Journal of Signal Processing, Vol.21, No.4, pp.183-186, July 2017

SELECTED PAPER AT NCSP'17

Prediction of Wind Speed Fluctuation Using Deep Belief Network with Ensemble Learning Method

Shogo Yoshida, Hiroshi Suzuki, Takahiro Kitajima and Takashi Yasuno

Graduate School of Advanced Technology and Science, Tokushima University
2–1 Minami-Josanjima, Tokushima 770-8506, Japan
E-mail: {yoshida-s, kitajima}ee.tokushima-u.ac.jp
{suzuki.hiroshi, yasuno.takashi}@tokushima-u.ac.jp

Abstract

This paper describes a prediction method for wind speed fluctuation using a deep belief network (DBN) trained with ensemble learning. In particular, we investigate the usefulness of the ensemble learning for an prediction accuracy improvement of wind speed fluctuation. Bootstrap aggregating (the bagging method), which is a typical algorithm of ensemble learning, has been applied to train the DBN. The prediction result is decided by a majority vote of each DBN output. In addition, two bagging methods with different selection methods of training data have been proposed. These proposed methods have been evaluated from several prediction results by comparison with a conventional method.

1. Introduction

In recent years, wind power generators have been introduced to rapidly provide renewable energy to electric power systems. Since the output of a wind power generator depends on weather conditions, in order to supply a stable electric power, it is necessary to control the output of other power generation equipment according to the output fluctuation of the wind power generator. If the output of the wind power generator can be predicted, stable and efficient operation of the electric power system can be expected.

Our research group has already proposed several prediction systems for wind power generation, such as a hierarchical neural network using the classification of the observed wind speed [1]. However, a prediction error between the predicted wind speed and the observed wind speed (time delay) was confirmed [2]. Therefore, we focused on the timing of the wind speed change and predicted, the wind speed fluctuation to reduce the time delay.

In this paper, we propose a prediction system for wind speed fluctuation using a deep belief network (DBN), which is a deep learning. Deep belief network is a type of neural network and can realize high performances in sound and image recognition fields because of connections between neurons in a number of layers. A DBN has some features: one is a pre-training function by unsupervised learning in forward propagation, the other is the ability to handle many input data in a multi layered structure. We focus on these features and demonstrate that a DBN is an effective means of for predicting wind speed fluctuation.

The generalization capability of neural networks can be enhanced by a DBN. A DBN is trained by a bagging method to the avoid over-training of networks. In addition, the prediction accuracy is expected to improve by using multiple DBNs as weak discriminators. In this method, training data is divided into some data sets, and the prediction result is used to evaluate multiple DBN outputs. The training data of each DBN are generated by two methods. One is simple division in the considered time series. The other is random sampling in the unconsidered time series. The prediction results for wind speed fluctuation were verified in the generation methods of training data by ensemble learning.

2. Prediction System for Wind Speed Fluctuation

2.1 Structure of deep belief network

Figure 1 shows the structure of the DBN that predicts the wind speed fluctuation $\hat{d}_{out}(t+n)$ from time-delayed fluctuations of the wind speed $d_{in}(t)$–$d_{in}(t-11)$ and the wind direction w_D. In order to convert linguistic expressions to numerical expressions, the wind speed fluctuation (increase/decrease) is expressed as a discrete value (1/0). For example, if the predicted value d_{out} is 0.5 or more, then it is classified as an increase (1), on the other hand, if $\hat{d}_{out}$ is less than 0.5, it is classified as a decrease (0).

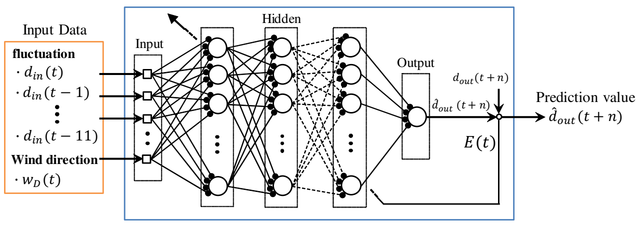


Figure 1: Structure of deep belief network

The number of layers of the DBN is set to seven by trial and error. Each hidden layer has 15 neurons. As training data for wind speed fluctuation, we classified the wind speed change for two modes: increase and decrease. Tables 1 and 2 show the definitions of the input data d_{in} and the training data d_{out}, respectively. Here, a is an integer value within the range of 0 to 11. Wind speed fluctuation is predicted after 10 min. ($n = 1$) and 60 min. ($n = 6$).

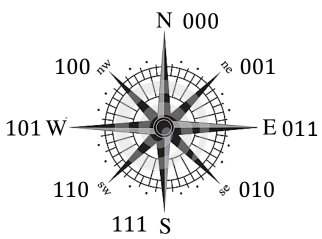
Table 1: Definition of input data d_{in}

	Condition	$d_{in}(t-a)$
Increase	$v(t-a){-}v(t-a-1){\geq}0$	1
Decrease	$v(t-a){-}v(t-a-1){<}0$	0

Table 2: Definition of training data d_{out}

	Condition	$d_{out}(t+n)$
Increase	$v(t+n){-}v(t){\geq}0$	1
Decrease	$v(t+n){-}v(t){<}0$	0

Figure 2 shows the definition of the wind speed direction $w_D(t)$, which is expressed by gray codes (3-bit data) in order to consider position information.



	w_D (gray codes)		
North	0	0	0
NE	0	0	1
East	0	1	1
SE	0	1	0
South	1	1	0
SW	1	1	1
West	1	0	1
NW	1	0	0

Figure 2: Definition of wind speed direction $w_D(t)$

2.2 Restricted Boltzmann machine

Figure 3 shows a graphical depiction of a restricted Boltzmann machine (RBM). An RBM is a probabilistic neural network that can calculate more effectively by restricting the connections of the original Boltzmann machine between the input (visible) layer and hidden layer. The network is constructed by stacking multiple RBMs and implementing unsupervised training (pre-training).

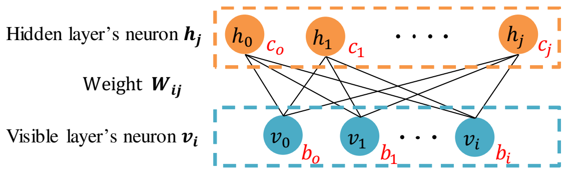


Figure 3: Restricted Boltzmann machine

Feature extraction is carried out by sampling on the basis of the conditional probability represented by Eqs. (1) and (2).

$$P(h_i = 1\,|\boldsymbol{v}\,) = \text{sigmoid}\left(\sum_{j=1}^{m} W_{ij}v_j + c_i\right) \qquad (1)$$

$$P(v_i = 1\,|\boldsymbol{h}\,) = \text{sigmoid}\left(\sum_{i=1}^{n} W_{ij}h_i + b_j\right) \qquad (2)$$

Where W_{ij} stands for the weight between the input layer and hidden layer, v_j is a neuron of the input layer and h_i is a neuron of the hidden layer. b_i is the bias of the input layer and c_j is the bias of the hidden layer. Then, each weight and bias are updated using the following equations:

$$W \leftarrow W + \epsilon\left(\boldsymbol{h}^{(1)T}v - P(\boldsymbol{h}^{(k+1)} = 1\,\Big|\boldsymbol{v}^{(k)})^T\boldsymbol{v}^{(k)}\right) \qquad (3)$$

$$\boldsymbol{b} \leftarrow \boldsymbol{b} + \epsilon\left(\boldsymbol{v} - \boldsymbol{v}^{(k)}\right) \qquad (4)$$

$$\boldsymbol{c} \leftarrow \boldsymbol{c} + \epsilon\left(\boldsymbol{h}^{(1)} - P(\boldsymbol{h}^{(k+1)} = 1\,\Big|\boldsymbol{h}^{(k)})\right) \qquad (5)$$

where ϵ is the learning rate of pre-training, k is the number of times of Gibbs sampling, which is performed on the basis of Eqs. (1) and (2). Empirically, it is assumed that omit k is set to 1.

2.3 Ensemble learning

Ensemble learning is a machine learning method that enhances the accuracy by combining multiple weak discriminators (DBNs). Simple training models are constructed from different samples and combined by some method. Therefore, the training model can achieve to both versatility and accuracy.

In this paper, a bagging method is applied to create multiple data sets and multiple weak learning machines are trained by these data sets.

The multiple weak DBNs are combined by a majority decision of each DBN output, then the prediction result of wind speed fluctuation is decided. The prediction of wind speed fluctuation using ensemble learning is evaluated by two methods: a simple majority decision and an average evaluation decision. The average evaluation decision is defined by

$$y = \frac{1}{M}\sum_{m=1}^{M} d_{out}^{(m)}(t+n) \qquad (6)$$

$$\text{PredictionResult} = \begin{cases} 1 & (y \geq 0.5) \\ 0 & (y < 0.5) \end{cases} \qquad (7)$$

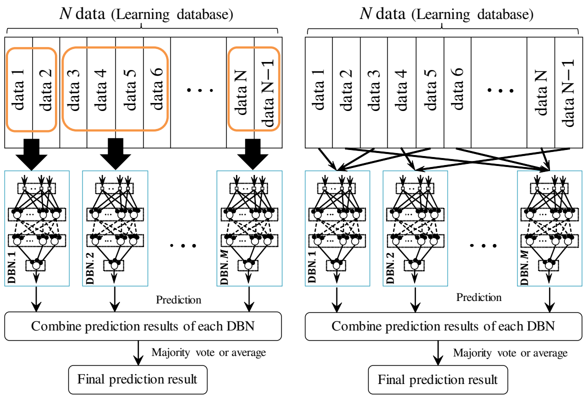
where M is the number of weak learning machines (DBNs), $d_{out}^{(m)}(t+n)$ is the output of each DBN and y is the average outputs of each DBN. In this paper, we propose two methods of creating data sets as shown in Figs.4(a) and 4(b).

2.3.1 Simple division model

In the simple division model, N training data which are decided randomly are divided into M DBNs. The seasonality and time dependence of wind speed fluctuation are confirmed from each of the divided training data holding the time series information.

2.3.2 Random sampling model

In the random sampling model, each DBN extracts data at random from the training database. In this case, the time series information of the data is ignored. Therefore, the information of training data is distributed uniformly to each DBN.



(a) Simple division model (b) Random sampling model

Figure 4: Method of generating training data using bagging

3. Prediction Result

Each DBN is trained using data from 2012 to 2014 and predicts the wind speed fluctuation for 2015. In ensemble learning, five DBNs are used as weak learning machines ($M = 5$). The five training data are generated by extracting data from the training database from 2012 to 2014. Table 3 shows the training parameters of the DBNs. Backpropagation is applied to fine-tuning as supervised learning.

Table 3: Training parameter of DBNs

	Pre-Training	Fine-Tuning
Learning rate	0.01	0.002
Epoch	5000	10000

3.1 Evaluation method for fluctuation prediction

The prediction result for wind speed fluctuation is evaluated from the concordance rate of the fluctuation, which is defined as the concordance rate of the predicted wind speed fluctuation $\hat{d}(t+n)$ and the fluctuation of the observed wind speed $d(t+n)$ (as a percentage).

3.2 Comparison of prediction accuracy of wind speed fluctuation

Table 4 shows the fluctuation prediction accuracy of an hierarchical neural network (HNN) and DBN. The seven-layered HNN is not able to learn because the root mean squared error (RMSE) did not decrease during training and all the prediction results increased. In contrast, the seven-layered DBN can learn and the prediction accuracy was improved. It was found that the three-layered HNN is better than the seven-layered DBN for 10-min-later prediction. On the other hand, the seven-layered DBN is better than three layered HNN for 60-min-later prediction. At this stage, the search for some parameters of the DBN is not satisfactory. Therefore, it is necessary to confirm the maximum performance of the DBN for suitable parameters.

Table 4: Prediction accuracy of HNN and DBN [%]

Prediction time	10min.	60min.
Three layered HNN	**59.50**	56.28
Seven layered HNN	51.15	51.15
Seven layered DBN	55.72	**57.30**

Figure 5 shows the distribution of the DBN outputs of the seven-layered DBN for 60-min-later prediction. The DBN outputs are concentrated in around 0.5. In addition, the prediction accuracy of wind speed fluctuation is poor around 0.5, and high near 0 or 1. Taking these results into account, the combined method of ensemble learning using the average value of all DBN outputs is employed (see sect. 2.3). Herewith, an output value of near 0 or 1 is more strongly reflected in the prediction result.

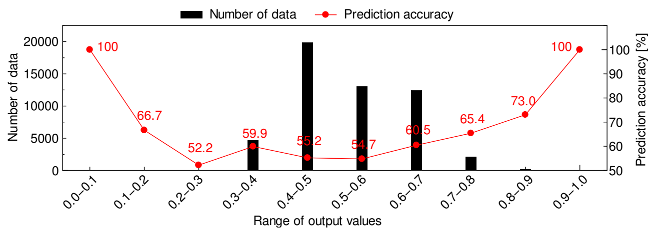


Figure 5: Distribution of DBN outputs (seven-layered DBN, prediction after 60 min)

3.3 Comparison of ensemble learning between the simple division mode and the random sampling mode

Table 5 shows the prediction accuracy with the bagging methods of two data extraction methods (simple division and random sampling). It is confirmed that the prediction accuracy of the wind speed fluctuation is improved using multiple weak DBNs. For 10-min-later fluctuation prediction, the prediction accuracy is superior to that of the seven-layered DBN without ensemble learning. However, the prediction accuracy of wind speed fluctuation is inferior to the three-layered HNN.

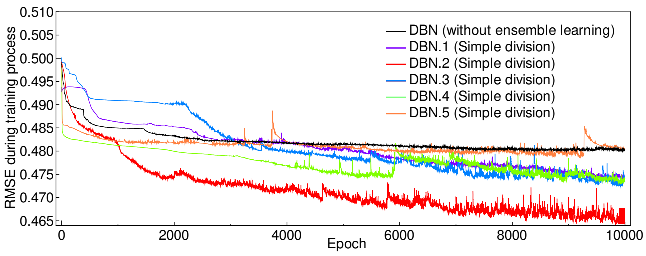
For 60-min-later fluctuation prediction, the prediction accuracy is improved from 57.30% (see table 4) to 58.12% using the random sampling model, which is better than the result for the simple division model (57.34%). Consequently, the prediction accuracy is improved by using part of the training database without training the whole training database. Furthermore, it is thought that extracting data uniformly from the training database is effective for the prediction of the wind speed fluctuation in order to learn various input data patterns. Accordingly, the proposed fluctuation prediction system, which uses DBNs with ensemble learning, is effective for 60-min-later fluctuation prediction.

Table 5: Prediction accuracy of HNN and DBN [%]

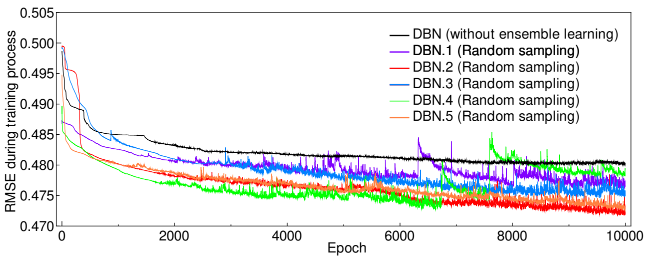
	Simple division		Random sampling	
	10min	60min	10min	60min
DBN.1	55.30	54.92	53.14	57.24
DBN.2	54.64	55.75	56.08	57.68
DBN.3	54.24	56.23	55.16	55.60
DBN.4	54.32	55.05	55.87	56.32
DBN.5	55.66	57.31	54.94	56.45
Simple majority	56.00	57.34	55.98	58.12
Average evaluation	55.80	57.42	56.10	57.82

3.4 Analysis of the prediction with ensemble learning

The transitions of the RMSE during the training process are shown in Figs.6(a) and 6(b). From Figs.6(a) and 6(b), the training efficiency of each DBN is improved by applying ensemble learning. The RMSE that of each weak DBN is smaller than the DBN without ensemble learning. Thus, we can confirm that ensemble learning is effective for training big data.



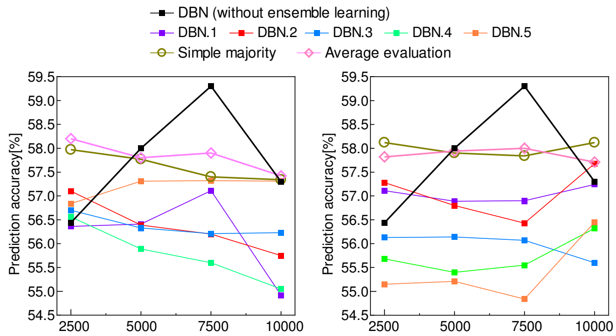
(a) Simple division model



(b) Random sampling model

Figure 6: Transition of RMSE of each DBN in the training process

Finally, Figs.7(a) and 7(b) show the fluctuation prediction accuracy of each weak DBN per epoch. The tendency of over-learning is confirmed for the simple division model. Hence, random sampling is a suitable way of generating training data. Moreover, the fluctuation prediction result of the DBN (epoch: 7500) without ensemble learning is the most accurate (59.32%). The prediction result of ensemble learning is inferior to that of the DBN is without ensemble learning. However, the DBN with ensemble learnings expected to have satisfactory accuracy in the circumstances in spite of the training epoch. However, it is necessary to thoroughly investigate some network parameters because the training of the network depends on whether or nor ensemble learning is incorporated.



(a) Simple division (b) Random sampling

Figure 7: Prediction accuracy of each weak DBN per training epoch (prediction after 60 min)

4. Conclusions

In this paper, we predicted wind speed fluctuation using a DBN with ensemble learning. The 60-min-later fluctuation prediction was improved by using the DBN. Also, for the bagging method used in the ensemble learning, we confirmed that a random sampling model was better than a simple division model.

In the future, it will be necessary to search for DBN parameters that are suitable for fluctuation prediction. In addition, the method combining an ensemble of outputs and a number of weak DBNs should be optimized.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 15K00048.

References

[1] Y. Miyabe, T. Kitajima and T. Yasuno: Wind speed prediction model using neural networks classified by observed wind speed, 2015 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP'15), No. 28PM2-2-2, pp. 230-233, Feb. 2015.

[2] S. Yoshida, H. Suzuki, T. Kitajima, A. M. Kassim, T. Yasuno, Correctiomn Method of Wind Speed Prediction System using Predicted Wind Speed Fluctuation, SICE Annual Conference 2016, No.Th2bPo5.5, pp.1054-1059, Sept. 2016.