SVRs and Uncertainty Estimates in Wind Energy Prediction

Jesús Prada(✉) and José Ramón Dorronsoro

Universidad Autónoma de Madrid, Cantoblanco, Madrid, Spain
jesus.prada@estudiante.uam.es, jose.dorronsoro@uam.es

Abstract. While Support Vector Regression, SVR, is one of the algorithms of choice in modeling problems, construction of its error intervals seems to have received less attention. On the other hand, general noise cost functions for SVR have been recently proposed. Taking this into account, this paper describes a direct approach to build error intervals for different choices of residual distributions. We also discuss how to fit these noise models and estimate their parameters, proceeding then to give a comparison between intervals obtained using this method. under different ways to estimate SVR parameters as well as the intervals obtained by employing a full SVR Bayesian framework. The proposed approach is shown on a synthetic problem to provide better accuracy when models fitted coincide with the noise injected into the problem. Finally, we apply it to wind energy forecasting, exploiting predicted energy magnitudes to define intervals with different widths.

1 Introduction

Support vector regression, SVR, [1] has been widely used in regression problems such as stock market [2], wind energy [3] or solar radiation [4] forecasting. Classical SVR, however, does not give probability intervals to address the uncertainty in the predictions and, in fact, error interval estimation for SVR has received a somewhat limited attention in the literature. Notice that here approaches such as the well known ones for linear regression under Gaussian models completely break down, not only by the difficulty of ensuring normal random variables but, above all, by the fact that the familiar analytic estimates of the linear coefficients are simply impossible in SVR and, of course, less so, any asymptotic analysis.

In [5], a Bayesian interpretation of SVR is described and then used to propose methods to determine, first, SVR parameters by maximizing an evidence function and, second, to derive probability intervals for predictions. A drawback of these methods is that they modify the classical SVR formulation of the problem to solve, and hence existing SVR software, such as the popular LIBSVM [6] cannot be used, at least without modifying it first.

A more direct approach is proposed in [7], which assumes prediction errors to follow a specific probability distribution that, in turn, is used to define the probability intervals. Zero-mean Gaussian and Laplace families are proposed in [7] as noise models and fitted by maximum likelihood estimation, MLE, using out-of-sample residuals of SVR models; optimal SVR parameters are obtained simply

I. Rojas et al. (Eds.): IWANN 2015, Part II, LNCS 9095, pp. 564–577, 2015.
DOI: 10.1007/978-3-319-19222-2_47

by cross validation. In this paper, we follow this methodology to give probability intervals under the assumption of both zero-mean Laplace and Gaussian distributions, as well as for their non–zero mean counterparts plus the Beta and Weibull distributions. A difficulty with this approach is that it assumes that the residual distribution is independent of the predicted value and, therefore, probability intervals have exactly the same width for all input instances. General error models for SVR other than the well known ϵ–insensitive loss have been proposed in [8]. This suggest that noise distribution might be different across particular problems and it should be reflected in the particular SVR model to be used. If the assumption is true and the underlying noise distribution is accurately estimated, one should expect a reduction in interval prediction errors. We study if the proposed method can estimate this noise distribution.

Our main contributions can be summarized as follows:

- We enlarge, as mentioned, the noise models considered in [7].
- We discuss Newton–Rapshon maximum likelihood estimates for the Beta and, particularly, Weibull distributions, as well as the definition of uncertainty intervals for them.
- We show on a synthetic problem how the proposed approach is able to pair the models fitted to the residuals with the specific noise injected on the problem targets. We also compare these results to the ones obtained by a statistic test for distribution hypotheses.
- We apply the methods proposed to the estimation of uncertainty intervals for the wind energy prediction of peninsular Spain, where we also consider the use of different intervals according to predicted energy magnitudes, showing how a two group data split results in more accurate intervals.

The rest of this paper is organized as follows. Section 2 briefly reviews both classical and Bayesian SVR formulations. In Section 3 there is an in-depth description of the proposed approach for error interval estimations and experiments are carried in Section 4, where we also consider four publicly available regression datasets besides the already mentioned synthetic and wind energy problems. The paper ends with a short section on conclusions and pointers for further work.

2 Support Vector Regression

2.1 Classical SVR Review

Given a sample $D = \{(x_i, y_i) : 1 \leq i \leq N\}$ of inputs $x_i \in R^n$ and targets $y_i \in R$, the SVR problem is that of minimizing the loss function $L(w, b, \xi)$ defined as

$$L(w, b, \xi) = \frac{1}{2}\|w\|^2 + C \sum_i (\xi_i + \xi_i^*), \tag{1}$$

over w, b and ξ subject to the constraints $-\xi_i - \epsilon \leq w \cdot x_i + b - y_i \leq \xi_i^* + \epsilon, \ \ \xi_i, \xi_i^* \geq 0$ .This is known as the SVR primal problem. It can also be seen as a variant of the standard L_2 regularized regression where instead of the familiar $z_i^2 = (y_i - w \cdot x_i - b)^2$ square error, we use the ϵ–insensitive loss function [9]