Int. J. Anal. Appl. (2024), 22:30

International Journal of Analysis and Applications

Application of Deep Belief Network in Weather Modeling: $PM_{2.5}$ Concentration in Thailand

Wandee Wanishsakpong[1], Suwanna Atsawachanakan[1], Thammarat Panityakul[2,*]

[1]*Department of Statistic, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand*
[2]*Division of Computational Science, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand*
**Corresponding author:* thammarat.p@psu.ac.th

ABSTRACT. In Thailand, the number of particles matter with diameter of less than 2.5 microns or $PM_{2.5}$ concentration exceed the standard in many areas, especially in Chiang Mai. This affects the image of the country in terms of economy, health, and environment. The objective of this research is to study the structure of model for $PM_{2.5}$ concentration by using a Deep Belief Network (DBN) with the daily data set of $PM_{2.5}$ concentration from the air quality monitoring station at Yupparaj Wittayalai School, Chiang Mai. The data was analyzed through an unsupervised path using the Minimizing Contrastive Divergence (MCD) algorithm, followed by a supervised path using Back-Propagation Neural Network (BPNN) algorithm to estimate the parameters of DBN. The result shows that the optimal DBN structure has 5 input nodes and 20 hidden neurons in the first hidden layer. This model has an 88.4 percent accuracy in forecasting $PM_{2.5}$ concentration. In addition, this model can be applied for other weather forecasting such as rainfall or water level in a basin.

1. Introduction

Recently, air pollution problems turned into a serious issue around the world and have become more serious for all living organisms on the earth. From the past few decades, air pollution is still increasing because of urbanization, industrialization, automobiles, power plants, chemical activities and other natural activities such as agricultural burning, volcanic eruptions and wildfires (Doreswamy et al., [3]). The cause of death from air pollution is increasing every year as well. According to the World Health Organization, about 7 million people die each year from exposure to small airborne particles, and more than 90% of those deaths are caused by

Received Sep. 13, 2023

2020 Mathematics Subject Classification. 62M10.

Key words and phrases. PM2.5; deep belief network; continuous restricted Boltzmann machines; minimizing contrastive divergence; back-propagation neural network.

https://doi.org/10.28924/2291-8639-22-2024-30
© 2024 the author(s)
ISSN: 2291-8639

exposure to air pollution (WHO, [11]). The small airborne particles affect to health hazards, especially those at risk for cardiovascular and respiratory disease (Lu et al., [6]).

In Thailand, air pollution problems tend to be more serious especially in Bangkok and the northern part of Thailand. During the dry season of the year 2019, it was found that the smog situation in 9 provinces in the upper northern region found small particulate matter ($PM_{2.5}$) which exceeded the average standard of Thailand, which is set to not exceed 50 micrograms per cubic meter almost the entire month of March of every station (Chantra and Wiriya, [1]). Chiang Mai province is one of the Northern provinces was ranked as the most air-polluted city in the world. The problem of $PM_{2.5}$ affects the image of the country in terms of economy, health, the environment, and tourism. The report of Director of the Tourism Authority of Thailand (TAT) revealed that Chiang Mai has approximately 5.7 million tourists since Jan until July 2019. The number of tourists in March decreased 12.21%, and April decreased 6.46% due to the problem of $PM_{2.5}$ dust concentration (Prachachat Business, [8]).

Many statistical methods have been used to study and forecast PM concentrations using air pollutants data, such as Holt-Winters exponential smoothing, autoregressive integrated moving average (ARIMA), linear regression model. These methods use collections of probability distribution and assumption to make predictions (Wongrin et al., [10]). In contrast, machine learning is another method to focus on prediction by using learning algorithms to find patterns and can apply to deal with data and make a prediction. It uses a layered structure of an algorithm called an Artificial Neural Network (ANN). ANN is a simulation of the neural network in the human brain. The ability of ANN to learn problems that are non-linear equations (Non-linear problems) and can learn data without parameters (Nonparametric data). Nowadays, artificial neural networks are widely applied but there are still some problems in implementing ANN, namely 1) the proper structure for the data, 2) initial weight, and 3) learning rate were suitable for learning information. Hecht-Nielasen [4] solved the first problem is multilayer perception and hidden layer able to learn non-linear problems. Hinton and Salakhutdinov [5] presented a learning process of Deep Belief Network (DBN) consisting of two learning processes. The first learning process of the unsupervised path is learning without labels, the weight in this process will be used to initial weight value in the supervised path for solved the second problem. The third problem of ANN is the learning rate value, if the learning rate is too high the performance of learning will be unstable, and if the learning rate is too low the learning time will be longer. These ANN problems are solved by using Deep Belief Network (DBN) with Continuous Restricted Boltzmann Machines (CRBMs).

Therefore, this study aims to study structure of model for $PM_{2.5}$ concentration by using Deep Belief Network (DBN) with the daily $PM_{2.5}$ concentration in Chiang Mai, Thailand.

2. Materials and Methods

2.1 Data

The data used in this study is the daily average (24-hour average) of $PM_{2.5}$ concentration (micrograms per cubic meter, $\mu g/m^3$) from the air quality monitoring station in Chiang Mai, the Pollution Control Department, Air Quality and Noise Management Bureau, The Ministry of Natural Resources and Environment of Thailand (www.pcd.go.th). The dataset contains the daily average of $PM_{2.5}$ concentration from January 1, 2012, to June 30, 2022, a total of 3,833 days. However, the dataset contains some missing values. These missing data were estimate by using linear interpolation method. The data were divided into two data sets, The first 80% of the data is denoted by the training set, and the remaining 20% is used for evaluating the models' performance, denoted by the testing set.

2.2 Methods

In this section, Deep Belief Network (DBN) is presented to study the suitable structure of model for prediction. There are 2 processes to develop model which are unsupervised path and supervised path. To validate the prediction performance of the models, Mean Square Error (MSE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-square (R^2) are used.

2.2.1 Unsupervised path

Unsupervised path uses Continuous Restricted Boltzmann Machine (CRBM) to train the dataset by Minimizing Contrastive Divergence (MCD) algorithm. The intuitive motivation to use MCD is using one step Gibbs sampling to estimate the value of 3 unknown parameters.

Continuous Restricted Boltzmann Machine (CRBM)

A continuous restricted Boltzmann machine (CRBM) was introduced by Chen and Murry (2003). The structure of CRBM which consists of 2 layers. The first layer is visible layer consists of m neurons $(V_1, V_2, V_3, \ldots, V_m)$. The second layer is hidden layer consists of n neurons $(H_1, H_2, H_3, \ldots, H_n)$ as shown in Fig. 1.

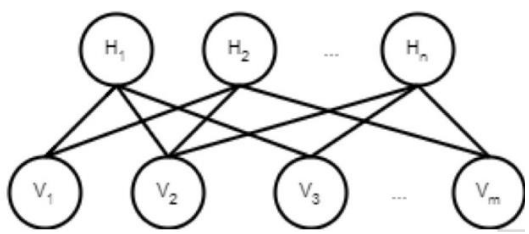

Fig. 1 CRBM's structure

Single neuron of Continuous Restricted Boltzmann Machine

The structure of CRBM consists of 2 layers, which are visible layer and hidden layer. A neuron in visible layer is called visible neuron and a neuron in hidden layer is called hidden neuron. Each visible neuron and hidden neuron associates with 2 values which are input value and output value. (Chen, H. and Murray, A.F., [2])

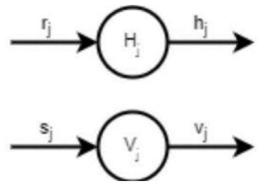

Fig. 2 Visible neuron and Hidden neuron

The input and output values of visible neuron, V_j and the input and output values of hidden neuron, H_j as shown in Fig. 2. The input value r_j and output value $V_j = \varphi(r_j)$ of visible neuron V_j can be computed as show in Equation (1) and Equation (2) respectively.

$$r_j = \sum_{i=1}^{n} w_{ij} h_i + \sigma N_j(0,1) \tag{1}$$

$$\varphi(r_j) = \theta_L + (\theta_H + \theta_L)\frac{1}{1+e^{-Kr_j}} \tag{2}$$

The input value s_j and output value $h_j = \varphi(s_j)$ of hidden neuron H_j can be computed as shown in Equation (3) and Equation (4) respectively.

$$s_j = \sum_{i=1}^{n} w_{ij}\, v_i + \sigma N_j(0,1) \tag{3}$$

$$\varphi(s_j) = \theta_L + (\theta_H + \theta_L)\frac{1}{1+e^{-\lambda_j s_j}} \tag{4}$$

where w_{ij} is the weight value, σ is the constant, $N_j(0,1)$ is the Gaussian random variable, θ_L, θ_H are constant lower and upper asymptotes.

Deep Belief Network (DBN)

Deep Belief Network (Chen, H. and Murray, A.F., [2]) is stacked of CRBM, the DBN structure is shown in Fig. 3. The CRBM consists of two layers, when two CRBM stack together, it means that the hidden layer of the first CRBM becomes visible layer of the second CRBM. Therefore, DBN can be formed by stacking of many CRBMs.

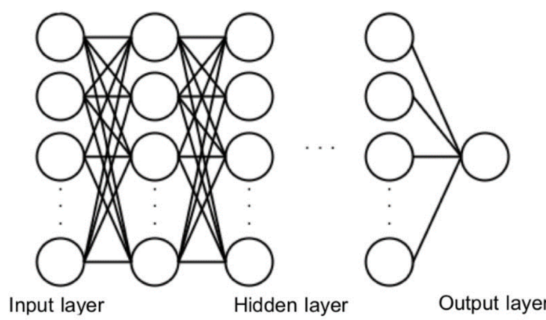

Fig. 3 DBN's structure

Minimizing Contrastive Divergence for CRBM

This algorithm works under the process of one-step Gibbs sampling. The MCD algorithm used the training rule in combination with learning rates to incrementally update the parameters values to achieve the final values. Therefore, we obtain the update rule for weight parameter w_{ij}, noise control parameter κ_i of visible neuron V_j and noise control parameter λ_j of hidden neuron

H_j which are shown in Equation (5), Equation (6) and Equation (7) respectively. (Chen, H. and Murray, A.F., [2])

$$\Delta\widehat{w_{ij}} = \eta_w \left(\langle v_i^{(0)} h_j^{(0)} \rangle - \langle v_i^{(1)} h_j^{(1)} \rangle\right) \tag{5}$$

$$\Delta\widehat{\kappa_j} = \frac{\eta_w}{\kappa_j^2} \left(\langle {v_i^{(0)}}^2 \rangle - \langle {v_i^{(1)}}^2 \rangle\right) \tag{6}$$

$$\Delta\widehat{\lambda_j} = \frac{\eta_\lambda}{\lambda_j^2} \left(\langle {h_j^{(0)}}^2 \rangle - \langle {v_j^{(1)}}^2 \rangle\right) \tag{7}$$

Learning DBN with CRBMs

The target of learning a deep belief network (Chen, H. and Murray, A.F., [2]) is to find the appropriate weight values between each layer. To learn the weight values of a DBN encompass 2 parts, unsupervised path, and supervised path. In unsupervised, we will learn each CRBM in order to use the Minimizing Contrastive Divergence algorithm (MCD) to obtain the weight values for all connections in the network. The final weight value from unsupervised path will be used to initialize weight value in supervised path. To train the deep belief network in supervised path, we will use the back-propagation neuron network algorithm (BPNN)

2.2.2 Supervised path

Supervised path uses feedforward backpropagation neural network to train the data set by back propagation neural network algorithm to find the optimal weight value.

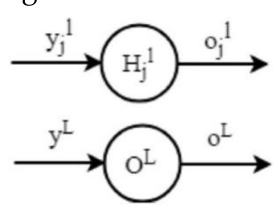

Fig. 4 Single hidden neuron H_j^l and single output neuron O^L in FFBNN

Feed-Forward Backpropagation Neural Network: FFBNN (Chen, H. and Murray, A.F., [2])

The first layer is the input layer, denoted l=1, which consists of input nodes. The last layer is output layer, denoted l=L, which consists of one output neuron. The hidden layers are all layers between input and output layers, denoted $2 \leq 1 \leq L-1$. There may be many hidden layers and each hidden layer consists of hidden neurons. The number of hidden neurons in each hidden layer can be different number as shown in Fig.4.

Back-propagation neural network: BPNN

Back-Propagation Neuron Network algorithm loos for the weight values which minimize error using Equation (8). In order to get the appropriates weight values using BPNN, 2 steps are required. The first step is Forward propagation, and the second step is backward propagation. (Chen, H. and Murray, A.F., [2])

$$E^L(t,o) = \frac{1}{2}(t - o^L)^2 \tag{8}$$

where E^L is the error function in layer L, t is the target value, o^L is the output value of output neuron in layer L.

Forward propagation computes the input Input (y_i) and Output (o_j) values of each neuron. The equations to compute are shown as Equation (9), and Equation (10) respectively.

$$y_j = \sum_{i=1}^{m} w_{ij} o_i + b_j \tag{9}$$

$$o_j = \frac{1}{1+e^{-y_j}} \tag{10}$$

where w_{ij} is the weight value and b_j the bias value.

Back propagation is the process of adjusting and updating the weight value by Equation (11).

$$\Delta w_{ij} = \eta_w\left[(t - o_j)o_i\right] \tag{11}$$

where η_w is the learning rate, t is the target value, o_j is the output value of neuron at present layer, o_i is the output value of neuron at layer before.

2.2.3 The performance indicators:

Mean Square Error: MSE (Phanida et al., 2017)

$$\text{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Root Mean Square Error: RMSE (Saichon, 2017)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Mean Absolute Error: MAE (Saichon, 2017)

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

R square: R^2 (Nikom, 2020)

$$R^2=\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{\sum_{i=1}^{n}(y_i - \bar{y})}$$

where y_i is actual data value, $\hat{y}_i$ is predicted data value, n is number of data set.

2.3 Research Method

2.3.1 To study the $PM_{2.5}$ concentration data from the air quality monitoring station at Yupparaj Wittayalai School, Chiang Mai since January 1, 2012, until June 30, 2022.

2.3.2 Check the missing data, if the dataset has missing value, then estimate missing data by using linear interpolation method.

2.3.3 Develop a model for prediction by learning Deep Belief Network with training set. The target of learning a deep belief network is to find the appropriate weight values.

There are 2 processes to develop model which are unsupervised path and supervised path.

2.3.4 Test the model with testing set.

2.3.5 Compare the performance of the model from the calculated values by selecting the model with the lowest MSE, MAE, RMSE, and the highest R-square values.

3. Results and Discussion

The dataset contains the daily average of $PM_{2.5}$ concentration from January 1, 2012, to June 30, 2022, a total of 3,833 days. Therefore, the complete dataset consists of n = 3833. However, the dataset has some missing values. These missing data were estimate by using linear interpolation method. The summary statistics of PM2.5 concentration for each year are provided in Table 1.

Year	Starting	Ending	n	Missing	Mean	S.D.	Median	Min	Max
2012	1 Jan	31 Dec	354	12	28.61	23.18	21.00	8.00	147.00
2013	1 Jan	31 Dec	346	19	34.59	22.80	26.00	11.00	188.00
2014	1 Jan	31 Dec	341	24	33.11	23.93	27.00	6.00	188.00
2015	1 Jan	31 Dec	353	12	32.65	36.64	18.00	6.00	266.00
2016	1 Jan	31 Dec	357	9	31.97	27.24	20.00	6.00	144.00
2017	1 Jan	31 Dec	339	26	23.09	16.36	16.00	4.00	98.00
2018	1 Jan	31 Dec	298	67	29.57	21.98	21.00	7.00	108.00
2019	1 Jan	31 Dec	362	3	30.37	28.47	21.00	5.00	210.00
2020	1 Jan	31 Dec	357	9	28.48	28.76	17.00	4.00	174.00
2021	1 Jan	31 Dec	363	2	24.06	23.44	14.00	4.00	116.00
2022	1 Jan	30 Jun	181	0	23.89	14.45	19.00	5.00	72.00

Table 1 Descriptive statistics of $PM_{2.5}$ concentration.

From Table 1, the average of $PM_{2.5}$ concentration are between 23 to 35 $\mu g/m^3$, and the medians are slightly lower. Most years have standard deviation between 14 to 29 $\mu g/m^3$, except year 2015, which show high variation. The highest mean, standard deviation, median, minimum, and maximum $PM_{2.5}$ concentration were found in year 2013, 2015, 2014, 2013 and 2015, respectively. The smallest mean, standard deviation, median, maximum $PM_{2.5}$ concentration were found in year 2022, 2022, 2021 respectively and the smallest minimum was found in year 2017, 2020 and 2021.

The time series plot and boxplots of $PM_{2.5}$ concentration are displayed in Fig. 5 and 6, respectively. The graphs illustrated that every year having seasonality with high concentration period from November to April, and low concentration period from May to October. The box plot show that the data in every year are right skewed with a large number of extreme values.

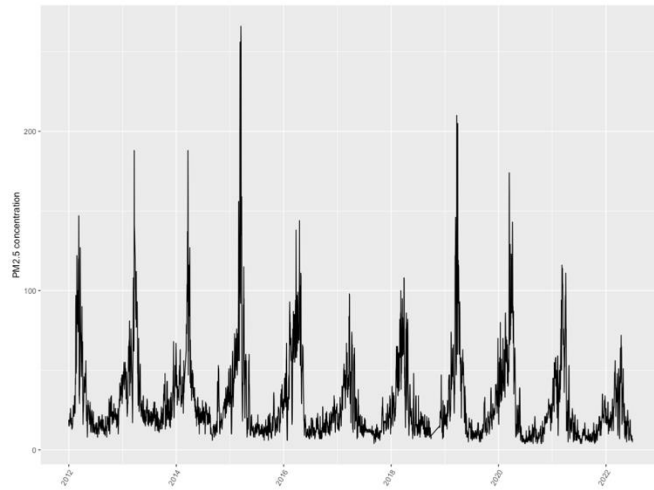

Fig. 5 Time series plot of $PM_{2.5}$ concentration

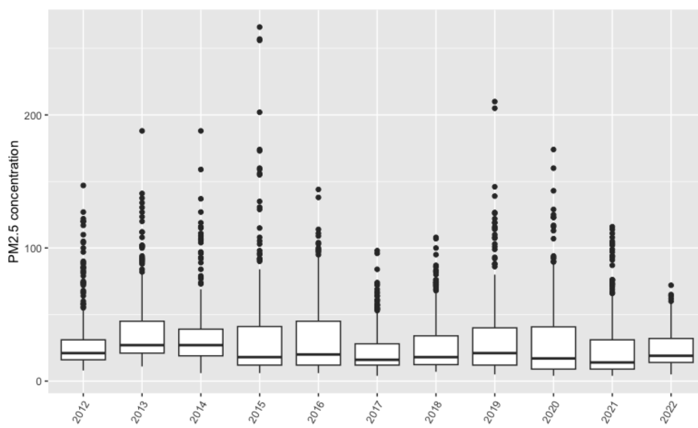

Fig. 6 Boxplots of $PM_{2.5}$ concentration

The dataset were learned by using the Minimizing Contrastive Divergence (MCD) algorithm to determine the suitable learning rate (η) for the data set. The learning rate = (0,0.1) was investigated as shown in Table 2.

η	R^2	TIME (sec)
0.0001	0.6081	39.0484
0.0010	0.8025	39.4391
0.0100	**0.8377**	**40.0568**
0.0250	0.8187	39.3103
0.0500	0.7771	39.5934
0.0750	0.7718	39.4840
0.1000	0.7464	39.4450

Table 2 Program processing time when η is different.

From Table 2, the suitable learning rate is 0.01 which contains the highest R^2 equal 0.8377 and takes 40.0568 seconds to develop the model.

The learning rate (η) was used to updates the parameters to determine the suitable DBN structure for the dataset. The learning rate was test with structure which have since 1 to 7, 14, 30, 60, and 90 input nodes, each input node consists of one hidden layer with hidden neurons 5, 10, 15, 20, 25, and 30 neurons as shown in Table 3.

From Table 3, R^2, MSE, MAE, and RMSE were considered for the efficient testing in predictable of $PM_{2.5}$ concentration. The result showed that the suitable structure with the daily $PM_{2.5}$ concentration data has 5 input nodes and 20 hidden neurons in the first hidden layer as shown in Fig. 7. This DBN structure gives the lowest value of MSE, RMSE, MAE were 75.5206, 8.6903, 5.2448 respectively and the highest value of R^2 were 0.8874. The plot of R^2 with difference structures were shown in Fig. 6. It found that the 5 input nodes and 20 hidden neurons of the DBN structure were presented the highest R^2 value.

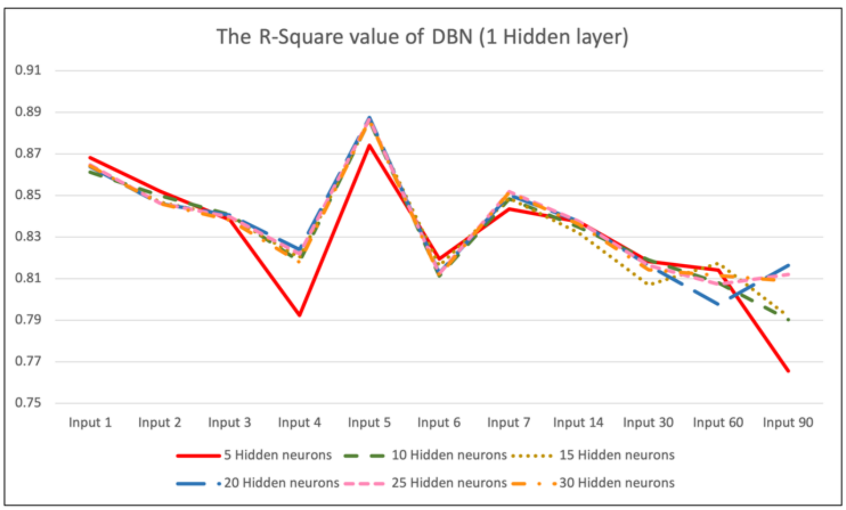

Fig. 6 R^2 of DBN structure

Input	Hidden neuron	R^2	MSE	RMSE	MAE
1	5	0.8681	61.9798	7.8727	5.0843
	10	0.8611	65.2648	8.0787	5.1351
	15	0.8642	63.8370	7.9898	5.0524
	20	0.8638	63.9955	7.9997	5.1226
	25	0.8647	63.5831	7.9739	5.0803
	30	0.8641	63.8733	7.9921	5.1304
2	5	0.8519	73.5017	8.5733	5.2243
	10	0.8499	74.4614	8.6291	5.1991
	15	0.8467	76.0799	8.7224	5.2627
	20	0.846	76.4191	8.7418	5.2686
	25	0.8462	76.2969	8.7348	5.2528
	30	0.8461	76.3630	8.7286	5.2540
3	5	0.8383	101.3963	10.0696	5.6426
	10	0.8407	99.8874	9.9944	5.5141
	15	0.8396	100.5822	10.0291	5.4788
	20	0.8404	100.1134	10.0057	5.4476
	25	0.8397	100.5368	10.0268	5.4545
	30	0.8381	100.5627	10.0778	5.4710
4	5	0.7923	97.0952	9.8537	6.9341
	10	0.8182	84.9919	9.2191	6.3696
	15	0.8202	84.0499	9.1679	6.3252
	20	0.8240	82.2806	9.0709	6.2549
	25	0.8219	83.2700	9.1252	6.3324
	30	0.8178	85.1785	9.2292	6.501
5	5	0.8740	84.5440	9.1948	5.7764
	10	0.8860	76.5101	8.7470	5.3289
	15	0.8864	76.2520	8.7322	5.2837
	20	**0.8874**	**75.5206**	**8.6903**	**5.2448**
	25	0.8867	76.0042	8.7180	5.2468
	30	0.8862	76.3575	8.7383	5.2769
6	5	0.8194	99.5131	9.9756	5.7442
	10	0.8113	103.9827	10.1972	5.8151
	15	0.8161	101.2901	10.0643	5.7615
	20	0.8127	103.1757	10.1575	5.8070
	25	0.8123	103.4301	10.1701	5.8487
	30	0.8115	103.8747	10.1919	5.8618
7	5	0.8435	92.9968	9.6435	5.5373
	10	0.8483	90.1337	9.4939	5.3616
	15	0.8486	89.9426	9.4838	5.3690
	20	0.8502	88.9813	9.4330	5.3552
	25	0.8518	88.0679	9.3844	5.3266
	30	0.8515	88.2322	9.3932	5.3465

Table 3 R^2, MSE, RMSE and MAE of Deep Belief Network.

Input	Hidden neuron	R^2	MSE	RMSE	MAE
14	5	0.8373	110.1368	10.4946	5.5735
	10	0.8347	111.9016	10.5784	5.5735
	15	0.8320	113.7668	10.6662	5.5494
	20	0.8376	109.9633	10.4863	5.5285
	25	0.8376	109.9842	10.4873	5.4995
	30	0.8362	110.9016	10.5310	5.5252
30	5	0.8182	98.2962	9.9144	5.5825
	10	0.8190	97.8726	9.8931	5.4578
	15	0.8069	104.3801	10.2167	5.5709
	20	0.8164	99.2666	9.9633	5.3797
	25	0.8163	99.2876	9.9643	5.4444
	30	0.8142	100.4377	10.0219	5.3833
60	5	0.8141	100.7502	10.0374	5.9045
	10	0.8080	104.0608	10.2010	5.8105
	15	0.8173	98.9982	9.9498	5.7565
	20	0.7976	109.6581	10.4718	5.8279
	25	0.8073	104.4323	10.2192	5.6899
	30	0.8115	102.1491	10.1069	5.645
90	5	0.7655	146.5098	12.1041	7.0614
	10	0.7902	131.1027	11.4500	7.1352
	15	0.7919	130.0277	11.4030	6.9797
	20	0.8164	114.6900	10.7093	6.4861
	25	0.8120	117.4571	10.8378	6.6282
	30	0.8089	118.8102	10.9000	6.7682

Table 3 R^2, MSE, RMSE and MAE of Deep Belief Network.

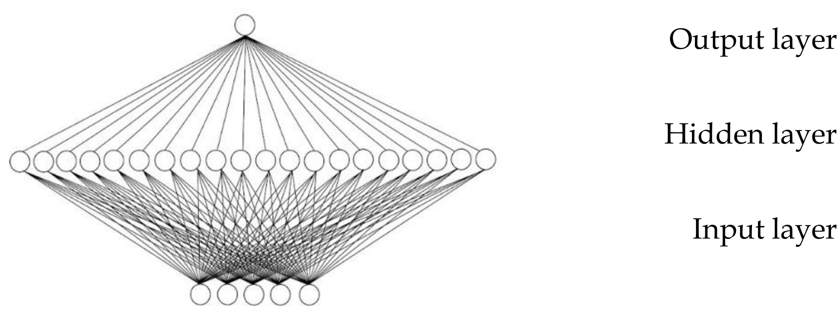

Fig. 7 The optimal DBN structure

4. **Conclusion**

Machine learning is an alternative approach that emphasizes the prediction aspect, employing learning algorithms to identify patterns and enabling the handling of data and prediction generation. This methodology can be applied through the utilization of Artificial Neural Networks. A type of artificial neural network called Deep Belief Network has used in modeling $PM_{2.5}$ concentration in Chiang Mai, Thailand by the 80% of the daily average of $PM_{2.5}$

concentration from January 1, 2012, to June 30, 2022, as the training set with linear interpolation method for missing data imputation. The models' performance is assessed by the remaining 20% of the dataset, these performance values were compared in order to find out the lowest MSE, MAE, RMSE, and the highest R-square values. The objective of training a deep belief network is to discover the correct weights for the network. This can be achieved through two distinct processes, the unsupervised pathway, and the supervised pathway.

Table 1 demonstrates that the average of $PM_{2.5}$ concentration is 23-35 $\mu g/m^3$ with a slightly lower median and standard deviation between 14-29 $\mu g/m^3$. The years with the highest average, variability, middle value, lowest value, and highest value of PM25 concentration were identified as 2013, 2015, 2014, 2013, and 2015, respectively. The years with the lowest average, variability, middle value, and highest PM2.5 concentration were determined to be 2022, 2022, and 2021, respectively. Additionally, the smallest minimum values were observed in the years 2017, 2020, and 2021. Figure 5 illustrates that each year exhibits a recurring pattern, characterized by a period of elevated concentration spanning from November to April, followed by a period of lower concentration from May to October. The time series plot in figure 6 reveals a prevalence of extreme values, resulting in a right-skewed distribution.

The model performance measurements in Table 3 state that, for the daily $PM_{2.5}$ concentration data, the optimal architecture consists of 5 input nodes and a first hidden layer comprising 20 hidden neurons, illustrated in Fig. 7 as well, with the lowest MSE, RMSE and MAE and the highest R^2 with respect to the others.

In summary, ANN modeling with optimized DBN structure, 5 input nodes and 20 hidden neurons in the first hidden layer, is a high-performance forecasting approach for $PM_{2.5}$ concentration. It should be compared with other forecasting methods that consider seasonal influences as well.

The DBN approach established in this study has versatile applications in weather forecasting, such as improving rainfall prediction through model refinement. By integrating supplementary meteorological data, including humidity levels, wind patterns, and historical precipitation records, it is possible to enhance the accuracy of rainfall forecasts.

Additionally, there is significant potential in extending the model's utility to assess flood risks. This can be achieved by combining the rainfall prediction model with topographical information, soil composition data, and historical flood records. This integration allows researchers to develop a comprehensive flood forecasting system. Furthermore, the incorporation of hydrological models into this framework offers the opportunity to make more precise water level predictions within river basins and watersheds. By coupling the meteorological model with hydrological simulations, valuable insights into the influence of rainfall on river discharge and water levels can be gained. This, in turn, supports reservoir management and flood control efforts. Ultimately, this system has the capacity to assist communities and authorities in preparing for and mitigating the potential impact of floods, thereby safeguarding lives and property.

Acknowledgment: This research was supported by National Science, Research and Innovation Fund (NSRF) and Prince of Songkla University (Grant No SCI6601264S). We would like to thanks Pollution Control Department, Air Quality and Noise Management Bureau, The Ministry of Natural Resources and Environment and Thai Meteorological Department, The Ministry of Digital Economy and Society for valuable data. And Miss Sureeluk Ma for guiding the computer program.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] S. Chantra, W. Wiriya, From the Problem of Open Burning to Building an Integrated Network on Smog, Community Res. Newsletter, 24 (2019), 10-15.

[2] H. Chen, A.F. Murray, Continuous Restricted Boltzmann Machine with an Implementable Training Algorithm, IEE Proc. Vis. Image Process. 150 (2003), 153-158. https://doi.org/10.1049/ip-vis:20030362.

[3] Doreswamy, K.S. Harishkumar, K.M. Yogesh, I. Gad, Forecasting Air Pollution Particulate Matter (PM2.5) Using Machine Learning Regression Models, Procedia Computer Sci. 171 (2020), 2057–2066. https://doi.org/10.1016/j.procs.2020.04.221.

[4] R. Hecht-Nielsen, Theory of the Backpropagation Neural Network, in: Neural Networks for Perception, Elsevier, 1992: pp. 65–93. https://doi.org/10.1016/B978-0-12-741252-8.50010-8.

[5] G.E. Hinton, R.R. Salakhutdinov, Reducing the Dimensionality of Data with Neural Networks, Science. 313 (2006), 504–507. https://doi.org/10.1126/science.1127647.

[6] F. Lu, D. Xu, Y. Cheng, S. Dong, C. Guo, X. Jiang, X. Zheng, Systematic Review and Meta-Analysis of the Adverse Health Effects of Ambient PM2.5 and PM10 Pollution in the Chinese Population, Environ. Res. 136 (2015), 196–204. https://doi.org/10.1016/j.envres.2014.06.029.

[7] N. Thanomsieng, Simple Linear Regression, Faculty of Public Health, Khon Kaen University, 2020.

[8] Prachachat Business, Statistics of Visitors to Chiang Mai, January-July 2019, 2019. https://www.prachachat.net/tourism/news-374725, Accessed 21 August 2022.

[9] P. Sombatmak. et al. The Data Is Out of the Criteria for Data Mining, Department of Statistics, Faculty of Science. King Mongkut's Institute of Technology Ladkrabang, 2017.

[10] W. Wongrin, K. Chaisee, K. Suphawan, Comparison of Statistical and Deep Learning Methods for Forecasting $PM_{2.5}$ Concentration in Northern Thailand, Pol. J. Environ. Stud. 32 (2023), 1419–1431. https://doi.org/10.15244/pjoes/157072.

[11] World Health Organization, 9 Out of 10 People Worldwide Breathe Polluted Air, But More Countries Are Taking Action, 2018. https://www.who.int/news/item/02-05-2018-9-out-of-10-people-worldwide-breathe-polluted-air-but-more-countries-are-taking-action.