Expert Systems with Applications 38 (2011) 4052–4057

Contents lists available at ScienceDirect

Expert Systems with Applications

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

# Short term wind speed prediction based on evolutionary support vector regression algorithms

Sancho Salcedo-Sanz [a,*], Emilio G. Ortiz-García [a], Ángel M. Pérez-Bellido [a], Antonio Portilla-Figueras [a], Luis Prieto [b]

[a] Department of Signal Theory and Communications, Universidad de Alcalá, Madrid, Spain
[b] Energy Resource Department, Iberdrola Renovables, Madrid, Spain

ARTICLE INFO

*Keywords:*
Support vector regression algorithms
Evolutionary algorithms
Hyper-parameters estimation
Short-term wind speed forecasting

ABSTRACT

Hyper-parameters estimation in regression Support Vector Machines (SVMr) is one of the main problems in the application of this type of algorithms to learning problems. This is a hot topic in which very recent approaches have shown very good results in different applications in fields such as bio-medicine, manufacturing, control, etc. Different evolutionary approaches have been tested to be hybridized with SVMr, though the most used are evolutionary approaches for continuous problems, such as evolutionary strategies or particle swarm optimization algorithms. In this paper we discuss the application of two different evolutionary computation techniques to tackle the hyper-parameters estimation problem in SVMrs. Specifically we test an Evolutionary Programming algorithm (EP) and a Particle Swarm Optimization approach (PSO). We focus the paper on the discussion of the application of the complete evolutionary-SVMr algorithm to a real problem of wind speed prediction in wind turbines of a Spanish wind farm.



## 1. Introduction

The Support Vector Machine (SVM) (Vapnik, 1998) is a powerful and robust methodology in statistical machine learning, that has been successfully applied to regression problems (SVMr) (Akay, 2009; Cherkassky & Ma, 2004; He, Wang, & Jiang, 2008; Lázaro, Santamaría, Pérez-Cruz, & Artés-Rodrıguez, 2005; Wu, Chau, & Li, 2008), including problems of wind speed prediction (Mohandes, Halawani, Rehman, & Hussain, 2004; Salcedo-Sanz et al., 2009). The SVMr is considered a good methodology because it allows the use of the kernel theory to increase the quality of regression models and also because, in most cases, it can be solved as a convex optimization problem. Several fast algorithms there exist to carry out the SVMr training, such as the sequential minimal optimization algorithm (Smola & Schölkopf, 1998). In spite of this, the SVMr performance heavily depends on the choice of several hyper-parameters, necessary to define the optimization problem and the final SVMr model. Unfortunately, there is not an exact method to obtain the optimal set of SVMr hyper-parameters, so that a search algorithm must be applied to obtain the best possible set of hyper-parameters which guarantees the maximum possible quality of the final model.

In general, the search algorithms used to obtain SVMr hyper-parameters can be divided in three groups. The first group of algorithms for SVMr hyper-parameters search is based on *grid searches* (Akay, 2009; Mohandes et al., 2004), where the search space of parameters is divided into groups of possible parameters to be tested, usually in an uniform fashion. The second group of search algorithms is formed by local search type approaches, such as *pattern search* proposed in Momma and Bennett (2002). Finally, the third group of algorithms applied to obtain SVMr hyper-parameters is based on metaheuristics, or global optimization algorithms, such as evolutionary computation (Hou & Li, 2009; Wang, Yang, Qin, & Gui, 2005; Wu, Tzeng, & Lin, 2009). This latter set of methodologies include several approaches, such as genetic algorithms, evolutionary algorithms, particle swarm or differential evolution, etc., which are applied to implement a robust research on the hyper-parameters search space. This work is focused on this type of methodologies, because they have shown very good performance in previous applications. For example Friedrichs and Igel (2005) was one of the first approaches where the evolutionary optimization of SVM parameters was proposed. More recently, similar approaches have also been successfully applied to different problems, as in Hou and Li (2009), where an evolutionary strategy algorithm has been applied to obtain the SVM parameters in a problem of short-term fault prediction, or in Cheng and Wu (2009), where an evolutionary SVM approach has been proposed to a problem of construction management. Novel evolutionary computation algorithms have also been tested, for example a quan-

* Corresponding author. Address: Departamento de Teoría de la Señal y Comunicaciones, Universidad de Alcalá de Henares, Campus Universitario, 28871 Alcalá de Henares, Madrid, Spain. Tel.: +34 91 885 6731; fax: +34 91 624 8749.
*E-mail address:* sancho.salcedo@uah.es (S. Salcedo-Sanz).

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.09.067

tum-inspired evolutionary approach has been proposed in Luo et al. (2008), and a novel hybrid genetic algorithm for both selecting the optimal kernel function and optimizing its corresponding hyper-parameters has been proposed in Wu et al. (2009). A hybrid approach involving a genetic algorithm for selecting SVM hyper-parameters has been recently applied to a problem of mitochondrial toxicity prediction (Zhang et al., 2009). In Wu, Tzeng, Goo, and Fang (2007) a real-coded genetic algorithm has been successfully applied to a problem of bankruptcy prediction. Also, different approaches based on Particle Swarm Optimization (PSO) algorithms can be found in the literature for the SVM hyper-parameters tuning. In Lin, Ying, Chen, and Lee (2008) a PSO has been successfully applied to the optimization of the SVM hyper-parameters together with a feature selection problem. A similar approach also including feature selection has been presented in Huang and Dun (2008). Novel PSO approaches have been proposed and tested in specific applications within the frame of SVM hyper-parameters tuning. For example, in Wang et al. (2009), an immune PSO approach is proposed to optimize the SVM parameters in an interesting problem of forest fire prediction, and in Wu (2010) and PSO with Gaussian mutation hybridized with a SVM is applied to a problem of power load forecast. It is easy to see that the evolutionary-based optimization of SVM hyper-parameters is really a hot topic, not only because the large amount of works applying these techniques, but also because the majority of works are very recent.

In this paper we deal with evolutionary-based approaches for SVMr hyper-parameters tuning in a real problem of wind speed prediction. We compare the performance of an Evolutionary Programming algorithm (Yao, Liu, & Lin, 1999) and a PSO approach (Eberhart & Shi, 2001) for SVMr parameter tuning, and how the resulting algorithm can be integrated in a complete system for wind speed prediction in wind farms. The specific model for SVMr parameters estimation described in the paper starts from a coarse grid search to reduce the SVMr hyper-parameters search space size. Then, in a second step, the evolutionary-based algorithms are applied to refine the search, improving the performance of the SVMr. Note that the regression SVMr algorithm has been previously applied to a problem of wind speed in Mohandes et al. (2004). In that paper, the authors successfully applied the SVMr standard algorithm with grid search to obtain the best set of kernel parameters. In this paper we show that the evolutionary estimation of hyper-parameters performs better than the grid search. The complete Evolutionary SVMr algorithm has been tested integrated into the forecasting model presented in Salcedo-Sanz et al. (2009), in a wind farm located at the south of Spain.

The rest of this paper is structured as follows: next section describes the main characteristics of the regression SVMr algorithm used in this paper. Section 3 presents the SVMr hyper-parameters model considered in this paper, consisting of a first coarse grid search, that will be refined using the evolutionary computation algorithms. Section 4 presents the basic prediction system where the proposed evolutionary SVMr approach will be integrated, and shows the good performance of this model in a problem of wind speed forecast in several turbines of a Spanish wind farm. Section 5 closes the paper with some final considerations and remarks.

## 2. $\epsilon$-SVMr formulation

The $\epsilon$-SVMr method for regression (Smola, Murata, Scholkopf, & Muller, 1998) consists of, given a set of training vectors $C = \{(\mathbf{x_i}, y_i), i = 1,\ldots,l\}$, obtaining a model of the form $y(\mathbf{x}) = f(\mathbf{x}) + b = \mathbf{w}^T\phi(\mathbf{x}) + b$, to minimize the following general risk function:

$$R[f] = \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{2}C\sum_{i=1}^{l} L(y_i, f(\mathbf{x})), \tag{1}$$

where $\mathbf{w}$ controls the smoothness of the model, $\phi(\mathbf{x})$ is a function of projection of the input space to the feature space, $b$ is a parameter of bias, $\mathbf{x_i}$ is a feature vector of the input space with dimension $N$, $y_i$ is the output value to be estimated and $L(y_i,f(\mathbf{x}))$ is the loss function selected. In this paper, we use the L1-SVMr (L1 support vector regression), characterized by an $\epsilon$-insensitive loss function (Smola & Schölkopf, 1998)

$$L(y_i, f(\mathbf{x})) = |y_i - f(\mathbf{x_i})|_\epsilon. \tag{2}$$

In order to train this model, it is necessary to solve the following optimization problem (Smola & Schölkopf, 1998):

$$\min\left(\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*)\right) \tag{3}$$

subject to

$$y_i - \mathbf{w}^T\phi(\mathbf{x_i}) - b \leqslant \epsilon + \xi_i, \quad i = 1,\ldots,l, \tag{4}$$

$$-y_i + \mathbf{w}^T\phi(\mathbf{x_i}) + b \leqslant \epsilon + \xi_i^*, \quad i = 1,\ldots,l, \tag{5}$$

$$\xi_i, \xi_i^* \geqslant 0, \quad i = 1,\ldots,l. \tag{6}$$

The dual form of this optimization problem is usually obtained through the minimization of the Lagrange function, constructed from the objective function and the problem constraints. In this case, the dual form of the optimization problem is the following:

$$\max \quad \left(-\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)\left(\alpha_j - \alpha_j^*\right)K(\mathbf{x_i}, \mathbf{x_j}) - \epsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*)\right) \tag{7}$$

$$\text{subject to} \quad \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0, \tag{8}$$

$$\alpha_i, \alpha_i^* \in [0, C]. \tag{9}$$

In addition to these constraints, the Karush–Kuhn–Tucker conditions must be fulfilled, and also the bias variable, $b$, must be obtained. We do not detail this process for simplicity, the interested reader can consult (Smola & Schölkopf, 1998) for reference. In the dual formulation of the problem the function $K(\mathbf{x_i},\mathbf{x_j})$ is the kernel matrix, which is formed by the evaluation of a kernel function, equivalent to the dot product $\langle\phi(\mathbf{x_i}),\phi(\mathbf{x_j})\rangle$. An usual election for this kernel function is a Gaussian function, as follows:

$$K(\mathbf{x_i}, \mathbf{x_j}) = e^{-\gamma\cdot\|\mathbf{x_i}-\mathbf{x_j}\|^2}. \tag{10}$$

The final form of function $f(\mathbf{x})$ depends on the Lagrange multipliers $\alpha_i,\ \alpha_i^*$, as follows:

$$f(\mathbf{x}) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(\mathbf{x_i}, \mathbf{x}). \tag{11}$$

Thus, the $\epsilon$-SVMr model depends on parameters $\epsilon$, which controls the width of the error margin allowed (see Eqs. (4) and (5)). Parameter $C$, which controls the regularization of the model and parameter $\gamma$, which determines the final width of the Gaussians in the final model. Note that all these parameters have influence in the set of support vectors, and their effect within the final regression model, which also influences the accuracy and robustness of the final model.

## 3. Proposed model for SVMr hyper-parameters estimation

This section describes the proposal of this technical note about the estimation of SVMr hyper-parameters estimation. The idea is to start with a coarse-grain grid search of parameters, and then carry