Chemometrics and Intelligent Laboratory Systems 180 (2018) 36–41

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Just-in-time semi-supervised soft sensor for quality prediction in industrial rubber mixers

Wenjian Zheng, Yi Liu *, Zengliang Gao, Jianguo Yang

Institute of Process Equipment and Control Engineering, Zhejiang University of Technology, Hangzhou, 310014, China

ARTICLE INFO

Keywords:
Soft sensor
Just-in-time learning
Semi-supervised learning
Extreme learning machine
Rubber mixing process

ABSTRACT

Increasing data-driven soft sensors have been adopted to online predict the quality indices in polymerization processes to improve the availability of measurements and efficiency. However, in industrial rubber mixing processes, most existing soft sensors for online prediction of the Mooney viscosity only utilized the limited labeled data. By exploring the unlabeled data, a novel soft sensor, namely just-in-time semi-supervised extreme learning machine (JSELM), is proposed to online predict the Mooney viscosity with multiple recipes. It integrates the just-in-time learning, extreme learning machine (ELM), and the graph Laplacian regularization into a unified online modeling framework. When a test sample is inquired online, the useful information in both of similar labeled and unlabeled data is absorbed into its prediction model. Unlike traditional just-in-time learning models only utilizing labeled data (e.g., just-in-time ELM and just-in-time support vector regression), the prediction performance of JSELM can be enhanced by taking advantage of the information in lots of unlabeled data. Moreover, an efficient model selection strategy is formulated for online construction of the JSELM prediction model. Compared with traditional soft sensor methods, the superiority of JSELM is validated via the Mooney viscosity prediction in an industrial rubber mixer.

1. Introduction

Accurate and reliable measurements of process variables and quality indices in chemical processes can ensure the success in their products. The industrial rubber mixing process is a fast (only 2–5 min), nonlinear, and time-varying batch process performed in an internal mixer. The Mooney viscosity is a key quality index which represents the viscoelastic behavior of an elastomer. Up to date, an accurate first-principles Mooney viscosity model is still not available. On the other hand, in most rubber/tire factories, the Mooney viscosity can be only obtained offline using laboratory analysis several hours later after a batch has been discharged [1,2]. In the absence of an economical or effective online measurement, soft sensors (or inferential sensors) could serve as an alternative solution [3–7]. Additionally, with the wide availability of process data in rubber/tire factories, increasing data-driven soft sensors have been adopted to predict the Mooney viscosity information [8–16].

Nowadays, common data-driven soft sensors for the Mooney viscosity prediction include neural networks (NN) [8–10], multivariable regression (e.g., partial least squares regression), and kernel learning-based regression approaches [11–16]. However, up to now, most of them are constructed in a supervised learning manner. This means that complete data samples including both input and output variables are need. Here, the labeled dataset denotes the one containing both of input and output data. And the unlabeled dataset represents the one only has input data. In the machine learning area, training a model with both of labeled and unlabeled data is known as semi-supervised learning. By suitably incorporating the information of unlabeled data into the supervised regression model, more accurate prediction of semi-supervised models can often be obtained than related supervised methods [17,18]. However, in chemical processes, semi-supervised soft sensor applications are still much fewer than those with supervised models [19–22]. From a practical modeling viewpoint, with lots of unlabeled data available, the development of novel semi-supervised soft sensors is promising.

As shown in Fig. 1, due to the time-consuming and costly lab assaying process in industrial rubber/tire plants, the assayed Mooney viscosity are delayed and limited. On the other hand, the mixing pressure, the mixing temperature, and other mixing variables are online measured continually during every production batch in an industrial mixer. This means lots of unlabeled data are available, while labeled data are limited. To cope with the gap, recently, a semi-supervised extreme learning machine (SELM) [23] and its improved version were applied to predict the Mooney viscosity with one recipe and showed better prediction performance than

* Corresponding author.
E-mail address: yliuzju@zjut.edu.cn (Y. Liu).

https://doi.org/10.1016/j.chemolab.2018.07.002
Received 30 January 2018; Received in revised form 20 June 2018; Accepted 2 July 2018
Available online 3 July 2018
0169-7439/© 2018 Elsevier B.V. All rights reserved.

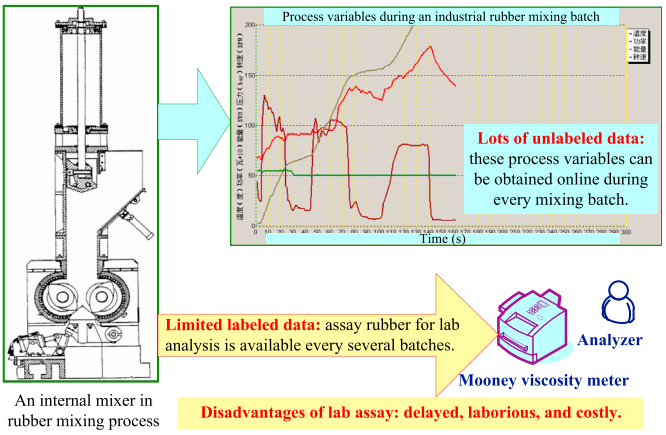

Fig. 1. Lots of unlabeled data during every batch and limited labeled data only available offline every several batches.

only using the soft sensor with labeled data [24]. However, for multiple mixing recipes with different characteristics, only using a single model is still not enough. Additionally, the varying properties of raw materials and mixing operating conditions indeed introduce batch-to-batch variations to both of the quality index and process data. As pointed out by Lu and Chiang [7], the life spans of most data-driven soft sensors are limited in practice. Consequently, a flexible model with adaptive structure is more suitable and attractive in industrial processes.

Just-in-time learning (JITL) methods, as an alternative solution, have been utilized in the development of data-driven soft sensors in chemical processes, especially for those with multiple modes/grades [25–29]. However, most traditional JITL-based soft sensors were built in a supervised learning manner. For online prediction of a query sample, only the labeled data are considered as the similar data. The information hidden in lots of unlabeled data (e.g., those online measured process variables) is omitted in selection and modeling of similar samples. As a result, the performance of JITL-based soft sensors may be restricted to some extent, especially when the labeled data are limited in industrial practice. To our best knowledge, the JITL-based semi-supervised soft sensor has not yet been developed, especially for rubber mixing process applications.

In this work, a JITL-based SELM (denoted as JSELM) soft sensor is proposed to predict the Mooney viscosity. The JSELM method integrates the JITL modeling manner, extreme learning machine (ELM) [30,31], and the graph Laplacian regularization into a unified framework. For online inquiry of a test sample, the useful information in both of similar labeled and unlabeled data is absorbed into its prediction model. Additionally, a fast cross-validation strategy is formulated for online construction of the JSELM model efficiently. Consequently, compared with traditional JITL models only using labeled data (e.g., JITL-based ELM [29]), the prediction performance of JSELM can be enhanced by taking advantage of the information in unlabeled data.

The remainder is so organized. In Section 2, after a brief introduction of ELM, the SELM soft sensing method is formulated. In Section 3, the proposed JSELM method with its model selection strategy is developed. The JSELM approach is validated via the industrial Mooney viscosity prediction in Section 4. Finally, the conclusion is drawn in Section 5.

2. SELM soft sensor modeling method

2.1. Brief overview of ELM model

In chemical processes, NN-based soft sensors have been widely utilized in past two decades. With a single layer feed-forward NN structure, the weights of the hidden neurons in ELM can be obtained very fast [30,31]. Due to its easy-to-use and nonlinear modeling properties, increasing ELM-based soft sensor applications have recently caught more attention [29,32–34]. Here, the ELM-based supervised modeling algorithm is briefly described. The labeled set is denoted as $\{\mathbf{S}^l\} = \{\mathbf{X}^l, \mathbf{Y}^l\}$, where $\{\mathbf{X}^l\} = \{\mathbf{x}_i^l\}_{i=1}^L$ and $\{\mathbf{Y}^l\} = \{y_i^l\}_{i=1}^L$ are the input and output datasets with L samples, respectively. ELM with N hidden nodes and the activation function $g(\cdot)$ can approximate the training data with zero error, which means $\sum_{i=1}^L \left\| y_i^l - \hat{y}_i^l \right\| = 0$, where y_i^l and $\hat{y}_i^l$ denote the measured output and predicted output, respectively. Compactly, the regression formulation of ELM is represented as [30,31]:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{Y}^l \tag{1}$$

where the output matrix of hidden-layer $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_N]_{L\times N}$ with $\mathbf{h}_i = \begin{bmatrix} g(\langle \mathbf{a}_i, \mathbf{x}_1^l \rangle + b_i) \\ \vdots \\ g(\langle \mathbf{a}_i, \mathbf{x}_L^l \rangle + b_i) \end{bmatrix}_{L\times 1}, i = 1, \cdots, N$; $g(\langle \mathbf{a}_i, \mathbf{x}_j^l \rangle + b_i)$ is the output of the ith hidden node related to the jth input $\mathbf{x}_j^l$. The terms $\mathbf{a}_i$ and b_i denote the input weight and the bias of the ith hidden node, respectively; and $\langle \mathbf{a}_i, \mathbf{x}_j^l \rangle$ denotes the inner product of $\mathbf{a}_i$ and $\mathbf{x}_j^l$. Here, the common sigmoidal function $g(v) = \frac{1}{1+\exp(-v)}$ is adopted for its nonlinear modeling ability. And $\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix}_{N\times 1}$ is the output weight parameter.

The weights of ELM are not necessarily retuned using some complex training algorithms [30,31]. For many regression cases, the number of training samples is much more than the number of hidden nodes, i.e., $L >> N$. As a result, the output weights $\hat{\boldsymbol{\beta}}$ are formulated as [30,31]:

$$\hat{\boldsymbol{\beta}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{Y}^l \tag{2}$$

Furthermore, to avoid the problem of $\mathbf{H}^T\mathbf{H}$ being noninvertible, a regularized ELM (RELM) model was constructed by simply adding a small value of ridge parameter $\gamma > 0$ to obtain $\hat{\boldsymbol{\beta}}$ [29,32].

$$\hat{\boldsymbol{\beta}} = (\mathbf{H}^T\mathbf{H} + \gamma\mathbf{I})^{-1}\mathbf{H}^T\mathbf{Y}^l \tag{3}$$

where $\mathbf{I}$ is a unit matrix.

Finally, for a test sample $\mathbf{x}_t = [x_{t1}, x_{t2}, \cdots, x_{tn}]^T \in R^n$, its prediction $\hat{y}_t$ is obtained below:

$$\hat{y}_t = \mathbf{h}_t\hat{\boldsymbol{\beta}} = \mathbf{h}_t(\mathbf{H}^T\mathbf{H} + \gamma\mathbf{I})^{-1}\mathbf{H}^T\mathbf{Y}^l \tag{4}$$

where $\mathbf{h}_t$ is the output hidden-layer vector related to $\mathbf{x}_t$.

2.2. SELM modeling approach

Semi-supervised learning algorithms assume that the input patterns from both labeled and unlabeled data are drawn from the same marginal distribution. Additionally, with the smoothness assumption [17,18], the data in the local region should have similar labels. In such a situation, the unlabeled data naturally provide useful information for exploring the data structure in the input space. By assuming that the input data follow the same manifold in the input space, semi-supervised learning algorithms can incorporate both labeled and unlabeled data into the learning process. Consequently, benefiting from lots of unlabeled data, the SELM model could provide more accurate prediction performance than ELM [23,24].

The input and output data samples are denoted as $\{\mathbf{X}\} = \{\mathbf{X}^l \cup \mathbf{X}^u\}$ and $\mathbf{Y} = \begin{bmatrix} \mathbf{Y}^l \\ \mathbf{Y}^u \end{bmatrix} = \begin{bmatrix} y_1^l \\ \vdots \\ y_L^l \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(L+U)\times 1}$, respectively. Correspondingly, the hid-

Download English Version:

https://daneshyari.com/en/article/7561773

Download Persian Version:

https://daneshyari.com/article/7561773

Daneshyari.com