Expert Systems with Applications 37 (2010) 233–239

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals

Elif Derya Übeyli̇ *

Department of Electrical and Electronics Engineering, Faculty of Engineering, TOBB Ekonomi ve Teknoloji Üniversitesi, 06530 Söğütözü, Ankara, Turkey

ARTICLE INFO

Keywords:
Least squares support vector machines
Model-based methods
Electroencephalogram (EEG) signals

ABSTRACT

The aim of the study is classification of the electroencephalogram (EEG) signals by combination of the model-based methods and the least squares support vector machines (LS-SVMs). The LS-SVMs were implemented for classification of two types of EEG signals (set A – EEG signals recorded from healthy volunteers with eyes open and set E – EEG signals recorded from epilepsy patients during epileptic seizures). In order to extract the features representing the EEG signals, the spectral analysis of the EEG signals was performed by using the three model-based methods (Burg autoregressive – AR, moving average – MA, least squares modified Yule–Walker autoregressive moving average – ARMA methods). The present research demonstrated that the Burg AR coefficients are the features which well represent the EEG signals and the LS-SVM trained on these features achieved high classification accuracies.

1. Introduction

The electroencephalogram (EEG), a highly complex signal, is one of the most common sources of information used to study brain function and neurological disorders (Agarwal, Gotman, Flanagan, & Rosenblatt, 1998; Adeli, Zhou, & Dadmehr, 2003; Hazarika, Chen, Tsoi, & Sergejew, 1997). Large amounts of data are generated by EEG monitoring systems for electroencephalographic changes, and their complete visual analysis is not routinely possible. Computers have long been proposed to solve this problem and thus, automated systems to recognize electroencephalographic changes have been under study for several years. There is a strong demand for the development of such automated devices, due to the increased use of prolonged and long-term video EEG recordings for proper evaluation and treatment of neurological diseases and prevention of the possibility of the analyst missing (or misreading) information (Agarwal et al., 1998; Adeli et al., 2003; Hazarika et al., 1997).

Support vector machines (SVMs) proposed by Vapnik (1995) are trained by solving a quadratic optimization problem. Least squares support vector machines (LS-SVMs) proposed by Suykens and Vandewalle (1999) are trained by solving a set of linear equations. In the present study, the LS-SVMs were implemented for classification of two types of EEG signals (set A – EEG signals recorded from healthy volunteers with eyes open and set E – EEG signals recorded from epilepsy patients during epileptic seizures) (Andrzejak et al., 2001; http://www.meb.uni-bonn.de/epileptologie/science/physik/eegdata.html). Feature extraction/selection plays an important role in classifying systems such as neural networks. Therefore, in order to extract the features representing the EEG signals, the spectral analysis of the EEG signals was performed by using the three model-based methods (Burg autoregressive – AR, moving average – MA, least squares modified Yule–Walker autoregressive moving average – ARMA methods). The model-based methods (parametric) are based on modeling the data sequence $x(n)$ as the output of a linear system characterized by a rational system. In the model-based methods, the spectrum estimation procedure consists of two steps. Given the data sequence $x(n), 0 \leqslant n \leqslant N-1$, the parameters of the method are estimated. Then from these estimates, the power spectral density (PSD) estimate is computed. The model-based methods spectra have a good statistical stability for short segments of signal and have a good spectral resolution and the resolution is less dependent on the length of the record (Kay & Marple, 1981; Kay, 1988; Proakis & Manolakis, 1996; Stoica & Moses, 1997). The implemented LS-SVM was trained on the Burg AR coefficients and the high accuracy was achieved in classifying the EEG signals (sets A and E).

The outline of this study is as follows. In Section 2, brief description of data is presented. In Section 3, brief review of classifiers (SVMs and LS-SVMs) is given with the related references for further reading. In Section 4, computation of the coefficients of the model-based methods and the results of application of the LS-SVM trained on the Burg AR coefficients to the EEG signals are presented. Discussion of the presented results is performed in the light of existing studies in the literature. Finally, in Section 5 the drawn conclusions are emphasized.

* Tel.: +90 312 2924080; fax: +90 312 2924091.
E-mail address: edubeyli@etu.edu.tr

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.05.012

2. Description of studied EEG signals

The data described in reference Andrzejak et al. (2001), which is publicly available (http://www.meb.uni-bonn.de/epileptologie/science/physik/eegdata.html) was used. In this section, only a short description is presented and refer to reference Andrzejak et al. (2001) for further details. The complete dataset consists of five sets (denoted A–E), each containing 100 single-channel EEG signals of 23.6 s. Each signal has been selected after visual inspection for artifacts and has passed a weak stationarity criterion. Sets A and B have been taken from surface EEG recordings of five healthy volunteers with eyes open and closed, respectively. Signals in two sets have been measured in seizure-free intervals from five patients in the epileptogenic zone (D) and from the hippocampal formation of the opposite hemisphere of the brain (C). Set E contains seizure activity, selected from all recording sites exhibiting ictal activity. Sets A and B have been recorded extracranially, whereas sets C, D, and E have been recorded intracranially. In the applications, performance degraded for a more detailed classification which further dissociated between sets A (healthy volunteer, eyes open) and B (healthy volunteer, eyes closed), and sets D (epileptogenic zone) and C (hippocampal formation of opposite hemisphere). Therefore, in the present study two dataset (A and E) of the complete dataset were classified. The implemented LS-SVM was formulated for two-class classification problem. The waveforms of the two types of EEG segments analyzed in the present study are shown in Fig. 1a and b.

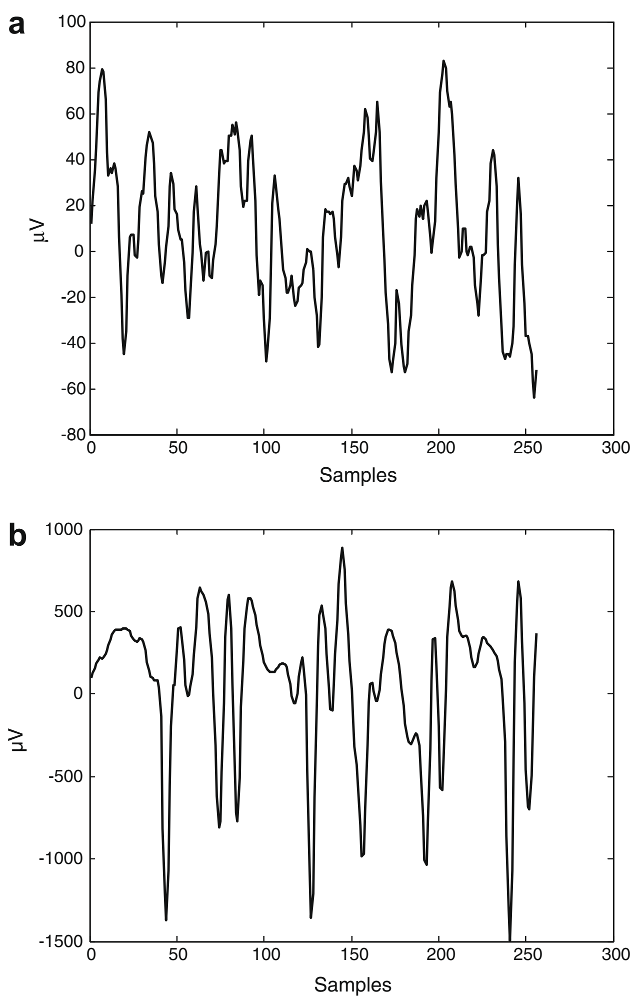

Fig. 1. Waveforms of the EEG segments (a) set A, (b) set E.

3. Brief review of classifiers

3.1. Support vector machines (SVMs)

The SVM proposed by Vapnik (1995) has been studied extensively for classification, regression and density estimation. Fig. 2 shows the architecture of the SVM. SVM maps the input patterns into a higher dimensional feature space through some nonlinear mapping chosen a priori. A linear decision surface is then constructed in this high dimensional feature space. Thus, SVM is a linear classifier in the parameter space, but it becomes a nonlinear classifier as a result of the nonlinear mapping of the space of the input patterns into the high dimensional feature space. Let m-dimensional training data be $\mathbf{x}_i (i = 1, \ldots, M)$ and their class labels be y_i, where $y_i = 1$ and $y_i = -1$ for classes 1 and 2, respectively. If these input data are linearly separable in the feature space, then the following decision function can be determined:

$$D(\mathbf{x}) = \mathbf{w}^t \mathbf{g}(\mathbf{x}) + b \tag{1}$$

where $\mathbf{g}(\mathbf{x})$ is a mapping function that maps $\mathbf{x}$ into the l-dimensional space, $\mathbf{w}$ is the l-dimensional vector and b is a scalar. To separate data linearly, the decision function satisfies the following condition:

$$y_i(\mathbf{w}^t \mathbf{g}(\mathbf{x}_i) + b) \geqslant 1 \text{ for } i = 1, \ldots, M \tag{2}$$

If the problem is linearly separable in the feature space, there are an infinite number of decision functions that satisfy Eq. (2). among them we require that the hyperplane has the largest margin between two classes. The margin is the minimum distance from the separating hyperplane to the input data and this is given by $|D(\mathbf{x})|/\|\mathbf{w}\|$. Then we call the separating hyperplane with the maximum margin optimal separating hyperplane.

Assuming that the margin is ρ, the following condition needs to be satisfied:

$$\frac{y_i D(\mathbf{x}_i)}{\|\mathbf{w}\|} \geqslant \rho \text{ for } i = 1, \ldots, M \tag{3}$$

The product of ρ and $\|\mathbf{w}\|$ is fixed:

$$\rho \|\mathbf{w}\| = 1 \tag{4}$$

In order to obtain the optimal separating hyperplane with the maximum margin, $\mathbf{w}$ with the minimum $\|\mathbf{w}\|$ that satisfying Eq. (3) should be found. From Eq. (4), this leads to solving the following optimization problem. Minimizing

$$\frac{1}{2} \mathbf{w}^t \mathbf{w} \tag{5}$$

subject to the constraints:

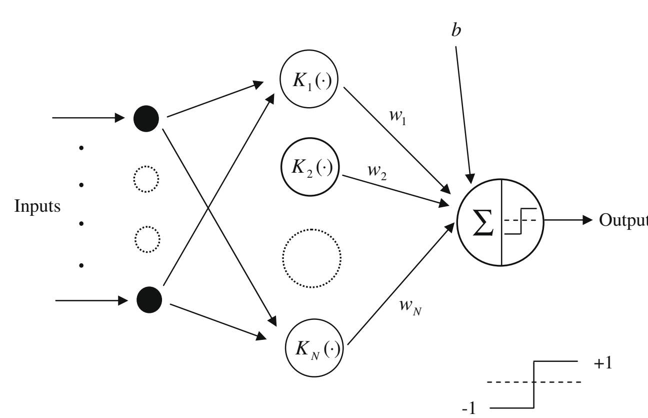

Fig. 2. Architecture of the SVM (N is the number of support vectors).