Artificial Intelligence Review (2023) 56:13691–13742
https://doi.org/10.1007/s10462-023-10478-4

A Review of multilayer extreme learning machine neural networks

José A. Vásquez-Coronel[1,2] · **Marco Mora**[2,3] · **Karina Vilches**[2,4]

Published online: 19 April 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract
The Extreme Learning Machine is a single-hidden-layer feedforward learning algorithm, which has been successfully applied in regression and classification problems in different research fields. The traditional algorithm assigns random weights and biases in the hidden layer, and the Moore–Penrose inverse matrix in the regularized least-squares method is adopted to compute the weights of the output layer. Training speed, generalization ability, and robustness are the advantages that characterize this algorithm, but it has some shortcomings in solving highly nonlinear problems. The scientific community adopted non-iterative multilayer learning models as effective and efficient measures, starting with the Multilayer Extreme Learning Machine, which incorporates an unsupervised extreme learning Autoencoder into its architecture for feature mapping. Since the literature does not present an in-depth review of non-iterative multilayer models, this paper focuses on a current description of the evolution of multilayer models, which are grouped into random mappings, kernel-correntropy strategies, and conditional probability techniques. In addition to showing the mathematical fundamentals of each model, a list of databases widely used in training multilayer networks is included. Finally, we present a class of fast iterative algorithms called Shrinkage-Thresholding, which solve the minimization problem associated with an Autoencoder.

Keywords Multilayer extreme learning · Autoencoder · Regression and classification · Composite optimization problems · Fast gradient variants

1 Introduction

The Extreme Learning Machine (ELM) is a single-hidden-layer feedforward neural network (Huang et al. 2004), which corrects some training limitations of artificial neural networks (Gupta and Sexton 1999; Ding et al. 2011). The ELM algorithm maps the input data to a representative feature space by randomly assigning weights and biases

Marco Mora and Karina Vilches have contributed equally to this work.

✉ Marco Mora
mmora@ucm.cl

Extended author information available on the last page of the article