Neurocomputing 350 (2019) 70–90

Contents lists available at ScienceDirect

Neurocomputing

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

Towards a more efficient and cost-sensitive extreme learning machine: A state-of-the-art review of recent trend

Peter Adeniyi Alaba[a,*], Segun Isaiah Popoola[b], Lanre Olatomiwa[c,d], Mathew Boladele Akanle[b], Olayinka S. Ohunakin[e,h], Emmanuel Adetiba[b,f], Opeoluwa David Alex[g], Aderemi A.A. Atayero[b], Wan Mohd Ashri Wan Daud[a]

[a] Department of Chemical Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
[b] Department of Electrical and Information Engineering, Covenant University, Ota, Ogun-State, Nigeria
[c] Department of Electrical and Electronics Engineering, Federal University of Technology, PMB 65, Minna, Nigeria
[d] Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough, United Kingdom
[e] The Energy and Environment Research Group (TEERG), Department of Mechanical Engineering, Covenant University, Ogun State, Nigeria
[f] HRA, Institute for Systems Science, Durban University of Technology, P.O. Box 1334, Durban, South Africa
[g] Department of Computer Science, University of the People, 225 S Lake Ave Suite 300, Pasadena, CA 91101, USA
[h] Senior Research Associate, Faculty of Engineering & the Built Environment, University of Johannesburg, South Africa

ARTICLE INFO

Article history:
Received 26 July 2018
Revised 21 February 2019
Accepted 12 March 2019
Available online 17 April 2019

Communicated by Miche Y

Keywords:
Extreme learning machine
Artificial intelligence
Big data analytics
Sample structure preserving
Imbalance data

ABSTRACT

In spite of the prominence of extreme learning machine model, as well as its excellent features such as insignificant intervention for learning and model tuning, the simplicity of implementation, and high learning speed, which makes it a fascinating alternative method for Artificial Intelligence, including Big Data Analytics, it is still limited in certain aspects. These aspects must be treated to achieve an effective and cost-sensitive model. This review discussed the major drawbacks of ELM, which include difficulty in determination of hidden layer structure, prediction instability and Imbalanced data distributions, the poor capability of sample structure preserving (SSP), and difficulty in accommodating lateral inhibition by direct random feature mapping. Other drawbacks include multi-graph complexity, global memory size, one-by-one or chuck-by-chuck (a block of data), global memory size limitation, and challenges with big data. The recent trend proposed by experts for each drawback is discussed in detail towards achieving an effective and cost-sensitive model.

1. Introduction

Extreme learning machine (ELM) is a kind of neural network (NN) characterized by biologically inspired single-hidden-layer feedforward network (SLFN), using biological learning techniques rather than artificial learning techniques. It is a biological learning technique that involves the use of kernels, random neurons (with or without unknown modeling/shape), and optimization constraint. ELM is more effective in terms of speed, generalization performance, simplicity and efficiency than the traditional NN in practical applications. The word “extreme” implies beyond conventional artificial learning methods, towards brain-like learning [1]. ELM helps to fill the gap between biological learning and machine learning mechanism [1,2]. Rather than using known activation function such as sigmoid, ELM uses unknown nonlinear piecewise continuous functions h(x) being the real activation functions of most living brain neurons [1]. Theoretically, ELM somehow combines brain learning features, matrix theory, control theory, neural network theory, and linear system theory, which were previously regarded to be isolated with big gaps.

Due to the capability for a wide range of activation functions h(x), ELM exhibits universal classification capability and universal approximation capability [1]. ELM can be used in solving problems pertaining to regression, classification, representational learning, feature selection, clustering, and several other learning tasks. Successful applications of ELM have been reported in several domains, such as output power forecasting [3], system identification [2], function approximation [4,5], biomedical engineering [2], biological information processing, data classification [6], computer vision, pattern recognition [7,8], robotics and control [2]. ELM generates the input layer (sensory layer) weights and the hidden nodes biases randomly and determines the output layer weights rationally by solving a generalized inverse matrix. The study of Huang et al. [9,10], substantiated that SLFNs with randomly generated hidden node parameters and with radial or additive basis function hidden

* Corresponding author.
E-mail address: adeniyipee@live.com (P.A. Alaba).

https://doi.org/10.1016/j.neucom.2019.03.086
0925-2312/