Information Sciences xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Information Sciences

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

Random vector functional link network for short-term electricity load demand forecasting

Q1 Ye Ren[a], P.N. Suganthan[a,*], N. Srikanth[b], Gehan Amaratunga[c]

[a] School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Ave., 639798 Singapore, Singapore
[b] Energy Research Institute @ NTU (ERI@N), #06-04 CleanTech One 1 CleanTech Loop, 637141 Singapore, Singapore
[c] Department of Engineering, University of Cambridge, 9 JJ Thomson Ave., Cambridge CB3 0FA, UK

ARTICLE INFO

Article history:
Received 13 March 2015
Revised 8 November 2015
Accepted 29 November 2015
Available online xxx

Keywords:
Random weights
Random vector functional link
Neural network
Time series forecasting
Electricity load demand forecasting

ABSTRACT

Short-term electricity load forecasting plays an important role in the energy market as accurate forecasting is beneficial for power dispatching, unit commitment, fuel allocation and so on. This paper reviews a few single hidden layer network configurations with random weights (RWSLFN). The RWSLFN was extended to eight variants based on the presence or absence of input layer bias, hidden layer bias and direct input–output connections. In order to avoid mapping the weighted inputs into the saturation region of the enhancement nodes' activation function and to suppress the outliers in the input data, a quantile scaling algorithm to re-distribute the randomly weighted inputs is proposed. The eight variations of RWSLFN are assessed using six generic time series datasets and 12 load demand time series datasets. The result shows that the RWSLFNs with direct input–output connections (known as the random vector functional link network or RVFL network) have statistically significantly better performance than the RWSLFN configurations without direct input–output connections, possibly due to the fact that the direct input–output connections in the RVFL network emulate the time delayed finite impulse response (FIR) filter. However the RVFL network has simpler training and higher accuracy than the FIR based two stage neural network. The RVFL network is also compared with some reported forecasting methods. The RVFL network overall outperforms the non-ensemble methods, namely the persistence method, seasonal autoregressive integrated moving average (sARIMA), artificial neural network (ANN). In addition, the testing time of the RVFL network is the shortest while the training time is comparable to the other reported methods. Finally, possible future research directions are pointed out.

1. Introduction

Short-term electricity load demand forecasting targets to predict the future load demand ranging from a few minutes to 1 day ahead [2,19]. In a contemporary competitive energy market, as power systems become deregulated and involve distributed power sources, electricity demand becomes more complicated in both temporally and spatially. Since electricity load demand is closely related to the electricity price, load forecasting plays an important role in the energy market

Q2 * Corresponding author. Tel.: +65 6705404; fax.: +65 67933318.
E-mail addresses: re0003ye@ntu.edu.sg (Y. Ren), epnsugan@ntu.edu.sg (P.N. Suganthan), nsrikanth@ntu.edu.sg (N. Srikanth), gaja1@hermes.cam.ac.uk (G. Amaratunga).
URL: http://www.ntu.edu.sg/home/epnsugan (P.N. Suganthan)

http://dx.doi.org/10.1016/j.ins.2015.11.039
0020-0255/

[19]. From the energy generation point of view, load forecasting is beneficial for electric utility operations such as power dispatching, unit commitment, fuel allocation and network diagnosis [2]. The significance of forecasting error with respect to power system operational cost was stated in [2,19]: 1% increase in error can result in wastage of millions of dollars. Therefore, improving the forecasting accuracy can significantly reduce the power system operational cost.

Fig. 3a in Section 5.1 shows a small fraction of load demand time series (TS). From the plot, we can observe the non-linearity of the load demand TS because it is usually affected by a lot of exogenous factors such as weather, special occasions, economy, and so on [9], thereby making the modeling of load demand TS stochastic. However, it also reveals strong seasonal components such as daily, weekly, monthly or yearly repetitions. These properties, on the other hand, makes the modeling somewhat deterministic [13].

In the literature, there are numerous methods for load forecasting and they can be categorized into (i) statistical methods, (ii) machine learning methods and (iii) hybrid methods. Statistical methods use mathematical equations and statistical theories to model and forecast the TS: it is fast but with limitations (distributions, stationarity, linearity, etc.). The commonly used statistical methods are exponential smoothing [32] and autoregressive moving average (ARMA) and so on [33]. Machine learning methods use supervised learning to model a portion of historical data (training data). Once the model is optimized based on the training data, it can be applied to forecast the unknown future data (testing data). Some well-known machine learning methods are artificial neural networks (ANN) [19,23], support vector regression (SVR) [12,18,22], etc. Combining more than one forecasting method to form a single forecasting method is known as a hybrid method. The combination can be sequential or parallel. For a sequentially combined hybrid method, one forecasting method's output is another forecasting method's input, and the final output is the output of the last forecasting method. For a parallel combined hybrid method, the training data is either bootstrapped to multiple training data sets or decomposed to a collection of training data sets first. Subsequently, each data set is trained by an individual method. Finally, the trained forecasting methods are used to predict the future data and the outputs are aggregated to form the final prediction. Negative correlation learning [3,9] and wavelet [17,20,25,35,36] and Empirical Mode Decomposition (EMD) [14,16,30] based methods are in this category.

ANN is a popular method among the machine learning based load forecasting methods. There are several ANN based forecasting methods reported in the literature [15,28,34]. The usual way to train the ANN is through back propagation (BP). But in [31], a single hidden layer neural network with random weights (RWSLFN) was reported with random weight assignment between input and hidden layers and least square estimation on the output weights as the training method. In addition, we will discuss another type of ANN which is called random vector functional link network (RVFL network) [5,6,27]. Although it is similar to RWSLFN reported in [31], it has direct input–output connections between input and output neurons. This paper will compare these neural networks and their forecasting capability in short-term load demand forecasting.

The remaining of the paper is organized as follows: Section 2 reviews different variants of RWSLFN/RVFL network; Section 3 states the essentials of the experiment configurations, feature selection methods, parameter optimization procedure and error measures. Section 4 assesses the performance of different variants on six generic TS data; Section 5 discusses the performance of RWSLFN/RVFL network with eight variants on 12 load demand TS, and Section 6 concludes the paper and suggests possible future research directions.

2. Review of related neural network structures

ANN is one of the most popular algorithms in machine learning. ANN can be categorized according to the connectionism: feedforward and recurrent. The main difference between feedforward neural network (FNN) and recurrent neural network (RNN) is that the FNN does not have connections from the hidden layers or output layer back to the hidden layer. RNN not only has the connections that the FNN has but also has the connections from hidden layers or output layer back to the input layer. The recurrent connections are usually delayed so that the RNN can exhibit better dynamic temporal behavior. There are two typical RNN architectures: Elman RNN and Jordan RNN. If there exists connections from the hidden layer to the input layer, it is an Elman RNN [8]. If the feedback connections are from the output layer to the input layer, it is a Jordan RNN [8].

In the literature, FNN is more widely used. It usually has one input layer, one output layer and several hidden layers as shown in Fig. 1. The neurons in the adjacent layers are connected. But there is no interconnection of neurons within the same layer or across non-adjacent layers.

In the input layer, each neuron i_m, $m \in \{1, \ldots, M\}$ takes a feature of the input vector and passes to the hidden layer neurons. Each neuron in the nth hidden layer h_{n,k_n} $n \in \{1, \ldots, N\}$, $k_n \in \{1, \ldots, K_N\}$ is formed by a nonlinear weighted sum of the outputs of the input layer or the preceding hidden layer (except the last hidden layer):

$$h_{1,k_1} = f\left(\sum_{m=0}^{M} w_{m,k_1} i_m\right), \quad \forall k_1 \in \{1, \ldots, K_1\} \tag{1}$$

$$h_{n,k_n} = f\left(\sum_{k_{n-1}=0}^{K_{n-1}} w_{k_{n-1},k_n} h_{n-1,k_{n-1}}\right), \quad \forall k_n \in \{2, \ldots, K_n\} \tag{2}$$

where $f(\cdot)$ is a nonlinear activation function, $w_{0,k_n} = 1$, $k_n \in \{1, \ldots, K_n\}$ denotes the input layer and hidden layer biases, M is the number of input layer neurons, N is the number of hidden layers, K_n is number of nth hidden layer neurons,

Please cite this article as: Y. Ren et al., Random vector functional link network for short-term electricity load demand forecasting, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2015.11.039