Journal of the Taiwan Institute of Chemical Engineers 45 (2014) 2225–2232

Contents lists available at ScienceDirect

Journal of the Taiwan Institute of Chemical Engineers

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

High reliability estimation of product quality using support vector regression and hybrid meta-heuristic algorithms

Saeid Shokri [a], Mohammad Taghi Sadeghi [b,*], Mahdi Ahmadi Marvast [c]

[a] Process Simulation and Control Laboratory, Department of Chemical Engineering, Iran University of Science and Technology (IUST), Narmak, 16765-163 Tehran, Iran

[b] Department of Chemical Engineering, Iran University of Science and Technology(IUST), Tehran, Iran

[c] Process & Equipment Technology Development Division, Research Institute of Petroleum Industry (RIPI), Tehran, Iran

ARTICLE INFO

Article history:
Received 26 November 2013
Received in revised form 17 April 2014
Accepted 19 April 2014
Available online 21 May 2014

Keywords:
Soft sensor
Hydrodesulfurization
Hybrid meta-heuristic
Support vector regression
Product quality estimation
Hyper-parameters optimization

ABSTRACT

Online estimation of product quality is a complicated task in refining processes. Data driven soft sensors have been successfully employed as a supplement to the online hardware analyzers that are often expensive and require high maintenance. Support Vector Regression (SVR) is an efficient machine learning technique that can be used for soft sensor design. However, choosing optimal hyper-parameter values for the SVR is a hard optimization problem. In order to determine the parameters as fast and accurate as possible, some Hybrid Meta-Heuristic (HMH) algorithms have been developed in this study. A comprehensive study has been carried out comparing the meta-heuristic algorithms of GA and PSO to the HMH algorithms of GA–SQP and PSO–SQP for prediction of sulfur quality in treated gas oil using the SVR technique. Experimental data from a hydrodesulfurization (HDS) setup were collected to validate the proposed SVR model. The SVR model yields better performances both in accuracy and computation time (CT) for predicting the sulfur quality with hyper parameters optimized by HMH algorithms. Applying the PSO–SQP algorithm gives the best performance with AARE = 0.133 and CT = 15.88 s compared to the other methods.

1. Introduction

In order to have a reliable operation with controlled product quality in modern industrial plants, many indexes and variables are routinely measured and automatically recorded. Hardware sensors are expensive and unreliable while there are many problems with hardware sensors such as time consuming maintenance, calibration requirements, insufficient accuracy and long dead time. Soft sensors are key technologies for estimation of product quality when hardware process analyzers are not available. They can be applied as an alternative to laboratory tests [1]. Moreover, soft sensors are used to establish an online monitoring of unmeasured product indices [2].

Artificial Neural Network (ANN) has been widely used as a useful tool for nonlinear soft sensing models. However, it gives no guarantee of high convergence speed, avoiding local minima, and over-fitting phenomenon, thus lacks the generalization capability [3]. Recently, Support Vector Machines (SVM) based on machine learning formalism, developed by Vapnik [4] is gaining popularity over the ANN due to its many attractive features [5]. SVM has been efficiently employed in various prediction and classification problems. Applying the SVM to solve the regression problems is called SVR [6]. SVR has recently received much attention in soft sensor design due to its successes in building nonlinear data-driven models [7]. Moreover, SVR is a powerful tool for problems characterized by small samples, nonlinearity, high dimension space and local minima. It tries to find an optimal hyper-plane function in a high dimensional space [8].

The main problem with SVR is that its hyper-parameters has to be set properly in advance. Unsuitably chosen kernel functions or hyper-parameters settings may lead to significantly poor performance [9]. These parameters are usually set by non-heuristics algorithms (*i.e.* grid search method (GSM)) [10]. The GSM suffers from the main drawbacks of being very time-consuming, lacking guarantee of convergence to a global optimum, and reliance on parameters of the boundary selection.

In hyper-parameter optimization problem, the optimization algorithm is the most critical factor to determine the convergence speed and ability to search the global optimum. Optimization algorithms can be classified into two major categories, namely non-heuristic or exact algorithms as well as approximation algorithms. Approximation algorithms are classified into three

* Corresponding author. Tel.: +98 2177240496; fax: +98 21 77240495.
E-mail address: sadeghi@iust.ac.ir (M.T. Sadeghi).

http://dx.doi.org/10.1016/j.jtice.2014.04.016
1876-1070/

Nomenclature

AARE	average absolute relative error
iter	current iteration number
itermax	maximum number of iterations
W	inertia weight factor
Wmax	initial weight
Wmin	final weight
C_1,C_2	acceleration coefficients
P_{besti}	best position of particle *i*
G_{besti}	best position of the group
V_i	velocity of particle *i*
CT	computation time

main categories including heuristic, meta-heuristic and hyper-heuristic. Complex optimization problems cannot be solved to optimality or any guaranteed bound, by exact (deterministic) methods within a "reasonable" time limit [11]. Meta-heuristic algorithms are approximate and usually non-deterministic. There is a wide variety of the meta-heuristic algorithms. Population-based meta-heuristics include genetic algorithms and particle swarm optimization [12].

Hybrid global–local optimization algorithms have become popular approaches for solving the optimization problems. These algorithms integrate the global exploring feature of the heuristic algorithms with local fine tuning feature of non-heuristic algorithms [13]. Li et al. [14] applied PSO algorithm with a modified Broyden–Fletcher–Goldfarb–Shanno (BFGS) method to tackle the problem. Aruldoss et al. [15] applied the PSO with SQP method. Liu and Qui [16] proposed the hybrid PSO–BP algorithm that combined the PSO mechanism with the Levenberg–Marquardt algorithm or the conjugate gradient technique. Zhao et al. [17] proposed a hybrid PSO with SA method for partner selection in virtual enterprises. Liu et al. [18] developed a novel hybrid algorithm named PSO–DE, in which DE was incorporated to update the previous best positions of PSO particles forcing them to jump out of local attractors. Furthermore, there have been many efforts to hybridize GA and SQP. Mansoornejad et al. [19] developed a hybrid optimization technique based on GA–SQP to determine the kinetic parameters of a set of highly nonlinear hydrogenation reactions. Chen and Wang [20] applied a hybrid of genetic algorithm to the optimization of short-time gasoline blending scheduling problem.

Different optimization algorithms have been employed to calculate the best hyper parameters of an SVR model. Most of studies on the parameter selection of SVR are focused on global optimization algorithms [21,22]. However, some optimization studies showed that the simultaneous application of local and global search techniques would improve the performance considerably either in CT or the accuracy of results [23].

Many researchers have developed heuristic algorithms for parameter optimization of SVR. Min et al. [24] studied the optimal parameter selection on SVM technique. Wu et al. [25] optimized the kernel parameters of SVR using GA. Huang [26] recommended a hybrid GA–SVR methodology to solve an important stock selection for an investment problem. Chen and Wang [27] optimized the SVR parameters using meta-heuristic algorithms such as real value GA and adopted the optimal parameters to construct the SVR models in forecasting tourism demand.

Tuning the hyper-parameters is a hard optimization problem. The parameters are the most important factors for efficiency and robustness of an SVR model. Therefore, potentials of the hybrid strategies for optimization of these parameters need to be further investigated. This study proposes a novel approach for using the SVR models in large scale optimization to increase their performance both in accuracy and CT by means of developing some HMH algorithms.

2. Methodology

2.1. The Support Vector Regression (SVR)

SVR is utilized to determine a nonlinear relation of the form $y = f(x)$ between the vectors of observation x and the desired y from a given set of training samples.

A number of cost functions such as the Laplacian, Huber's, Gaussian, and ε-insensitive can be used for the SVR formulation. Among the functions, the robust ε-insensitive loss function (L_ε) is more common [28]:

$$L_\varepsilon(f(x)-y)=\begin{cases}|f(x)-y|-\varepsilon & |f(x)-y|\geq\varepsilon\\ 0 & \text{otherwise}\end{cases} \tag{1}$$

where, ε is a precision parameter representing the radius of the tube located around the regression function, $f(x)$. The goal in using the ε-insensitive loss function is to find a function that fits the current training data with a deviation less than or equal to ε. C and ε are user-defined parameters in the empirical analysis. The constant $C > 0$ is a parameter determining the trade-off between generalization ability and accuracy in the training data while the parameter ε defines the degree of tolerance to errors. The optimization problem can be reformulated as:

$$Min\frac{1}{2}||w||^2+C\sum_{i=1}^{l}(\xi_i^-+\xi_i^+) \tag{2}$$

That is subject to the constraints given below:

$$y=\begin{cases}y_i-(\langle w,x_i\rangle+b)\leq\varepsilon+\xi_i\\ (\langle w,x_i\rangle+b)-y_i\leq\varepsilon+\xi_i^*\\ \xi_i,\xi_i^*\geq 0\end{cases}$$

The positive slack variables ξ and ξ^* represent the distance from actual values to the corresponding boundary values of the ε-tube.

For SVR based models, four different kernel functions, including linear, quadratic, Gaussian, and polynomial are used. Generally, the application of Gaussian function is shown to yield a better prediction performance [29].

$$K(x,y)=\exp\left(-\frac{||x_i-y_j||^2}{2\sigma^2}\right) \tag{3}$$

In order to build an SVR model efficiently, the SVR parameters must be specified carefully. These parameters include the kernel function, regularization parameter C, Bandwidth of the kernel function (σ^2) and tube size of the ε-insensitive loss function (ε).

2.2. Hybrid optimization approach

Hyper-parameter tuning is a complicated problem in SVR model. It has a great effect on predicted values and model performance. It is difficult to solve the complicated problem with merely a single algorithm; therefore, a hybrid optimization approach that combines multiple optimization algorithms is useful. For this purpose, hybridization of GA and PSO algorithms (as a global optimizer) with SQP (as a local fine tuner) are employed. The GA and PSO algorithms have been utilized to initialize the starting point as the input for SQP algorithm.

In summary, integration of GA/PSO and SQP with SVR is an effective method for hyper-parameter optimization of SVR model