Neurocomputing 177 (2016) 304–316

Contents lists available at ScienceDirect

Neurocomputing

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

Stochastic gradient based extreme learning machines for stable online learning of advanced combustion engines

Vijay Manikandan Janakiraman [a,*], XuanLong Nguyen [b], Dennis Assanis [c]

[a] Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI, USA
[b] Department of Statistics, University of Michigan, Ann Arbor, MI, USA
[c] Stony Brook University, NY, USA

ARTICLE INFO

Article history:
Received 6 May 2015
Received in revised form
1 November 2015
Accepted 7 November 2015
Available online 28 November 2015

Keywords:
Online learning
Extreme learning machine
System identification
Lyapunov stability
Engine control
Operating envelope

ABSTRACT

We propose and develop SG-ELM, a stable online learning algorithm based on stochastic gradients and Extreme Learning Machines (ELM). We propose SG-ELM particularly for systems that are required to be stable during learning; i.e., the estimated model parameters remain bounded during learning. We use a Lyapunov approach to prove both asymptotic stability of estimation error and boundedness in the model parameters suitable for identification of nonlinear dynamic systems. Using the Lyapunov approach, we determine an upper bound for the learning rate of SG-ELM. The SG-ELM algorithm not only guarantees a stable learning but also reduces the computational demand compared to the recursive least squares based OS-ELM algorithm (Liang et al., 2006). In order to demonstrate the working of SG-ELM on a real-world problem, an advanced combustion engine identification is considered. The algorithm is applied to two case studies: An online regression learning for system identification of a Homogeneous Charge Compression Ignition (HCCI) Engine and an online classification learning (with class imbalance) for identifying the dynamic operating envelope. The case studies demonstrate that the accuracy of the proposed SG-ELM is comparable to that of the OS-ELM approach but adds stability and a reduction in computational effort.

1. Introduction

Homogeneous Charge Compression Ignition (HCCI) Engines are of significant interest to the automotive industry owing to their ability to reduce emissions and fuel consumption significantly compared to existing methods such as spark ignition (SI) and compression ignition (CI) engines [1–3]. Although HCCI engines tend to do well in laboratory controlled tests, practical implementation is quite challenging because HCCI engines do not have a direct trigger for ignition (such as spark in SI or fuel injection in CI). Further, HCCI requires some special engine designs such as exhaust gas recirculation (EGR) [4], variable valve timings (VVT) [5], intake charge heating [6] among others. Such advanced designs also increase the complexity of the engine operation making it unstable and extremely sensitive to operational disturbances [7,8]. A model based control is typically opted to address the challenges involved in controlling HCCI [9,5,10]. For model development, both physics based approaches [9,5,10] and data based approaches [11–14] were shown to be effective. A key requirement for a model based control is the ability of the models to accurately predict the engine state variables for several operating cycles ahead of time, so that a control action with a known consequence can be applied to the engine. Further, in order to be vigilant against the engine drifting towards instabilities such as misfire, ringing, knock, etc. [15,16], the operating limits of the engine particularly in transients, is required to be known. In order to develop controllers and operate the engine in a stable manner, both models of the engine state variables as well as the operating envelope are necessary.

Data based modeling approaches for the HCCI engine state variables and dynamic operating envelope were demonstrated using neural networks [11], support vector machines [12], extreme learning machines [13,14] by the authors. However, previous research considered an offline approach where the data collected from engine experiments were taken offline and models were developed using computer workstations that had high processing and memory. However, a key requirement in advancing HCCI modeling is to perform online learning for the following reasons. The models developed offline are valid only in the controlled experimental conditions. For instance, the experiments are performed at a controlled ambient temperature, pressure and

* Corresponding author. Presently at: NASA Ames Research Center, MS 269-1, Moffett Field, CA 94035-1000, USA. Tel.: +1 734 358 6633.
E-mail address: vijai@umich.edu (V.M. Janakiraman).

http://dx.doi.org/10.1016/j.neucom.2015.11.024
0925-2312/

humidity conditions. As a result, the models developed are valid for the specified conditions and when the models are implemented on a vehicle, the expectation is that the model works on a wide range of climatic conditions that the vehicle is exposed to, possibly on conditions that were not experimented. Hence, an online adaptation to learn the behavior of the system at new/unfamiliar situations is required. Also, since the offline models are developed directly from experimental data, they may perform poorly in certain operating regions where the density of experimental data is low. As more data becomes available in such regions, an online mechanism can be used to adapt to such data. In addition, the engine produces high velocity streaming data; operating at about 2500 revolutions per minute, an in-cylinder pressure sensor can produce about 1.8 million data observations per day. It becomes infeasible to store this volume of data for offline model development. Thus, an online learning framework that processes every data observation, updates the model and discards the data is required for advanced engines like HCCI.

Online learning, as the name suggests, refers to obtaining a model online; i.e., learning happens when the system is in operation and as data is streaming in. Typically, the learning is sequential; i.e., the data from the system is processed one-by-one or batch-by-batch and the model parameters are updated. A data processor on-board a combustion engine usually is low on computation power and memory. Thus, simple linear models used to be the natural choice for combustion engines. However, for a system like the HCCI engine, linear models may be insufficient to capture the complex dynamics, particularly for predicting several steps ahead in time [11]. While numerous nonlinear methods for online learning do exist in machine learning literature, a complete survey is beyond the scope of this paper. The recent paper on online sequential extreme learning machines (OS-ELM) [17] surveys popular online learning algorithms in the context of classification and regression and develops an efficient algorithm based on recursive least squares. The OS-ELM algorithm appears to be the present state of the art (although some variants have been proposed such as [18–20]) for classification/regression problems achieving a global optimal solution, high generalization accuracies and most importantly, in quick time. Also, based on observations from our previous work [21], we choose extreme learning machines (ELM) over other popular methods such as neural networks and support vector machines for the HCCI engine problem. It has been shown that both polynomial and linear methods were inferior in terms of prediction accuracy [12,11] although they have simple algorithms suitable for online applications. The online variants of SVM usually work by approximating the batch (offline) loss function so that data can be processed sequentially [22,23] and achieve accuracies similar to that of the offline learning counterparts. However, SVMs Come with a high computation and memory requirement to be used efficiently on a memory limited system such as the engine control unit [13]. Thus we prefer ELM over SVM and other state of the art nonlinear models.

In spite of its known advantages, an over-parameterized ELM may suffer from ill-conditioning problem when a recursive least squares type update is performed (as in OS-ELM). This sometimes results in poor regularization behavior as reported in [24,25,20,26,27], which leads to an unbounded growth of the model parameters and unbounded model predictions. This may not be a serious problem for many applications as the model usually improves as more data becomes available. However, for control problems in particular, if decisions are made simultaneously based on the online learned model (as in adaptive control [28]), it is critical that the parameter estimation algorithm behaves in a stable manner so that control actions can be trusted at all times. Hence a guarantee of stability and parameter boundedness is of extreme importance. To address this issue, we propose the SG-ELM, a stable online learning algorithm based on stochastic gradient descent and extreme learning machines. By extending ELM to include a notion of stable learning, we hope that the simplicity and generalization power of ELM can be retained along with stability of identification, suitable for real-time control applications. We use a Lyapunov approach to prove both asymptotic stability of estimation error and boundedness in the estimated parameters suitable for identification of nonlinear dynamic systems. Using the Lyapunov approach, we determine an upper bound for the learning rate of SG-ELM that seems to avoid bad regularization that may arise during online learning. These are the main contributions of this paper. Further, we also apply the SG-ELM algorithm to two real-world HCCI identification problems including online state estimation and online operating boundary estimation which is a novel application of online extreme learning machines.

The remainder of the article is organized as follows. The ELM modeling approach is described in Section 2 along with algorithm details on batch (offline) learning as well as the present state of the art; the OS-ELM algorithm. In Section 3, the stochastic gradient based ELM algorithm is derived along with a stability proof. In Section 4, the background on HCCI engine and experimentation are discussed. Sections 5 and 6 cover the discussions on the application of the SG-ELM algorithm on the two applications, followed by conclusions in Section 7.

2. Extreme learning machines

Extreme Learning Machine (ELM) is an emerging learning paradigm for multi-class classification and regression problems [29,30]. An advantage of the ELM method is that the training speed is extremely fast, thanks to the random assignment of input layer parameters which do not require adaptation to the data. In such a setup, the output layer parameters can be analytically determined using a linear least squares approach. Some of the attractive features of ELM [29] include the universal approximation capability of ELM, the convex optimization problem of ELM resulting in the smallest training error without getting trapped in local minima, closed form solution of ELM eliminating iterative training and better generalization capability of ELM [30]. In comparison, a backpropagation neural network has the same objective function as that of ELM but they often get trapped in local minima whereas ELM do not. Support vector machines on the other hand, solves a convex optimization problem but the computation involved is quite high and running times are slow for large data-sets. Thus, ELM appears to be very efficient both in terms of accuracy and running times compared to several state-of-the-art algorithms.

Consider the following data set

$$\{(x_1, y_1), \ldots, (x_N, y_N)\} \in (\mathcal{X}, \mathcal{Y}), \tag{1}$$

where N denotes the number of training samples, $\mathcal{X}$ denotes the space of the input features and $\mathcal{Y}$ denotes labels whose nature differentiate the learning problem in hand. For instance, if $\mathcal{Y}$ takes integer values $(1, 2, 3 ,.. \}$ then the problem is referred to as classification and if $\mathcal{Y}$ takes real values, it becomes a regression problem. ELMs are well suited for solving both regression and classification problems faster than state of the art algorithms [30]. A further distinction could be made depending on the availability of training data during the learning process, as offline learning (or batch learning) and online learning (or sequential learning). Offline learning could make use of all training data simultaneously as all data is available to the algorithm and time is generally not a limiting factor. So it is possible to have the model see the data several times (iterations) so that the best accuracy can be