Comparative analysis of solar power generation prediction system using deep learning

So-yeong Kim
Department of Information and Comm. Engineering
Chosun University
Gwangju, Republic of Korea
kimsy9808@chosun.kr

Eun-ji Lee
Department of Information and Comm. Engineering
Chosun University
Gwangju, Republic of Korea
20175122@chosun.kr

Uttam Khatri
Department of Information and Comm. Engineering
Chosun University
Gwangju, Republic of Korea
uttamkhatri03@gmail.com

Seokjoo Shin
Department of Computer Engineering
Chosun University
Gwangju, Republic of Korea
sjshin@chosun.ac.kr

Ji-In Kim
Department of Information and Comm. Engineering
Chosun University
Gwangju, Republic of Korea
ji_kim87@naver.com

Goo-rak kwon
Department of Information and Comm. Engineering
Chosun University
Gwangju, Republic of Korea
grkwon@chosun.ac.kr

***Abstract*— Due to the seriousness of environmental pollution, modern society is making efforts to switch from fossil fuel-centered energy to new and renewable energy worldwide. Among these new and renewable energies, solar energy is currently attracting attention due to its high growth potential. Among the renewable energy currently used, solar energy can generate a large amount of power without air pollutants and is widely used due to its high utilization. However, it is difficult to accurately predict the amount of power generation because solar energy, which is the source of energy from the sun, is greatly affected by seasons, weather, and installation environment. Therefore, this paper compares and analyzes the prediction and accuracy of solar power generation according to the environment through weather forecasts provided by the Meteorological Administration, using Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) learning models.**

***Keywords*—Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Solar energy, Power generation**

I. Introduction

Eco-friendly energy refers to energy that can be used so that the environment is no longer polluted by escaping from carbon energy in modern society. Since wind power, earth power, sunlight, and the like are used instead of carbon such as coal and petroleum, pollutants such as carbon dioxide are not emitted in the process of converting the natural phenomenon into thermal energy, and thus research has been steadily conducted. Fig. 1 is obtained from the data of 「Renewable energy generation (excluding non-renewable waste, from the 4th quarter of 2019)」 provided by the National Statistical Office [1]. The graph compares the amount of power generated by each energy source of renewable energy among the total power generation of 47,805,649 (MWh) of renewable energy in 2019. On the graph, the proportion of waste is the highest, but there is a disadvantage that air pollutants are generated when along with power generation using waste [2]. However, in the case of power generation using solar power, which has the next highest amount of power generation, except for the disadvantage that the initial cost is high, there are no air pollutants and a large amount of electricity can be obtained, so power generation using solar power is valuable as an alternative to fossil fuels. Therefore, in this paper, we intend to compare and analyze the deep learning models that predict the amount of power generation according to the weather with accuracy so that solar power generation can be efficiently performed.

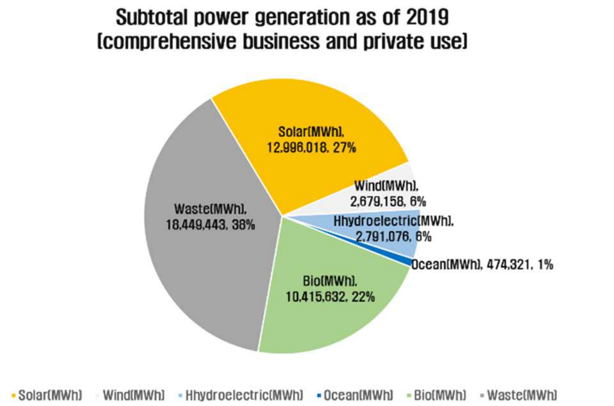

Fig. 1. Subtotal of power generation as of 2019 (Comprehensive for business and private use)

II. Dataset

Considering the fact that the data values differ depending on the Korea Meteorological Administration, for accurate prediction and verification of power generation, Cheongju, Chungcheongbuk-do, Republic of Korea was used as an observation point, and the Korea Meteorological Administration's synoptic meteorological observation data and the Norwegian Meteorological Agency's data were used together.

A. Training data

The training data used is the synoptic meteorological observation (ASOS) data provided by the Meteorological Data Open Portal of the Korea Meteorological Administration [3]. The observation point corresponds to Cheongju-si, Chungcheongbuk-do, and the observation period is from April 1, 2019 to September 1, 2021. The features used for training are 'Time', 'Temperature (°C)', 'Precipitation (mm)', 'Wind speed (m/s)', 'Humidity (%)', 'Dew point temperature (°C)', 'Amount of cloud cover (decile)' and 'Amount of power generation'.

B. Validation data

For the validation data, forecast data provided by the Norwegian Meteorological Agency was used [4]. The observation point corresponds to Cheongju-si, Chungcheongbuk-do as the observation point of the training data, and the observation period is from October 22, 2021 to December 21, 2021. About 31 days of data recorded during the period were used, and the features used for validation are 'temperature', 'precipitation', 'wind speed', 'humidity', 'dew point', 'cloud cover', 'power generation' and 'time'.

C. Output data

As the output data, the inverter data obtained through the monitoring system provided by MRT was used as the amount of power generation [5]. The location of the power plant corresponds to Cheongju-si, Chungcheongbuk-do, and the period is from April 1, 2019 to September 1, 2021. The total number of inverters is 11, and the total capacity of the inverters is 449.5kW. Features used for output data are 'accrue', 'start', 'day', and 'time'.

The amount of time generated = accrue – start – (the sum of days before the relevant day) (1)

The formula for calculating the amount of time generated is equation (1), and the unit is W.

III. DATA PROCESSING

A. Training Data

In the ASOS weather observation data, partial Nan values existed for each column, and the data values before and after Nan values were compared and corrected. In the case of cloud cover, it was provided in the tenth quartile, so the training was conducted by multiplying it by 10. Tables I and II correspond to the data set observed in Cheongju-si, Chungcheongbuk-do.

B. Validation Data.

In the forecast data of the Norwegian Meteorological Administration, wind direction and wind speed are provided in combination, so it was processed so that only the wind speed could be extracted separately. Tables III and IV correspond to the weather forecast data set corresponding to Cheongju-si, Chungcheongbuk-do.

TABLE I. WEATHER OBSERVATION DATA FOR ASOS IN CHEONGJU-SI, CHUNGCHEONGBUK-DO (1)

Date	Temp (°C)	Precipitation QC flag	Wind speed(m/s)
2019-04-01 0:00	2.7	9	0.4
2019-04-01 1:00	2.4		1
2019-04-01 2:00	2		1
2019-04-01 3:00	1.7		1.1
2019-04-01 4:00	1.3		0.9
2019-04-01 5:00	1.3		1.2
2019-04-01 6:00	1.1		0.8

TABLE II. WEATHER OBSERVATION DATA FOR ASOS IN CHEONGJU-SI, CHUNGCHEONGBUK-DO (2)

Humidity(%)	Dew point temp (°C)	Sunlight (hr)	Sunlight QC flag
48	-7.2		9
51	-6.7		9
53	-6.5		9
56	-6.1		9
59	-5.8		9
62	-5.1		9
66	-4.5		9

TABLE III. WEATHER FORECAST DATA IN CHEONGJU-SI, CHUNGCHEONGBUK-DO (1)

Date	Time	Temp	Wind speed	Precipitation
2021-10-22 0:00	0	7		
2021-10-22 1:00	1	6		
2021-10-22 2:00	2	5		
2021-10-22 3:00	3	5		
2021-10-22 4:00	4	5		
2021-10-22 5:00	5	4		
2021-10-22 6:00	6	4		
2021-10-22 7:00	7	5		
2021-10-22 8:00	8	6		
2021-10-22 9:00	9	8		
2021-10-22 10:00	10	10		
2021-10-22 11:00	11	11		
2021-10-22 12:00	12	12		

TABLE IV. WEATHER FORECAST DATA IN CHEONGJU-SI, CHUNGCHEONGBUK-DO (2)

Wind direction	Humidity	Dew point	Cloud cover	Power generation
1 from north west1	67	1	0	0
1 from north1	64	0	0	0
1 from west1	68	0	0	0
1 from west1	70	0	0	0
1 from west1	69	-1	0	0
1 from west1	71	0	0	0
1 from west1	71	-1	0	0
1 from west1	68	-1	0	0
1 from north west1	68	0	0	10.6
1 from north west1	63	1	0	31.1
2 from north west2	59	2	0	119.8
3 from north west3	55	3	0	216.2
3 from north west3	53	3	0	291.2

C. Output Data

As for the output data, the amount of data stored is flexible because the facility operates at the moment of power generation and has a large environmental influence.

The amount of power generation at the time = accumulated value of the time – accumulated value of the previous time (2)

The formula for calculating the amount of power generation at the time corresponds to equation (2), and when extracting data by time, if it is 0 o'clock, the accumulated value is initialized to 0W. The maximum power generation per hour of 1 inverter is 45,409W, and if the power generation exceeds 45,000W, it was estimated as an error value and performed by 11 inverters 1-11 respectively to delete the data on the relevant day. In addition, data on the day of partial omission due to the loss of communication at some time zones were deleted and processed for accurate prediction. Table V corresponds to a data set of 11 inverters installed in Cheongju-si, Chungcheongbuk-do.

TABLE V. INVERTER DATA IN CHEONGJU-SI, CHUNGCHEONGBUK-DO

Name	Unit	Start	Day	Accrue	Time
Inverter1	60	4748406	267	4748673	2019-04-01 6:00
Inverter1	60	4748406	1853	4750259	2019-04-01 7:00
Inverter1	60	4748406	5770	4754176	2019-04-01 8:00
Inverter1	60	4748406	20239	4768645	2019-04-01 9:00
Inverter1	60	4748406	43732	4792138	2019-04-01 10:00
Inverter1	60	4748406	70598	4819004	2019-04-01 11:00
Inverter1	60	4748406	95883	4844289	2019-04-01 12:00
Inverter1	60	4748406	115707	4864113	2019-04-01 13:00
Inverter1	60	4748406	130818	4879224	2019-04-01 14:00
Inverter1	60	4748406	146957	4895363	2019-04-01 15:00
Inverter1	60	4748406	162571	4910977	2019-04-01 16:00

IV. METHODS

Long Short-Term Memory (LSTM), one of the deep neural network algorithms, has the specificity of Cell State, which remembers the previous state of the system being simulated. Before the input vector (processed information) is stored in the current state along with the old state according to the Timestamp, the current power generation of the solar power generation system can be predicted through the previous weather information [6]. On the other hand, the Gated Recurrent Unit (GRU) is a modified model of the LSTM in which the structure of the LSTM is more simply processed. The forgetting gate and input gate of LSTM are integrated into update data, and the cell state and hidden state are integrated into one to become a simpler structure than LSTM, resulting in faster learning due to reduced weight count, but almost the same performance as LSTM [7]. And for more accurate comparison, a keras density layer is added. After providing an input layer to the model, it is a model in which a dense layer is added by activating a function that is generated to solve Dying ReLU (the phenomenon in which neurons die) of ReLU called Laeky ReLU.

A. Long Short-Term Memory (LSTM)

Four LSTM layers followed by one dense layer were used as conditions for the long short-term memory (LSTM) learning model. As an activation function, leakyReLU, alpha value was 0.01, lose was mse, and as an optimizer, 'adam' was used. Table VI corresponds to the LSTM learning model summary.

TABLE VI. LONG SHORT-TERM MEMORY LEARNING MODEL SUMMERY

Layer(type)	Output Shape	Param#
gru(GRU)	(None, 128)	52608
dense(Dense)	(None, 64)	8256
dense _1(Dense)	(None, 32)	2080
dense _2(Dense)	(None, 16)	528
dense _3(Dense)	(None, 1)	17

B. Gated Rucurrent Unit (GRU) Model

One GRU layer, three dense layers were used as conditions for the Gated Recurrent Unit (GRU) learning model. leakyReLU was used as the activation function, alpha value was 0.01, lose was mse, and 'adam' was used as the optimizer. Table VII corresponds to the GRU learning model summary.

TABLE VII. GATED RECURRENT UNIT LEARNING MODEL SUMMERY

Layer(type)	Output Shape	Param#
lstm(LSTM)	(None, 1, 32)	5120
lstm_1(LSTM)	(None, 1, 16)	3136
lstm_2(LSTM)	(None, 1, 8)	800
lstm_3(LSTM))	(None, 1, 4)	208
dense_4(Dense)	(None, 1, 1)	5

C. Dense Neural Network

Four Dense layers were used as conditions for the Dense Neural Network model. As an activation function, leakyReLU, alpha value was 0.01, lose was mse, and as an optimizer, 'adam' was used. Table VIII corresponds to the Dense Neural Network learning model summary.

TABLE VIII. DENSE NEURAL NETWOKR LEARNING MODEL SUMMERY

Layer(type)	Output Shape	Param#
dense_5(Dense)	(1, 1, 64)	512
dense_6(Dense)	(1, 1, 32)	2080
dense_7(Dense)	(1, 1, 16)	528
dense_8(Dense)	(1, 1, 8)	136
dense_9(Dense)	(1, 1, 1)	9

V. RESULTS AND DISCUSSION

To obtain the error rate applied to the demonstration system, the actual generation was subtracted from the predicted generation and divided by the installed capacity. Table I corresponds to the results when each model was trained. When Timestamp is 1 and Epoch is 50, the training loss values are 953 for LSTM, 974 for GRU, and 1003 for Dense Neural Network. The validation loss values were 2244 for LSTM, 2389 for GRU, and 2692 for Dense Neural Network. In terms of the error rate applied to the demonstration system, LSTM was 13.01%, GRU was 13.67%, and Dense Neural Network was 14.23%. Among them, LSTM showed the highest efficiency. When Timestamp is 24 and Epoch is 200, the training loss value is 31 for LSTM and 985 for Dense Neural Network. The validation loss values are 3769 for LSTM and 2174 for Dense Neural Network. In terms of the error rate applied to the demonstration system, the efficiency of the Dense Neural Network was higher, with LSTM being 17.03% and Dense Neural Network 13.03% as shown in TABLE IX.

TABLE IX. LONG SHORT-TERM MEMORY, GATED RECURRENT UNIT, DENSE NEURAL NETWORK LEARNING MODEL COMPARISON

Table Head	**LSTM**		**GRU**	**Dense Neural Network**	
	Timestamp : 1 Epoch : 50	***Timestamp : 24 Epoch : 200***	***Timestamp : 1 Epoch : 50***	***Timestamp : 1 Epoch : 50***	***Timestamp : 24 Epoch : 200***
Training Loss Value	953	31	974	1003	985
Validation Loss Value	2244	3769	2389	2692	2174
Error rate when applying the demonstration system	13.01%	17.03 %	13.67%	14.23%	13.03%

VI. CONCLUSION

Since the establishment of a solar energy generation prediction system is a key role in efficiently managing and distributing energy, many studies are being conducted. Therefore, accurate prediction of power generation for various variables such as season, weather, and installation environment is essential for solar energy generation. In this paper, we used ASOS data of Cheongju-si, Chungcheongbuk-do as training data, forecast data of Cheongju-si, Chungcheongbuk-do as validation data, and generation data of Cheongju-si, Chungcheongbuk-do as output data. It was shown that the amount of power generation can be predicted using LSTM (Long Short-Term Memory), GRU (Gated Recurrent Unit), and Dense Neural Network as the learning model algorithm. As a result, it was confirmed that the error rates of LSTM, GRU, and Dense Neural Network vary according to the detailed conditions of Timestamp and Epoch. When Timestamp is 1 and Epoch is 50, the efficiency of LSTM is the highest, and when Timestamp is 24 and Epoch is 200, the Dense Neural Network shows high efficiency.

ACKNOWLEGEMENT

This work was supported by the ICT R&D program of MSIP/IITP. [No. 2021-0-01264, AI-based photovoltaic system with real-time prediction of failure and generation amount]

REFERENCES

[1] Korea Energy Agency, " New and renewable energy supply performance survey," https://kosis.kr/statHtml/statHtml.do?orgId=337&tblId=DT_337N_A002.

[2] S. K. Kim, S. B. Lee, J. M. Park, C. Yu, J. H. Hong, H. C. Kim, N . E. Jung, M. R. Jo, J. H. Kim, and S. H. Heo, "A Study of Air Pollutants Reduction Plan for Waste Solid Fuel Fired Facilities," Journal of Korean Society for Atmospheric Environment, vol. 28, Issue 6, pp. 656-668, 2012.

[3] Meteorological Administration Weather Data Open Portal Jonggwan Weather Observation, https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36.

[4] Norwegian Meteorological Administration forecast data, https://www.yr.no/en/details/table/2869116/Republic%20of%20Korea/North%20Chungcheong/Miwon-myeon.

[5] MRT co. Ltd., http://www.mrt.co.kr/.

[6] Dongkyun Kim and Seokgu Kang. "Data Collection Method for Construction of Precipitation-Daily Runoff Estimation LSTM Model" Proceedings of the Korean Society for Water Resources 54, no.10, pp. 795-805, 2021.

[7] Tamal Datta Chaudhuri and Indranil Ghosh, "Artificial neural network and time series modeling based approach to forecasting the exchange rate in a multivariate framework," Journal of Insurance and Financial Management, vol. 1, no. 5, 2016.