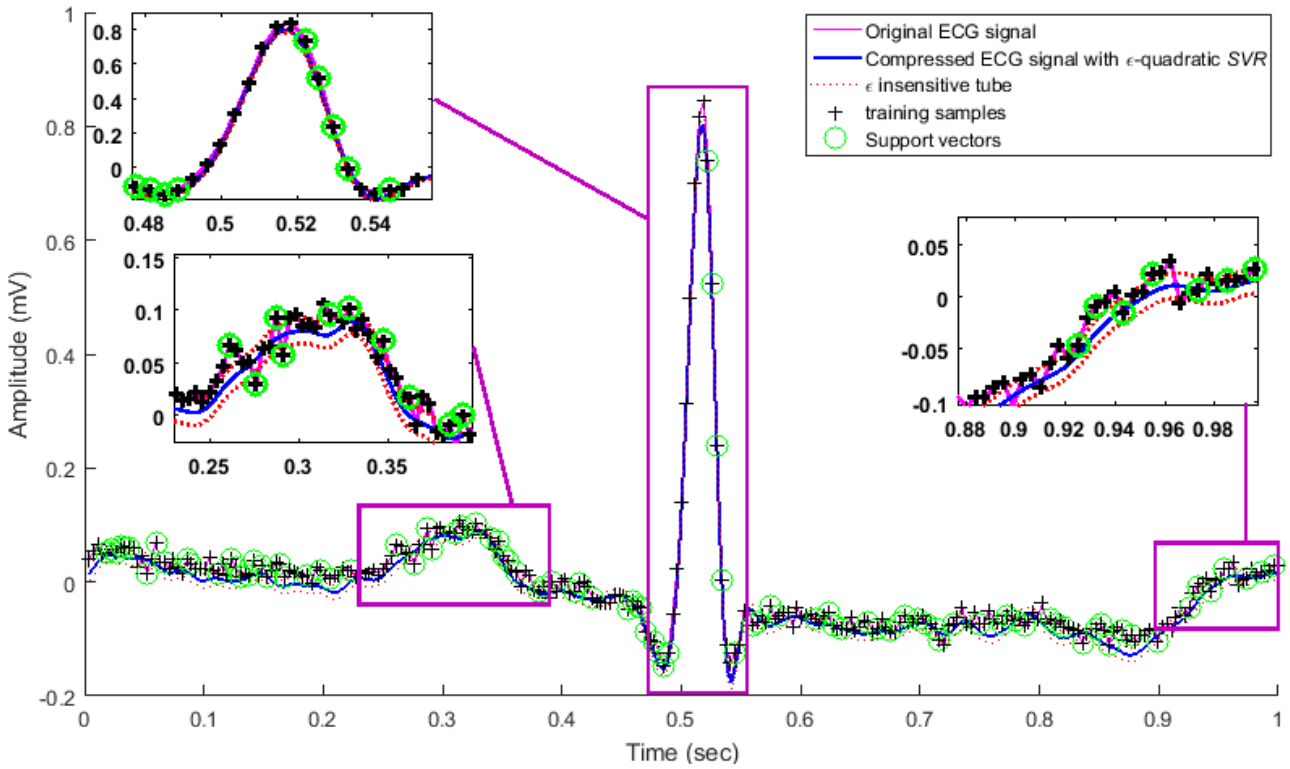


Figure 5. ECG (dark blue) signal compressed by ε-insensitive quadratic loss based SVR. The areas separated by the red dashed line below and above the compressed ECG signal show the ε-insensitive area. The '+' signs in the black color represent the training samples, and the '+' signs in the green ring represent the support vectors (the samples that yield the compressed ECG signal).

Table 1. Experimental results on ECG data compressed by ε insensitive quadratic loss based SVR.

Model Parameters		#TS	#SV	***w***	RMSE	PRD	CR
ε=0.010	C=30 σ=0.012	268	107	0,97	0,014	11,69	2,50
ε=0.012			83	0,96	0,016	12,77	3,22
ε=0.015			52	0,95	0,019	15,42	5,15
ε=0.020			29	0,93	0,024	19,68	9,24
ε=0.030			22	0,87	0,036	29,05	12,18
ε=0.040			17	0,78	0,048	39,29	15,76
ε=0.050			12	0,76	0,059	48,15	22,33

In order to show that the proposed method is robust against Gaussian distribution, the original ECG signal is polluted by Gaussian noise $\left(1/\sqrt{2\pi}\right) exp(-(x-\mu)^2/2\tau^2)$ with the mean μ=0 and the variance τ=0.01 (Figure 4 magenta spots).

Figure 5 demonstrates the plot of the ECG signal with Gaussian noise (magenta) and its reconstructed version (dark blue) using ε-insensitive quadratic loss function (with the parameters ε=0.02, C = 30, and σ = 0.012). As can be seen from Figure 5 that the training examples (185 black '+' signs) in the ε-insensitive region are not included in the solution representation of the SVR optimization algorithm. Since training samples (83 support vectors, black '+' signs in green circles) outside the ε-insensitive region are only included in the solution representation, the ECG signal reconstructed by the ε-insensitive quadratic loss function contains fewer training samples than the original signal. Therefore, SVR with ε-insensitive quadratic loss function gives a sparse (compressed) solution. The parameter ε allows us to control the selection of the support vectors which may be directly related to the sparseness. For different ε values, the numerical analysis in terms of the number of training samples (#TS), the number of support vectors (#SV), the flatness (***w***) estimated by (15), root mean square error (RMSE) estimated by (19), percent root mean square difference (PRD) calculated by (18), and compression ratio (CR) calculated by (17) are provided in Table 1.

As the value of the sparsity (compression) parameter ε is increases, the CR value shown in Table 1 also increases. However, this increase also causes the PRD and RMSE values to increase, which leads to distortions in the compressed signal.

It is obviously seen from Table 2 that the proposed SVR (ε-Quadratic loss) and standard SVR models with the same parameter (ε=0.02 and σ=0.012) values show different behaviors in various C parameters.

Table 2. Comparing the performance of the standard SVR (Vapnik loss) and the proposed SVR (ε-Quadratic loss) in compressing ECG data

Model parameters		#TS	Loss function	#SV	$\boldsymbol{w}$	RMSE	PRD	CR
C=0.5	ε=0.02 σ=0.012	268	ε-Quadratic	155	0,47	0,039	31,95	1,72
			Vapnik	82	0,95	0,024	19,59	3,26
C=1			ε-Quadratic	116	0,60	0,032	26,48	2,31
			Vapnik	71	0,97	0,024	19,46	3,77
C=5			ε-Quadratic	64	0,79	0,026	20,80	4,18
			Vapnik	58	0,98	0,024	19,59	4,62
C=10			ε-Quadratic	50	0,82	0,025	20,28	5,36
			Vapnik	53	0,98	0,024	19,41	5,05
C=20			ε-Quadratic	39	0,89	0,024	19,56	6,87
			Vapnik	46	0,99	0,024	19,53	5,82
C=30			ε-Quadratic	28	0,91	0,024	19,38	9,57
			Vapnik	33	1,02	0,024	19,28	8,12
C=50			ε-Quadratic	18	0,93	0,024	19,49	14,88
			Vapnik	25	1,03	0,024	19,78	10,72

When compared, the proposed SVR model performs better than the standard SVR model as the value of C increases. In other words, in the same RMSE and PRD, when the C value is 10 and above, the proposed SVR model provides a better CR than the standard SVR model, while at the same time it produces a smoother function (small $\boldsymbol{w}$). For example, the proposed SVR model yields SV=18, CR=9.57, and $\boldsymbol{w}$=0.93 while the standard SVR model yields SV=25, CR=8.12, and $\boldsymbol{w}$=1.03 for the same values of parameters (ε=0.02, σ=0.012, C=30, PRD=19,38, and RMSE=0,024). On the other hand, standard SVR performs better at small C values. This is due to the emphasis on the smoothness of the model rather than the empirical error in the optimization problem.

It can be concluded from these results that the proposed SVR model is more efficient than the standard SVR model if the data is contaminated with Gaussian noises, since the quadratic loss function is optimal for Gaussian noise.

4. Discussion

Long-term ECG recordings produce large amounts of data at a level that makes storage and transmission difficult. Many algorithms have recently been suggested for compressing, modeling and reconstructing ECG signals. Recently, support vector algorithm has received great interest in regression (compression) problems because of the attractive properties such as better generalization ability independent of dimensionality of problems and ensured global minimum solution. But, the performance of SVR usually depends on the selected loss function. As Table 2 shows, the proposed SVR model (CR=14,88) provides a better compression ratio than the standard SVR model (CR=10,72) for the same values of parameter (ε=0.02, σ=0.012, C=50, and RMSE=0,024). Most of the compression (regression) algorithms consider Gaussian error distribution which corresponds to a typical quadratic loss function. For this reason, standard SVR is not suitable for applications with Gaussian error distribution.

5. Conclusion and Future Work

In this study, the ε-insensitive quadratic loss based SVR technique is proposed to cope with Gaussian error distribution in ECG signals. The ε-insensitive quadratic loss function ignores the small noise and provides sparseness (compression) in solution representation.

For various epsilon (ε) values, computer simulations show that the proposed ε insensitive quadratic loss based SVR technique optimally yields good compression ratio.

Computational comparisons between the proposed SVR model and standard SVR model demonstrate the effectiveness of the proposed model especially large C values. For the same RMSE (0,024), PRD (19,38), C (30) ε (0.02), σ (0.012) values, the proposed SVR model provides SV=18, CR=9.57, and $\boldsymbol{w}$=0.93 while the standard SVR model yields SV=25, CR=8.12, and $\boldsymbol{w}$=1.03.

As a future work, from a theoretical point of view, the proposed technique can be effectively applied to larger ECG data with sequential minimal optimization algorithm. Moreover, since it does not need any additional algorithm (pre-processing of the ECG signal), it can be applied to other biomedical signals such as Elektroensefalografi (EEG), Elektromiyografi (EMG).

References

[1] Wei, J. J., Chang, C. J., Chou, N. K., Jan, G. J. 2001. ECG data compression using truncated singular value decomposition. IEEE Transactions on Information Technology in Biomedicine, 5(2001), 290-299.

[2] Jalaleddine, S. M., Hutchens, C. G., Strattan, R. D., Coberly, W. A. 1990. ECG data compression techniques-a unified approach. IEEE transactions on Biomedical Engineering, 37(1990), 329-343.

[3] Manikandan, M. S., Dandapat, S. 2014. Wavelet-based electrocardiogram signal compression methods and their performances: a prospective review. Biomedical Signal Processing and Control, 14(2014), 73-107.

[4] Singh, B., Kaur, A., Singh, J. 2015. A review of ecg data compression techniques. International journal of computer applications, 116(2015), 39-44.

[5] Olmos, S., MillAn, M., Garcia, J., Laguna, P. 1996. ECG data compression with the Karhunen-Loeve transform. In Computers in Cardiology, Indianapolis, 8-11 September, USA, 253-256.

[6] Reddy, B. S., Murthy, I. S. N. 1986. ECG data compression using Fourier descriptors. IEEE Transactions on Biomedical Engineering, 4(1986), 428-434.

[7] Benzid, R., Messaoudi, A., Boussaad, A. 2008. Constrained ECG compression algorithm using the block-based discrete cosine transform. Digital Signal Processing, 18(2008), 56-64.

[8] Shinde, A. A., Kanjalkar, P. 2011. The comparison of different transform based methods for ECG data compression, Uluslararası konferans, ICSCCN-IEEE, 21-22 June, Thuckafay, India, 332-335.

[9] Addison, P. S. 2005. Wavelet transforms and the ECG: a review. Physiological measurement, 26(2005), R155.

[10] Karal, O. 2018. Destek Vektör Regresyon ile EKG Verilerinin Sıkıştırılması, Journal of the Faculty of Engineering and Architecture of Gazi University (In press).

[11] Karal, O. 2017. Maximum likelihood optimal and robust Support Vector Regression with lncosh loss function. Neural Networks, 94(2017), 1-12.

[12] Cortes, C., Vapnik, V. 1995. Support-vector networks, Mach. Learn., 20(3), 273-297, 1995.

[13] Smola, A. J., Schölkopf, B. 2004. A tutorial on support vector regression. Statistics and computing, 14(2004), 199-222.

[14] Mahmoodian, H., Ebrahimian, L. 2016. Using support vector regression in gene selection and fuzzy rule generation for relapse time prediction of breast cancer. Biocybernetics and Biomedical Engineering, 36(2016), 466-472.

[15] Valizadeh, M., Sohrabi, M. R. 2018. The application of artificial neural networks and support vector regression for simultaneous spectrophotometric determination of commercial eye drop contents. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 193(2018), 297-304.

[16] Huber, M. B., Lancianese, S. L., Nagarajan, M. B., Ikpot, I. Z., Lerner, A. L., Wismuller, A. 2011. Prediction of biomechanical properties of trabecular bone in MR images with geometric features and support vector regression. IEEE Transactions on Biomedical Engineering, 58(2011), 1820-1826.

[17] Guermoui, M., Mekhalfi, M. L., Ferroudji, K. 2013. Heart sounds analysis using wavelets responses and support vector machines. In Systems, Signal Processing and their Applications (WoSSPA), IEEE, May 2013, 8th International Workshop on, 233-238.

[18] Hu, Q., Zhang, S., Yu, M., Xie, Z. 2016. Short-term wind speed or power forecasting with heteroscedastic support vector regression. IEEE Trans. Sustainable Energy, 7(2016), 241-249.

[19] Khosravi, A., Koury, R. N. N., Machado, L., Pabon, J. J. G. 2018. Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system. Sustainable Energy Technologies and Assessments, 25(2018), 146-160.

[20] Guermoui, M., Rabehi, A., Gairaa, K., Benkaciali, S. 2018. Support vector regression methodology for estimating global solar radiation in Algeria. The European Physical Journal Plus, 133(2018), 22.

[21] Das, U. K., Tey, K. S., Seyedmahmoudian, M., Mekhilef, S., Idris, M. Y. I., Van Deventer, W., Stojcevski, A. 2018. Forecasting of photovoltaic power generation and model optimization: A review. Renewable and Sustainable Energy Reviews, 81(2018), 912-928.

[22] Yaslan, Y., Bican, B. 2017. Empirical mode decomposition based denoising method with support vector regression for time series prediction: a case study for electricity load forecasting. Measurement, 103(2017), 52-61.

[23] Nava, N., Matteo, T. D., Aste, T. 2018. Financial Time Series Forecasting Using Empirical Mode Decomposition and Support Vector Regression. Risks, 6(2018), 7.

[24] Massana, J., Pous, C., Burgas, L., Melendez, J., Colomer, J. 2016. Short-term load forecasting for non-residential buildings contrasting artificial occupancy attributes. Energy and Buildings, 130(2016), 519-531.

[25] Moody, G. B., Mark, R. G. 2001. The impact of the MIT-BIH arrhythmia database. IEEE Engineering in Medicine and Biology Magazine, 20(2001), 45-50.