Deep Belief Networks for Electricity Price Forecasting

Aleksandra Dedinec*, Aleksandar Dedinec**
* Faculty of Computer Science and Engineering, Skopje, Macedonia
** Research Center for Energy and Sustainable Development
Macedonian Academy of Sciences and Arts, Skopje, Macedonia
aleksandra.kanevche@finki.ukim.mk, dedinec@manu.edu.mk

***Abstract*— In this paper, one of the aspects of the smart grids is analyzed. This aspect includes the utilization of the large amount of available digital information for creating smart models for planning and forecasting. The latest and new achievements in the field of machine learning are used for that purpose. Specifically, models based on deep belief networks are developed within this paper and it is examined whether these models may be applied for electricity price forecasting. For that purpose, the hourly data of the prices of the power exchanges in the region of Southeast Europe are used. The obtained results present the advantages of the developed models based on deep belief networks, compared to the traditional neural networks, when applied to electricity price forecasting. To this end, the mean average percent error of the deep belief network model is less than the minimum error of the traditional neural network model in each of the analyzed datasets.**

Keywords: deep belief network, electricity price forecasting, power exchange, neural networks

I. Introduction

One of the aspects that the smart grids involve is the utilization of the large amount of available digital information for creating smart models for planning and forecasting, which could be based on the latest and new achievements in the field of machine learning. With the increasing amount of data generated on a daily basis, as well as the rapid development of computer systems, especially in terms of processing power using the GPGPU-General Purpose Graphical Processing Unit, a new field in the machine learning is developed, and it is called deep learning. Modern deep-learning architectures, which include deep neural networks, deep belief networks and recurrent neural networks have been applied in a variety of areas with great success, such as in image classification, speech recognition, natural language processing and bioinformatics.

On the other hand, power systems are considered complex systems composed of interdependent components for production, transmission, storage and consumption of electricity, in which the balance between production and consumption must be constantly maintained. But, the ability to achieve this goal is very closely linked to the exact predictions and planning of the system. On the demand side, forecasting and planning is a nontrivial task due to various factors affecting consumers, such as outside temperature (where any cold or hot wave can cause a significant change in consumption and thus the price), work habits, everyday activities of people and also the market conditions that largely define the behavior of consumers. In addition, liberalization of markets and the development of power exchanges also plays a significant role in the development and behavior of the entities that make up the power systems.

In this paper, the goal is to answer the question of whether the latest methods in the field of deep learning, and concretely, the deep belief networks can lead to creation of models that will help in achieving more precise forecasting results in the field of energy systems. Particular emphasis is placed on the analysis of whether these models can be applied for short-term forecasting of the electricity prices on the power exchanges.

The electricity price depends on many variables, except on the historical data for the prices, such as the seasonality (daily, weekly or annually), system load, weather parameters, working or non-working days etc. Therefore, its forecasting represents a complex problem, which is non-linear and many different models have been developed to solve it. These models, used for electricity price forecasting, mainly belong to the following two groups: statistical [1-2] and artificial intelligence based models [3-7]. The most widely used methods are those in the field of artificial intelligence and are based on artificial neural networks applied in this area since the 1980s [8-9]. It is proven that the neural networks can approximate any complex function, if sufficient number of hidden layers and hidden nodes in the layers are included, which makes them suitable for the complex problem of electricity price forecasting. However, by using the backpropagation algorithm in order to calculate the gradient needed for determination of the weights used in the neural network, there are certain problems that may arise. These problems involve bad scalability – the model cannot be easily used for larger networks, which includes the time issue, the model depends on random initialization of the weights and biases and therefore it can get stuck in local optimum. As a solution for these problems, in this paper the model od deep belief networks (DBN) is used [10-11]. In this approach, the initial values of the weights are calculated using layer-by-layer unsupervised learning method, after which fine tuning is applied using a standard supervised method, which includes backpropagation. Compared to random initialization, the obtained initial parameters of the network are closer to the optimal solution, which also results in faster convergence. This led to the successful application of the deep belief networks model in many diverse areas. However, very few studies have approached the subject of using deep belief networks in electricity price forecasting.

As an application of the model, the data for electricity prices on the day-ahead power exchanges are used. Particularly, the possibilities for forecasting the prices on the Serbian, Croatian and Bulgarian power exchanges (SEEPEX, CROPEX and IBEX) were analyzed for the hourly data for 2016.

The paper is organized as follows. After the introduction, short description of the methodology used for electricity price forecasting is presented. The results and corresponding discussion are given in the following section, and the last section concludes the paper.

II. METHODOLOGY

As mentioned in the introduction, neural networks are used as a basic method for electricity price forecasting in this paper. For that purpose, the model developed in [12] is used. In this paper, a deep belief network model is developed and integrated into the traditional neural network model.

A deep belief network is a kind of deep learning architecture, that represents a probabilistic, generative model that can learn to probabilistically reconstruct its inputs. It is composed of multiple simple learning modules and in this paper, each pair of layers of the neural network is pre-trained by using restricted Boltzmann machine (RBM). Restricted Boltzmann machines represent a special type of generative energy based models that can learn a probability distribution over its set of inputs [11]. An RBM has a single layer of hidden units which have undirected, symmetrical connections to a layer of visible units. The name Restricted Boltzmann machines comes because of the restriction that their neurons must form a bipartite graph. There are no connections between the nodes in the hidden layer, so the main advantage of the RBMs is that the hidden units are conditionally independent given the visible states.

During the training phase, gradient descent is used, where the weights are updated according to the following equation:

$$\Delta w_{ij} = \epsilon \frac{\partial \log(p(v))}{\partial w_{ij}} \tag{1}$$

Where ϵ is the learning rate and $p(v)$ is the probability that the network assigns to a visible vector, v which is represented by the following equation:

$$p(\mathbf{v}) = \frac{1}{Z}\sum_{\mathbf{h}} e^{-E(\mathbf{v},\mathbf{h})} \tag{2}$$

Where Z is a partition function, which is a sum of $e^{-E(\mathbf{v},\mathbf{h})}$ over all possible configurations, and is used for normalizing:

$$Z = \sum_{\mathbf{v},\mathbf{h}} e^{-E(\mathbf{v},\mathbf{h})} \tag{3}$$

And $E(\mathbf{v},\mathbf{h})$ represent energy of the joint configuration $(\mathbf{v},\mathbf{h})$ of the visible and hidden units:

$$E(\mathbf{v},\mathbf{h}) = -\sum_{i\in visible} a_i v_i - \sum_{j\in hidden} b_j h_j - \sum_{i,j} v_i h_j w_{ij} \tag{4}$$

where v_i, h_j are the binary states of visible unit i and hidden unit j, a_i, b_j are the biases and w_{ij} is the weight between them. A lower energy indicates that the network is in a more desirable state. This energy function is used to calculate the probability that is assigned to every possible pair of a visible and a hidden vector.

The gradient or the derivative of the log probability of a training vector with respect to a weight has a simple form, so the weights can be calculated using the equation:

$$\Delta w_{ij} = \epsilon(\langle v_i h_j\rangle_{data} - \langle v_i h_j\rangle_{model}) \tag{5}$$

where $\langle \ldots \rangle_p$ represents averages with respect to distribution p.

For calculating the $\langle v_i h_j\rangle_{model}$ part of eq. (10), the following training process is used [11]. First, the visible units v_i are set to be equal to the training sample. Afterwards, the hidden states h_j are calculated according to the following equation:

$$P(h_i = 1|\mathbf{v}) = \sigma(b_i + \sum_i v_i w_{ij}) \tag{6}$$

where $\sigma(x)$ is the logistic sigmoid function. One step "reconstruction" of the visible v_i' and hidden h_j' units is produced after repeating the process once more, using (6) and the following equation:

$$P(v_i = 1|\mathbf{h}) = \sigma(a_i + \sum_j h_j w_{ij}) \tag{7}$$

Figure 1 presents the methodology used in this paper for layer-by-layer unsupervised learning method, when using three layers. RBM1 is composed of the input and the hidden layer of the neural network, and RBM2 is composed of the hidden and the output layer of the neural network. The values of the units in the hidden layer obtained from RBM1 are used as values for the input layer in RBM2. The model can accordingly be expanded for networks with more hidden layers.

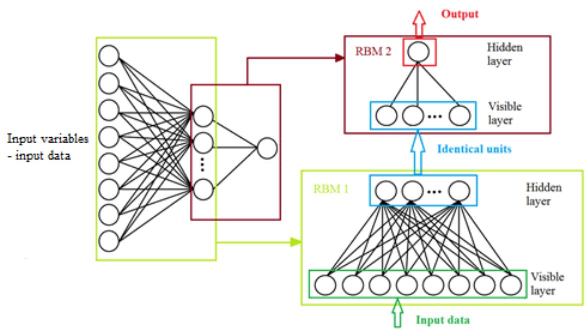

Figure 1. Example of deep belief network with three layers [13]

A. Selection of the input variables of the neural network

As the selection of input variables is one of the most important steps when developing a traditional neural network model, so it is for the deep belief network model. When forecasting time series variables, such as the electricity price forecasting, there are two possibilities. One is to use only historical data, and the other is to also include other variables on which the variable that is being forecasted depend on. In this paper, we have used the second approach, and based on a detail analysis of the input data, we have selected the following input variables (which are used for both – the traditional neural network model and the deep belief model):

- Hour of day
- Day of week
- Holiday flag
- Previous day's average price
- Price for the same hour of the previous day
- Temperature
- Price for the same hour-day combination of the previous week

III. RESULTS

Firstly, the optimal parameters of the deep belief network model are determined, using which the best results are obtained. For the IBEX data, the optimal number of layers is 4 with 24 neurons in each of them. The optimal results for the CROPEX data are obtained for 3 layers with 6 neurons each, and for the SEEPEX each layer has 20 neurons, and there are 3 layers in total. As it can be noticed, there are three, two and two restricted Boltzmann machines for each of the three datasets (IBEX, CROPEX and SEEPEX), respectively. Additionally, the obtained optimal values for the number of epochs, size of mini-batch, learning rate and momentum are presented in Table I.

TABLE I.
PARAMETERS USED FOR THE DEEP BELIEF NETWORK FOR ELECTRICITY PRICE FORECASTING

	IBEX	CROPEX	SEEPEX
Total number of layers	4	3	3
Number of neurons in the hidden layers	24	6	20
Number of epochs	[1,1,1]	[1,1]	[1,1]
Size of mini-batch	[1,1,1]	[1,1]	[1,1]
Learning rate	[0.01,0.01,0.01]	[0.02,0.02]	[0.94,0.94]
Momentum	[0,0,0]	[0,0,0]	[0,0]

Using these parameters, a comparison between the actual data and the forecasted data using the developed model for the electricity prices for SEEPEX is presented on Figure 2. It can be noted that the two curves match quite well. The Croatian power exchange (CROPEX) is more unpredictable, and has more peaks in the prices, that are not predicted by the DBN, as presented in Figure 3. The results for IBEX (Figure 4) show that there are a lot of peaks of the price in the testing period that are not predicted, mainly because these peaks were not present in the data for the training period. These results are very similar to the results of the price forecasting using the traditional neural network [12]. However, in order to compare more precisely the results obtained from the two models, the mean absolute percent error (MAPE) was calculated.

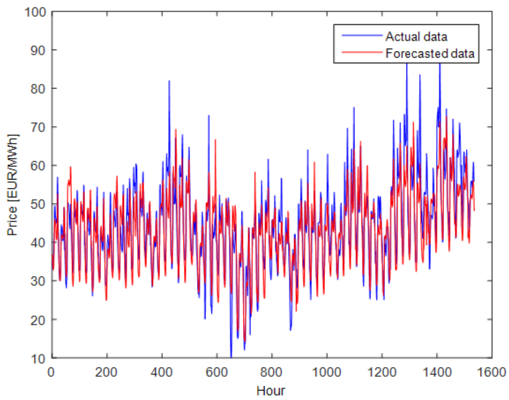

Figure 2. Actual and forecasted electricity prices for SEEPEX using DBN

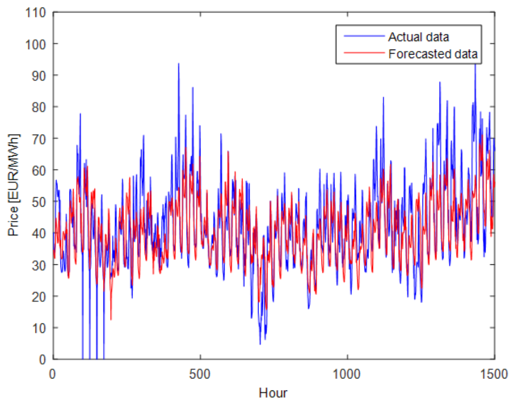

Figure 3. Actual and forecasted electricity prices for CROPEX using DBN

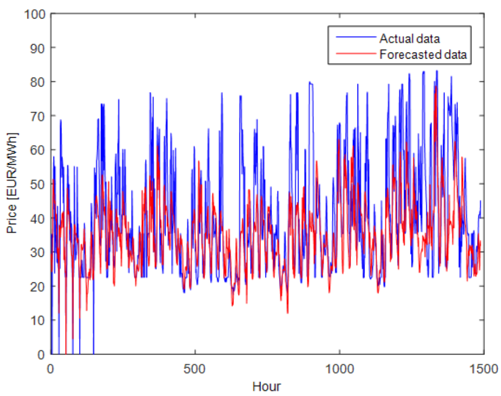

Figure 4. Actual and forecasted electricity prices for IBEX using DBN

The results for the MAPE for each of these cases are presented on Figure 5, Figure 6 and Figure 7. The MAPE in the traditional neural networks is different for each execution (as a result of the random initialization of the weights and biases) and in this case the best results that can be obtained from 100 runs were presented as a result of the neural network. That result is compared to the forecasting derived from the deep belief network. As it can be noted, in each of the three data sets there is an improvement in the forecasting when using deep belief networks. The biggest improvement was on the Serbian power exchange (SEEPEX), so the MAPE of 9.28% is decreased to 9.07%. For the Croatian power exchange (CROPEX) MAPE is decreased from 16.9% to 16.7%, and for the Bulgarian power exchange (IBEX) from 21.5% to 21.4%.

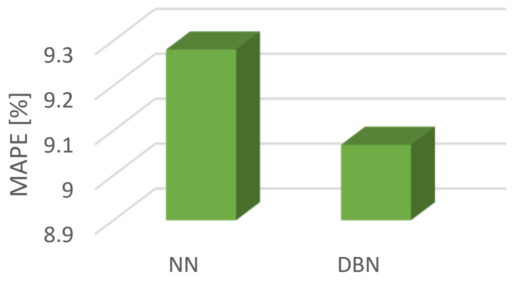

Figure 5. Comparison between the MAPE for the NN and DBN models for SEEPEX

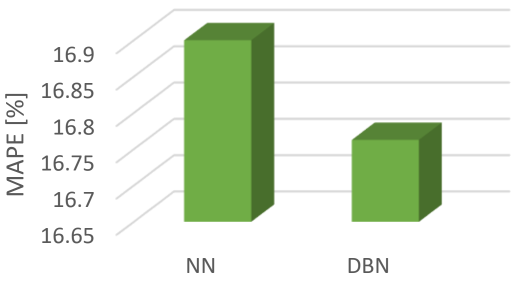

Figure 6. Comparison between the MAPE for the NN and DBN models for CROPEX

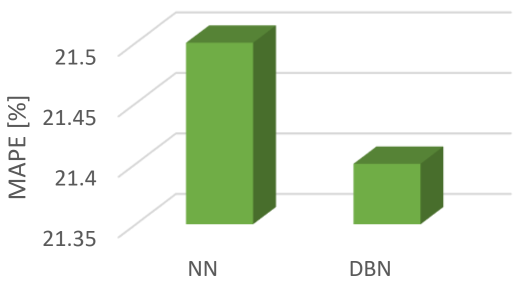

Figure 7. Comparison between the MAPE for the NN and DBN models for IBEX

IV. Conclusion and Future Work

In this paper, deep belief network models were developed for the electricity price forecasting of the day-ahead power exchanges. The results show that for each data set that was analyzed, the mean absolute percent error of the deep belief network model forecasting is smaller than the minimum mean absolute percent error obtained from the traditional neural network model. This shows that the deep belief network model is suitable for the application of electricity price forecasting.

Acknowledgment

This work was partially financed by the Faculty of Computer Science and Engineering at the "Ss. Cyril and Methodius" University.

References

[1] A. Goel, “Regression Based Forecast of Electricity Demand of New Delhi” *Int J Sci Res Publ*, 2014;4:1–7.

[2] SR. Huang, “Short-term load forecasting using threshold autoregressive models,” *Gener Transm Distrib IEE Proceedings-* 1997;144:477–81. doi:10.1049/ip-gtd:19971144.

[3] MSS. Rao, SA. Soman, BL Menezes, P. Chawande, P. Dipti and T. Ghanshyam, “An Expert System Approach to Short Term Load Forecasting for Reliance Energy Limited”, *Mumbai. 2006 IEEE Power India Conf* 2006:763–8. doi:10.1109/POWERI.2006.1632604.

[4] P. Xiong, Y. Dang, T. Yao and Z. Wang, “Optimal modeling and forecasting of the energy consumption and production in China,” *Energy* 2014;77:623–34. doi:10.1016/j.energy.2014.09.056.

[5] C. Hamzacebi and HA Es, “Forecasting the annual electricity consumption of Turkey using an optimized grey model,” *Energy* 2014;70:165–71. doi:10.1016/j.energy.2014.03.105.

[6] X. Qiu, L. Zhang, Y. Ren and PN. Suganthan. “Ensemble Deep Learning for Regression and Time Series Forecasting,” *Comput Intell Ensemble Learn (CIEL)*, 2014 IEEE Symp 2014:1–6.

[7] TS. Mahmoud, D. Habibi, MY. Hassan and O. Bass. “Modelling self-optimised short term load forecasting for medium voltage loads using tunning fuzzy systems and Artificial Neural Networks,” *Energy Convers Manag*, 2015;106:1396–408. doi:10.1016/j.enconman.2015.10.066.

[8] AR. Khan, A. Mahmood, A. Safdar, ZA. Khan and NA. Khan. “Load forecasting, dynamic pricing and DSM in smart grid: A review”. *Renew Sustain Energy Rev*, 2016;54:1311–22. doi:10.1016/j.rser.2015.10.117.

[9] R. Weron. “Electricity price forecasting : A review of the state-of-the-art with a look into the future,” *Int J Forecast* 2014;30:1030–81. doi:10.1016/j.ijforecast.2014.08.008.

[10] GE. Hinton, S. Osindero and YW. Teh. “A fast learning algorithm for deep belief nets,” *Neural Comput* 2006;18:1527–54. doi:10.1162/neco.2006.18.7.1527.

[11] G. Hinton. “A Practical Guide to Training Restricted Boltzmann Machines n.d.” https://www.cs.toronto.edu/~hinton/absps/guideTR.pdf (accessed May 18, 2018).

[12] A. Dedinec and A. Dedinec, “Electricity price forecasting of the South East European power exchanges”, 7th International Conference on Information Society and Technology, Serbia, 2017

[13] A. Dedinec, S. Filiposka, A. Dedinec and L. Kocarev, “Deep belief netwrok based electricity load forecasting: An analysis of Macedonian case”, *Energy*, Volume 115, Part 3 (2016), pp. 1688-1700.