Electric Power Systems Research 125 (2015) 109–115

Contents lists available at ScienceDirect

Electric Power Systems Research

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

Improved short-term load forecasting using bagged neural networks

A.S. Khwaja, M. Naeem, A. Anpalagan*, A. Venetsanopoulos, B. Venkatesh

Department of Electrical and Computer Engineering, Ryerson University, Toronto, Canada

ARTICLE INFO

Article history:
Received 1 January 2015
Received in revised form 12 March 2015
Accepted 30 March 2015
Available online 18 April 2015

Keywords:
Smart grid
Short-term load forecasting
Non-linear regression
Artificial neural networks
Bagging
Power systems

ABSTRACT

In this paper we present improved short-term load forecasting using bagged neural networks (BNNs). The BNNs consist of creating multiple sets of data by sampling randomly with replacement, training a neural network on each data set, and averaging the results obtained from each trained neural network. The bagging process reduces estimation errors and variation range of errors compared to using a single neural network for load forecasting. Examples with real data show the effectiveness of our proposed techniques by demonstrating that using BNNs can reduce load forecasting errors, compared to various existing techniques.

1. Introduction

Load refers to a device or a set of devices, which are consuming energy from the power system networks. The pattern of this energy use varies with respect to time [1]. Forecasting this energy load has been receiving a lot of attention for the last many decades. An accurate short term load forecast, which usually predicts one hour to a week's future load, is necessary for scheduling of functions such as coordination of power generation. This coordination can be used for establishing the hourly schedules of generation resources to minimize the operating cost of the system [2].

Additionally, short-term load forecasting is an important component of smart grids for cost savings and ensuring a continuous flow of electricity supply [3]. One of its usages can be for demand side management, where customers are given incentives to reduce their energy consumption or modify their energy usage pattern. This can require load forecasting so that the load patterns can be predicted in advance. It is also known that making the energy consumption publicly known can help reduce the load in both public buildings and households [4].

Economic and reliable operations of an electric utility depend greatly on how accurate the load forecast is. The costs of electricity buying and selling depend on the decision taken about the current electricity demand [5]. The load dispatcher at the main dispatch center must have an idea of the load pattern in advance so that sufficient generation may be dedicated to meet the customer requirements. Overestimation of future load pattern can cause starting of extra generating units that will lead to an unnecessary increase in the reserve and operating costs. Underestimation of future load forecasts will result in failure to provide the required operating reserve and stability to the system, which may lead to a breakdown of the power system network [1]. Similarly, the use of smart prepaid meters also requires the ability to accurately predict the amount of energy consumed on a daily basis for each end user [6], so that sufficient units of energy may be purchased in advance.

Improving accuracy of short-term load forecasting has a significant effect on power systems operations, as economy of operations and control of power systems may be very sensitive to forecasting errors. The presence of forecasting errors can increase the operating cost and depending on the type of power system, high forecasting errors can lead to scheduling that can either lead to a high cost or has high risk of system breakdown [7]. In general, forecasts with approximately 10% mean absolute percent error (MAPE) can be obtained easily; however, the cost of this error is significantly high and research efforts to help reduce it by a few percentage points would be considered as significant contributions [8]. If the accuracy of short-term load forecasts can be improved, it will result in significant financial savings for utilities and co-generators [9,10].

Load pattern is affected by several factors, such as weather, time, characteristics of population, economy, electricity prices, geographical conditions and type of consumers [11]. Many load forecasting techniques have been proposed and applied successfully to predict the load patterns. Due to the importance of load forecasting in reducing costs, a wide variety of models have been proposed in the last two decades. The conventional models for load forecasting are designed based on the relationship between load power and factors influencing load power. Some of these conventional methods include [12]:

* Corresponding author. Tel.: +1 416 979 5000; fax: +1 416 979 5280.
E-mail address: alagan@ee.ryerson.ca (A. Anpalagan).

http://dx.doi.org/10.1016/j.epsr.2015.03.027
0378-7796/

- Time series approach: The autoregressive, moving average and mixed autoregressive and moving average models are some of the widely used models in the time series approach. The time series approach may perform poorly for data with inherent special events such as holidays [11].
- Multi-variate regression approach: In multi-variate regression, the load is presented as a linear combination of several factors with different coefficients [13]. In [2], a multiple linear regression method is employed for load forecasting. In [5], regression tree method is utilized for load forecasting. In [14], a regression method is used alongwith a time series model to predict the daily electricity load.
- State estimation approach: In this case, the load may be seen as observations of an unknown state, where both observations and states are described by equations. The state may be estimated using Kalman filter or its other variants.

The conventional load forecasting methods have the advantage that load pattern can be forecasted using a simple prediction model. However, due to the nonlinear relationship between load pattern and the factors influencing it, it is difficult to find this nonlinear relationship using conventional methods [11]. The accuracy of load forecasting can be increased with the help of artificial intelligence techniques [1].

A few approaches based on artificial intelligence are:

- Fuzzy logic: A fuzzy logic is a set of "IF–THEN" logic statements. For the problem of load forecasting, historical load and weather data are converted into fuzzy information to build a model and then the estimates of forecast values are obtained [1,15]. In [9], fuzzy neural networks are used for load forecasting and simulated annealing is used to select the parameters of the fuzzy neural networks. In [12], short-term load forecasting is carried out using artificial neural networks (ANNs) combined with a fuzzy system.
- Support vector machines: Support vector machines (SVM) use training data and solve an optimization problem. This problem first uses a nonlinear transformation to map the training input data to a higher dimensional space, then constructs a linear model in space [16]. Load forecasting using SVM has been presented in [17]. An adaptive model based on SVM is proposed in [18]. This model optimizes required input parameters using an optimization algorithm and utilizes a feature selection algorithm to select model inputs. This method is referred to as strategic, seasonality-adjusted, support vector regression based model (SSASVR).
- Artificial neural networks: The ANNs are composed of nodes that are artificial neurons. These artificial neurons are computational models inspired from natural neurons. Efforts to forecast load patterns using ANN have shown promising results [19,20]. In [7], the authors presented a novel short-term load forecasting approach, which analyzed the weather sensitive components of the loads using wavelet transform. A similar method has been presented in [21] and is known as similar day-based wavelet neural network (SIWNN). Ref. [22] presents a hybrid unsupervised supervised ANNs, where the data are first classified and then uses for creating a load forecasting model using ANNs.

The most popular artificial intelligence based load forecasting methods are SVM and ANNs. However, despite their accuracy, artificial intelligence methods have a number of inconveniences such as difficulty in parameter selection, non-obvious selection of variables, over-fitting, etc. They also require a lot of historical data to learn the patterns present in the data [4].

In this paper, we present improved short-term load forecasting using bagging technique. This technique is based on a combination of a large number of load forecasting techniques, where each technique uses a bootstrap sample of the training set used to create the load forecasting model. We use ANNs as the load forecasting technique. The importance of this work is due to the fact that according to the best of our knowledge, bagging techniques have not been used for load forecasting using ANNs. We show that this technique leads to a reduction in forecasting error. It also reduces the variation in forecasting accuracy that occurs when ANNs are used for load forecasting.

This paper is organized as follows: Section 2 presents an overview of load forecasting using ANNs. Section 3 presents the load forecasting problem and bagged neural networks (BNNs). Section 4 presents results, and conclusion is given in Section 5.

2. Artificial neural networks

The ANNs are trained by adapting a network and comparing the output obtained with the input training and target data. The training is carried out until the network output matches the target data. Feedforward neural networks permit the signals to travel from input to output in only one way, which means that the output of any layer does not affect the same layer. The ANN consists of an input layer and an output layer. The layer in between these two layers is the hidden layer. The neural networks use weights for each input variable and a bias that acts as a threshold to produce outputs. They further use learning algorithms to fine tune the weights and biases of a network such that the output can be obtained from the weights and biases. The learning algorithm should be such that a small change in the weights and biases leads to a small change in the output. A sigmoid function is chosen as a function that calculates output using weights and biases as it shows a great similarity to real neurons:

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{1}$$

Let the feature vector that is input at layer 1 be denoted as $\boldsymbol{x}$. Let $w^l_{j,k}$ denote the weight for the connection from the kth neuron in the $(l-1)$th layer to the jth neuron in the lth layer. Let the total number of neurons in the $(l-1)$th and lth layers be K and J, respectively. Let L be the total number of layers, b^l_j denote the bias of the jth neuron in the lth layer and a^l_j denote the activation of the jth neuron in the lth layer. It is denoted as:

$$a^l_j = \sigma\left(\sum_k w^l_{j,k} a^{l-1}_k + b^l_j\right), \quad j = 1, \ldots, J \tag{2}$$

Given the last layer, the activation function for the jth neuron is computed as:

$$a^L_j = \rho\left(\sum_k w^L_{j,k} a^{L-1}_k + b^L_j\right), \quad j = 1, \ldots, J \tag{3}$$

where $\rho(z) = z$ is a linear function. The weight $w^l_{j,k}$ is an entry of the weight matrix $\boldsymbol{W}^l$ defined as:

$$\boldsymbol{W}^l = \begin{bmatrix} w^l_{1,1} & w^l_{1,2} & \cdot & \cdot & w^l_{1,K} \\ w^l_{2,1} & w^l_{2,2} & \cdot & \cdot & w^l_{2,K} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ w^l_{J,1} & w^l_{J,2} & \cdot & \cdot & w^l_{J,K} \end{bmatrix}_{J\times K} \tag{4}$$