Fuzzy Cost Support Vector Regression

Hadi Sadoghi Yazdi, Tahereh Royani, Mehri Sadoghi Yazdi, Sohrab Effati

Abstract—In this paper, a new version of support vector regression (SVR) is presented namely Fuzzy Cost SVR (FCSVR). Individual property of the FCSVR is operation over fuzzy data whereas fuzzy cost (fuzzy margin and fuzzy penalty) are maximized. This idea admits to have uncertainty in the penalty and margin terms jointly. Robustness against noise is shown in the experimental results as a property of the proposed method and superiority relative conventional SVR.

Keywords—Support vector regression; Fuzzy input; Fuzzy cost.

I. INTRODUCTION

The standard support vector machine works over crisp training samples. Chun-fu Lin [1, and 2] proposed fuzzy support vector machine (FSVM) by considering the noise in the training samples. They used the membership function to express the membership value of a sample to positive or negative class, but with crisp training data. So it remains a conventional support vector machine from view point of fuzzy theory. Importance degree of training data are modeled in the FSVM by insertion of membership value $\mu_i$ in penalty term of cost function to form of $\frac{1}{2}||w||^2 + C\sum_i \mu_i \xi_i$ . It is noted that the error term $\xi_i$ is scaled by $\mu_i$ . The fuzzy membership values are used to weight the soft penalty term in the cost function of SVM. The weighted soft penalty term reflects the relative fidelity of the training samples during training. Important sample with large membership value will have more emphasis in the FSVM training procedure and more effect over determination of hyperplanes.

Hong into [3] presents support vector fuzzy regression machines. This paper introduces the use of SVM for multivariate fuzzy linear and nonlinear regression models. Presented model in [3] for regression includes fuzzy input and output $\tilde{y}$ to form of $\tilde{y} = w^T \tilde{x} + \tilde{b}$ . Then a SVM model is used for calculation of crisp $w$ (weights). This model includes conventional fuzzy regression with new constraints. Upper and lower bound of fuzzy input and output are used for generation of constraints. But effect of fuzzy variables (input and output) over cost of SVR has not been considered. Assuredly, uncertainty in input data affects over margin and penalty maximization in the SVR which has not been studied in the previous works.

In [4] Ji studied support vector machine with fuzzy chance constrains to following form

$$\min \frac{1}{2}||w||^2 + C \sum_i \xi_i$$

subject to

$$\text{Pr}(\tilde{y}(W^T \tilde{x})_i + b + \xi_i \geq 1) \geq \lambda$$

$$\xi_i \geq 0, \quad i = 1, 2, ..., l$$

They showed that $\text{Pr}(\tilde{a} \leq 0) \geq \lambda$ with triangular fuzzy number $\tilde{a} = (r_1, r_2, r_3)$ and for any given level $\lambda (0 < \lambda \leq 1)$ is equivalent to: $(1 - \lambda)r_1 + r_2 \leq 0$. Thereupon, constraints of (2) are simplified.

In our previous work [5], we apply probabilistic constraints for reducing of noisy samples in maximization of margin. A Constraint is to form of $\text{Pr}(\tilde{d}_i (w^T x + b) \geq u_i \geq \delta_i$ where $u_i$ is independent random variable with known distribution functions and $0 \leq \delta_i \leq 1$ is value of effect of $i^{th}$ samples in fixation of the optimal hyperplane.

Liu in [6] presented total margin-based adaptive fuzzy support vector machines (TAF-SVM). TAF-SVM is a type of FSVM [1, and 2] which also corrects the skew of the optimal separating hyperplane due to the very imbalanced data sets by using different cost algorithm. This work is performed with dividing training data into two categories with different importance and result in dual problem is different boundary for Lagrange multipliers.

In [1], linear and quadratic functions are presented for $\mu_i$ in the FSVM which two main targets are followed, increasing margin and decreasing misclassification error. In [7] authors present two new methods for calculation of membership function of $\mu_i$ based on geometry distribution of the training samples. Those samples are near to optimal hyperplane, have similar geometry property. The main idea of FSVM [1] is that if the input is detected as an outlier or noisy sample, membership function decreases so total error term decrease. In [8] new method for $\mu_i$ of FSVM is presented which follows in the same idea that one input is assigned a low membership of the class if it is detected as an outlier. However, method of