Neural Computing and Applications (2022) 34:5923–5938
https://doi.org/10.1007/s00521-021-06727-8

ORIGINAL ARTICLE

Quick extreme learning machine for large-scale classification

Audi Albtoush[1] · Manuel Fernández-Delgado[1] · Eva Cernadas[1] · Senén Barro[1]

Received: 14 June 2021 / Accepted: 2 November 2021 / Published online: 18 January 2022

Abstract
The extreme learning machine (ELM) is a method to train single-layer feed-forward neural networks that became popular because it uses a fast closed-form expression for training that minimizes the training error with good generalization ability to new data. The ELM requires the tuning of the hidden layer size and the calculation of the pseudo-inverse of the hidden layer activation matrix for the whole training set. With large-scale classification problems, the computational overload caused by tuning becomes not affordable, and the activation matrix is extremely large, so the pseudo-inversion is very slow and eventually the matrix will not fit in memory. The quick extreme learning machine (QELM), proposed in the current paper, is able to manage large classification datasets because it: (1) avoids the tuning by using a bounded estimation of the hidden layer size from the data population; and (2) replaces the training patterns in the activation matrix by a reduced set of prototypes in order to avoid the storage and pseudo-inversion of large matrices. While ELM or even the linear SVM cannot be applied to large datasets, QELM can be executed on datasets up to 31 million data, 30,000 inputs and 131 classes, spending reasonable times (less than 1 h) in general purpose computers without special software nor hardware requirements and achieving performances similar to ELM.

Keywords Extreme learning machine · Classification · Large-scale datasets · Model selection

1 Introduction

The extreme learning machine (ELM) is a neural network model [21] related to the SVM [2, 31] that has been widely applied for classification, regression and time series prediction. Several variants of ELM have been proposed in the literature, such as twin ELM [42] for classification, that uses two non-parallel hyperplanes that simultaneously minimize the distance to one class while keeps away from the other class. The online sequential ELM (OS-ELM) allows to learn pattern-by-pattern or chunk-by-chunk [30] and has also been combined in ensembles [27]. In [50], the augmented OS-ELM is used for classification and regression of noncircular quaternion signals that provide a convenient way to represent 3D and 4D signals. Other ensembles of ELM have been proposed in [39], using negative correlation learning, and [34], a committee of voting ELMs trained with different bootstrap training samples for road lane landmark detection. The problems of ELM with low column rank data are solved by the effective ELM [44] that has been successfully applied to image classification [5]. Low-rank matrix factorization is also used in the ELM autoencoder [38] to learn optimal low-dimensional features for the same application. The weighted ELM [51] is oriented to classification of imbalanced data that has been applied in [20] for discriminative data clustering alongside with linear discriminant analysis and K-means, a clustering method that was used with ELM in [32] to forecast sales of computer servers. The ELM has also been combined with other paradigms like fuzzy logic, used in [17] to find the optimal ELM hyper-parameters (size of the hidden layer), and particle swarm optimization (PSO) for feature selection and hidden layer size estimation in the sleep stage classification over electrocardiogram signal [41]. The ELM and PSO are also combined in [28] with boosting for electric consumption time series forecasting. In [25], the ELM is applied instead of genetic algorithms for symbolic regression in system identification. Deep neural networks have also been combined with ELM

✉ Manuel Fernández-Delgado
manuel.fernandez.delgado@usc.es

[1] Centro Singular de Investigación en Tecnoloxías Intelixentes da USC (CiTIUS), University of Santiago de Compostela, 15782 Santiago de Compostela, Spain