Neurocomputing 131 (2014) 143–156

Contents lists available at ScienceDirect

Neurocomputing

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

Optimally pruned extreme learning machine with ensemble of regularization techniques and negative correlation penalty applied to automotive engine coldstart hydrocarbon emission identification

Ahmad Mozaffari, Nasser L. Azad *

Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada N2L, 3G1

ARTICLE INFO

Article history:
Received 25 July 2013
Received in revised form
28 September 2013
Accepted 26 October 2013
Communicated by G.-B. Huang
Available online 11 November 2013

Keywords:
Automotive engine modeling
Coldstart
Negative correlation based ensembles
Regularization
Extreme learning machines

ABSTRACT

In this investigation, the authors test the efficacy and reliability of optimally pruned extreme learning machine with ensemble of regularization techniques to identify the exhaust gas temperature (T_{exh}) and the engine-out hydrocarbon emission (HC_{raw}) during the coldstart operation of automotive engines. These variables have significant impact on the cumulative tailpipe emissions (HC_{cum}) during a coldstart phenomenon, which is the number one emission-related problem for today's spark-ignited (SI) engine vehicles. To do so, the concepts of ensemble computing with negative correlation learning (NCL) and pruning of neurons are used in tandem to cope with difficulties associated with extracting knowledge from collected database. In the proposed framework, the regularization strategies are adopted to help us increasing the numerical stability of identifier while mitigating the redundant complexity of hidden neurons. Moreover, to increase the generalization of identifier and also reduce the effects of uncertainty, an ensemble of independent OP-ELM with NCL selection criterion called OP-ELM-ER-NCL is taken into account. To endorse the valid performance of OP-ELM-ER-NCL for modeling the characteristics of engine over the coldstart phenomenon, its performance is compared to a set of well-known identification systems, i.e. standard extreme learning machine (ELM), back-propagation neural network (BPNN), OP-ELM with different types of regularization, ensemble of regularized OP-ELM without negative correlation (OP-ELM-ER), and an ensemble ELM with a constrained linear system of leave-one-out outputs (E-LL), in terms of both accuracy and computational complexity. The simulation results indicate that the proposed identifier is really capable of capturing the knowledge of collected database. It is observed that its resulted accuracy and robustness are comparable with those obtained by identification methods available in the literature. Besides, using NCL strategy aids the ensemble to select the most effective regularization techniques and remove the redundant (ineffective) ones, which consequently decreases the complexity of final ensemble.

1. Introduction

In recent decade, an obvious trend has been emerged towards using computationally efficient neural identifiers. The main motivation behind such an interest lies on an inefficient and time consuming training procedure of traditional neural networks. To be more to the point, promotion of industry, social communications, trades, financial issues and global competition have obliged engineers of computational intelligence (CI), system sciences and control societies to seek for efficient expert systems with highest possible generalization capabilities and practical implications. Nowadays, the scientists of CI society and system science engineers achieve a compromise on the main stream line of investigations to develop efficient approaches suited for prediction, classification, intelligent control, incremental data mining, time series forecasting, etc. (see Refs [7,13] for more details on chronological stream line of investigation).

Machine learning (ML) techniques play a predominant role on promotion of CI and expert systems. Through enormous number of researches, it has been proved that using efficient ML strategies can explicitly increase the accuracy of fuzzy systems, neural systems, and their hybridization consequents [11,31]. Different types of ML strategies are used to adjust the constructive parameters of identifiers. Among them, iterative learning techniques such as back propagation, Levenberg-Marquardt, heuristics (e.g. Nelder–Mead simplex and golden search) and swarm, evolutionary and memetic metaheuristics (in both single objective and multiobjective forms) have gained remarkable reputation [5]. Although the mentioned techniques have proved their practical

* Corresponding author. Tel.: +1 98 8845 6738797.
E-mail addresses: amozaffari10@yahoo.com (A. Mozaffari), nlashgar@uwaterloo.ca (N.L. Azad).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.10.030

authenticity through several researches, they suffer from several draw backs which hinder their applicability for cases in which there is an obvious need for incremental learning, fast computation, recurrent optimization, real-time control, etc. Here, to avoid any prejudice or biased claim, we elicit some of their main practical drawbacks from existed reports in the literature:

1. The need for optimizing all of the constructive parameters requires a high computational burden. This issue is even much more obvious when the structure of expert system is rather complex (i.e. neural network with high number of hidden nodes or fuzzy system with large rule base) [11].
2. One of the major flaws of such systems is in real-time and incremental circumstances in which they should be fast enough to be adapted to systematic changes. Unfortunately, most real-world systems have incremental characteristics, and therefore one should try his/her best to remarkably decrease the algorithmic redundancies to hasten the training process [11].
3. As these iterative techniques do not use analytical operators, their performance and convergence strictly relates to the efficacy of the controlling parameters, e.g. the learning rate and momentum in BP, exploration/exploitation balancing values in swarm and evolutionary algorithms, etc. The urge for adjusting the values of these controlling parameters requires further sensitivity analyzes or automatic supervisions [28].
4. As the iterative training schemes concern with use of algorithms, certain knowledge of the algorithm analysis and programming are required to guarantee the best performance of an expert system. Besides, their use of algorithms, especially stochastic algorithms such as swarm and evolutionary methods, produces un-desired sensitivity of the expert system to environmental spaces. As an example, the expert should have certain knowledge of stochastic computations to find out which modification on which algorithm is best suited for learning specific types of expert systems [5].

As it can be seen, in spite of high accuracy, iterative techniques may result different problems when dealing with large data sets. To cope with such a flaw, researchers have tried to develop an expert system which not only mitigates the computational difficulties associated with iterative approaches, but also yields acceptable results in terms of both accuracy and robustness. Proposed by Huang et al. [9], an extreme learning machine (ELM) is a really fast iterative-less neural training strategy which uses an analytical algebraic procedure to adjust the regression weights of a single hidden layer feedforward neural network (SLFN). Since its proposition, several theoretical studies have been done on its structural characteristics.

The main provocations behind theoretical researches on ELM are twofold: (1) improving its accuracy by devising several existed methodologies and (2) reducing its computational complexity by using pruning approaches [10]. The second motivation attracts much more percentage of researchers. This may be due to the fact that the training of ELM is conducted by algebraic approaches, and thus cannot be modified significantly. However, the studies on standard ELM have indicated that it suffers from irrelevant variables and information in database, which deemed to happen in real-world problems [17]. For the first time, this flaw was mainly reported in Ref. [16] and addressed by proposing a network pruning technique. The authors called their method optimally pruned extreme learning machine (OP-ELM). OP-ELM uses LASSO [4] regularization technique to rank the neurons of hidden layer. In spite of their promising findings, standard OP-ELM has a problem which relates to computations of the pruning criterion. In Ref. [17], the authors proposed a cascade regularization method based on LASSO and Tikhonov criteria [17] to cope with the computational difficulties of standard OP-ELM. They reported that their method significantly improves the performance of OP-ELM. Besides, through numerical experiments, they observed that their technique always remains less than one order of magnitude slower than OP-ELM.

Taking the results of prior numerical studies into account, in this paper, the authors aim at using a novel structure of ELM to not only avoid redundancies in the hidden layer, but increase its generalization capability. In other words, the authors would like to address both concerns (two criteria stated at the beginning of the last paragraph) regarding the use of ELM. This can be done by considering the negative selection based ensemble training into account. Since its proposition, the characteristics of such a topology design scheme have been thoroughly studied in the literature [14,26]. The negative correlation learning (NCL) obliges the ensemble to select sub-components of highest possible diversity. To be more to the point, NCL adds a penalty term to objective function which fosters the ensemble to choose sub-components which capture the knowledge of database in different fashions. This helps the entire ensemble to act as a universal identifier capable of identifying different characteristics of database. Hence, in our method, the main obligation of ensemble is to select the ELMs of different regularization properties. The details of proposed ensemble identifier are presented more closely in the rest of the paper.

To elaborate on the effectiveness of the designed ensemble identifier, it is used to derive a soft model for variations of T_{exh} and HC_{raw} during coldstart operations of automotive engines, as a highly nonlinear, transient and uncertain phenomenon. Based on comparative results, it is observed that optimally pruned extreme learning machine with ensemble of regularization techniques can cope with difficulties associated with modeling the coldstart main characteristics. Besides, it is observed that the method shows comparative or even better results as compared to the considered rival identifiers.

The rest of the paper is organized as follows. In Section 2, the authors provide a concise review which clearly outlines the main progressive strides taken for identifying the main characteristics of coldstart phenomenon. The experimental procedures required for collecting the database are presented in Sections 3. Section 4 describes the characteristics of the proposed ensemble identifier. Section 5 is devoted to numerical experiments and consequent discussions. Finally, the paper is concluded in Section 6, and some remarks are provided for future investigations.

2. Modeling coldstart phenomenon

2.1. Practical advances and open questions

In the recent decade, emission regulations for automotive engines have become more stringent. That would not be an exaggeration, if we state that the authorities of automotive industry exert most of their budgets on reducing the emission of un-desired gasses from automobile engines [21,22,32]. The importance of fulfilling such environmental provisions obliges the industrialists and engineers to decrease their financial interests to some degree, and instead, allocates a remarkable financial budget to improve the performance of engines. To do so, several exhaustive and practical efforts need to be done. Achieving low level tailpipe emissions in spark-ignited (SI) engines is one of the most important elements among them [22,32]. Unburned hydrocarbons (*HCs*), carbon monoxide and nitrogen oxides are enumerated as the most important pollutant materials expelled from engines during their performance.

By pursuing the main stream of investigations and reported documents on mitigating the engine's emission, it can be inferred