Estimation of Power Generation and Consumption based on eXplainable Artificial Intelligence

SooHyun Shin*, HyoSik Yang**

*Department of Computer Engineering, Sejong University, Republic of Korea

schz307@naver.com, hsyang@sejogn.edu

Abstract— **Recently, various policy and technical efforts have been underway around the world to solve global warming. Major companies are also converting 100% of their electricity used inside the workplace to renewable energy. The most widely used PV energy in renewable energy has uncertainty and instability due to the nature of climate change, so accurate power generation prediction is essential, and power load prediction is required to increase energy efficiency. Although artificial intelligence technology is used in various ways to accurately predict it, it does not provide the basis for the result and the validity of the derivation process due to the black box problem of artificial intelligence technology. In this paper, we develop a PV power generation and power load prediction model using an explainable artificial intelligence model that increases the energy efficiency inside the workplace and derive a basis for the prediction results using an explainable artificial intelligence.**

***Keywords*— Green gas, SHAP, XAI, LSTM, LGBM**

I. INTRODUCTION

Environmental problems directly related to global warming include energy production and consumption and increasing waste disposal. Efforts are active worldwide, including signing the Paris Climate Agreement, the first climate agreement jointly by the international community to prevent global warming in 2016, and introducing a carbon tax on using fossil fuels. Large companies also continue to make such efforts. Recently, ESG (Environmental, Social, and Governance) management has become a hot topic around the world, and non-financial factors have a profound effect on management evaluation. Major companies such as Apple, Volkwswagen and Samsung are converting 100% of their workplace's power to renewable energy to focus on managing ESG. Among new and renewable energy, PV(Photo Voltaic) energy is widely spread because it is easy to install and maintain in existing buildings. However, PV energy systems have time-season variability, uncertainty, and constraints that make stable power supply difficult and reduce efficiency.

It is essential to predict variability in cause-based generation for higher efficiency and stable power supply in PV energy systems. In addition, it is necessary to predict not only the amount of power generated but also the amount of power load for proper scheduling of PV energy and ESS (Energy Storage System). Recently, artificial intelligence technology was applied to derive high acuity for prediction, but a problem of reliability arose due to a black-box problem in which artificial intelligence's mindset was unknown on what basis the prediction value came out. For this interpretation of artificial intelligence, an explainable artificial intelligence technology called XAI (eXplainable Artificial Intelligence) is developed to represent the rationality of machine-to-human interactions in a way that decomposes black-box for human understanding [1-2].

In this paper, we develop a PV power generation and power load prediction model that increases the energy efficiency inside a company's workplace. We compare results by predicting generation and load using public weather data, PV power generation data, and power load data provided in Toronto, Canada, and deep learning and machine learning models LSTM (Long Short-Term Memory) and LGBM (Light Gradient Boosting Machine), which are effective for time series prediction. In addition, we use SHAP(SHapley Additive exPlanations), an explainable artificial intelligence model, to solve the black box problem of predicted values, to increase explainability.

In chapter 2, we explain XAI, in Chapter 3 analyzes data sets and develops prediction models, predicts PV power and power loads using prediction models is developed in Chapter 4, and Chapter 5 deals with conclusion.

II. EXPAINABLE ARTIFICIAL INTELLIGENCE

As AI technology moves from general tasks to more core tasks, such as making decisions, it is necessary to provide the rationale for the result determined and the validity of the derivation process. AI technology with black box problems makes it difficult to optimize while not immediately knowing the cause of the error and even developers cannot figure out how to derive the results. In other words, even if the model makes certain judgments and produces good results, it is also difficult to use vast data sets and understand the predictions of complex AI models due to the black box problem of artificial intelligence. Recently, various studies on the XAI algorithm, an explainable artificial intelligence, are being conducted to solve this problem, so it is possible to identify which variables have affected the results, providing reliability and evidence. Early in the XAI study, the proposed LIME paper could not be proved by a unique solution, but SHAP's Shapley sampling or Shapley regression could already prove a unique solution, which could be achieved through approximation methods that reduce computational time.

In this paper, SHAP, one of the XAI models, is used. SHAP proposes additive feature attribution methods, a new classification for explanatory models, by considering the explanation of a prediction as the model itself and defining it as an eXplainable model. Based on game theory, we show that there is a unique solution to the additive feature attribution method and propose SHAP as a method for feature importance [3]. The SHAP algorithm is divided into TreeSHAP, which helps interpret machine learning models, and DeepSHAP, which helps interpret deep learning models, and evaluates data to quantify and visualize how input variables have affected the model's results [4].

III. PV Power Generation and Power Load Forecasting

PV Power systems have time and season variability, uncertainty, and constraints that make it difficult to supply stable power, degrade efficiency, and make it difficult to supply stable and reliable power, making it essential to predict and manage power generation. In addition, it is necessary to predict the amount of power load to enable high-efficiency, low-carbon power consumption inside the workplace. By predicting the amount of power generation and load, it is possible to manage low power that may minimize the internal power consumption of the system.

A. Data set

For the PV power generation forecast, we use the Toronto Canada weather public data for 2016 and the Toronto PV power generation public data. The Power load forecast uses 2016 power load data provided by Trontohydro and weather public data, Canada. A summary of the data set used for the PV power generation and load prediction is shown in the following Table 1. The discomfort index was obtained and added using temperature and humidity from the weather data used to predict the load. The evening data, in which the power generation could not be measured, made the prediction results poor, and the time in winter and summer was classified and excluded.

TABLE 1. Summary of Dataset

Dataset	Data Type
Canada Toronto Weather Public Data	**Day, hour, rainfall, prevrainfall(1 hour ago), windspeed, prevwindspeed(1 hour ago), insolation(ilsa), previnsolation(1 hour ago, ilsa), humidity, temperature, sunshine, cloudiness, discomfort index**
Canada Toronto PV Power Generation Public Data	**Hourly PV Power Generation**
Canada Torontohydro Power Load Data	**Hourly Power Load**

B. Prediction Model

PV power generation is affected by daily weather conditions and seasonal solar radiation due to the nature of the sun. In the case of power loads, they are used in the same pattern by time zone in residential and commercial use. To predict these time series characteristics of PV power generation and power load, it is necessary to establish a forecasting model specialized for time series. One of the commonly used methods for time series forecasting is Recurrent Neural Networks (RNNs) [5-6]. Conventional RNN is a sequence model that processes input and output in sequence units and has excellent performance in various time series data predictions including energy consumption prediction [7-8]. It, however, becomes a multiplication form with time and becomes gradient growth or exponential, learning to have dependency causes gradient vanishing or exploding problems, making training difficult [9-10]. To solve this problem, LSTM is used [11]. In previous studies, it can be confirmed that LSTM showed better performance than other models in predicting PV power generation [12-13].

LGBM is a gradient boosting framework that uses the decision tree algorithm among various machine learning algorithms and was developed by Microsoft for performance and scalability purposes. The leaf-wise tree structure has the advantage of reducing losses compared to the level-wise tree structure and increasing accuracy while learning at high speed [14]. A study of LGBM combined with LSTM for short-term PV power generation forecasting showed model suitability for PV power generation prediction [15].

Research using various prediction models for forecasting the amount of PV power generation has been conducted [16-17]. These studies, however, have used all night-time data where power generation cannot be measured due to lack of progress, which negatively affects the learning of the model [18].

This paper develops a PV power generation and power load amount forecasting model using LSTM and LGBM by reflecting an appropriate time variable. First, we create a new dataset by reflecting the time variables of meteorological data and power demand datasets to remove missing values and proceed with scaling for standardization and normalization of the data. At this time, the data was inserted and scaled using MinMaxScaler so that the maximum and minimum values were set to 1 and 0, and the power demand data was inserted with Y data and scaled using standard scaler using the mean and standard deviation. Second, we divide the dataset by a ratio of 80% and 20% to divide the training set for learning and the test set for testing. Third, learning is carried out by putting the separated training set in the LSTM model. The proposed LSTM model structure is shown in Figure 1.

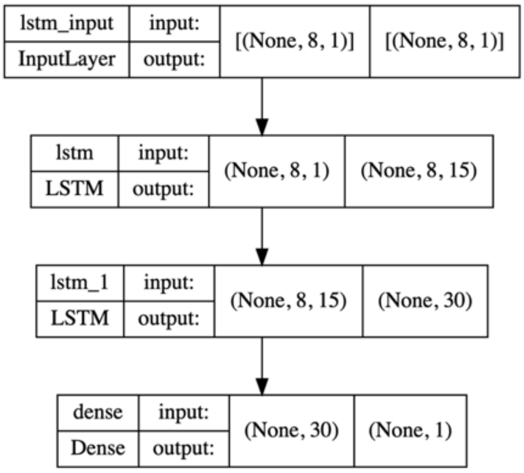


Figure 1. LSTM model structure

The following Figures 2 and 3 show the PV power generation and power load forecasting workflow.

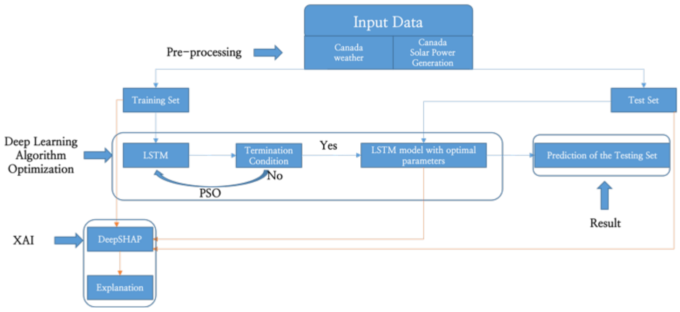


Figure 2. LSTM workflow

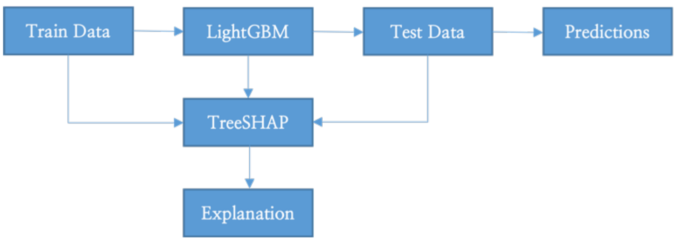


Figure 3. LGBM model structure

IV. Numerical Results

The results of PV power generation and load prediction using LSTM are shown in Figure 4 and 5. The blue line on the graph indicates the actual value and the orange line indicates the predicted results. Root Mean Square Error (RMSE) was used as an evaluation index. The RMSE of the generation amount was 0.0593, and the RMSE of the load amount was 0.0671, indicating the excellent performance of the model.

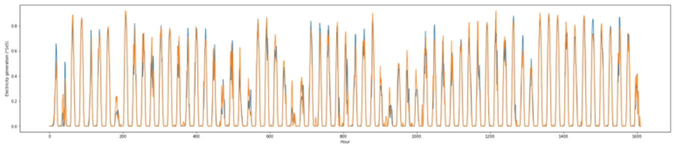


Figure 4. PV power generation forecasting result using LSTM

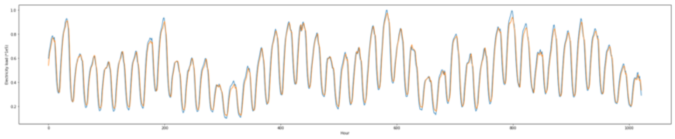


Figure 5. Power load forecasting result using LSTM

Figures 10 and 11 show the values expressed using SHAP to show how much the input values affect the results and are graphs that visualize the SHAP values of power generation and load, respectively.

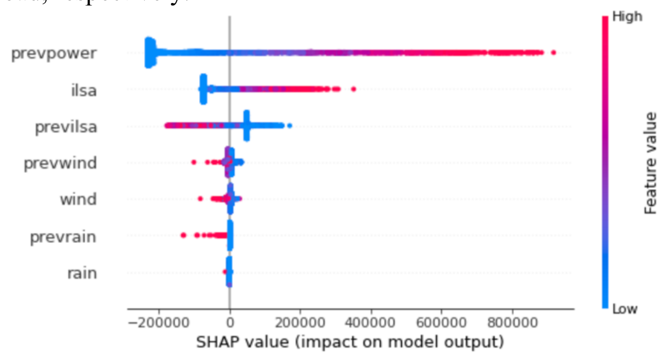


Figure 6. SHAP value of LSTM PV power generation

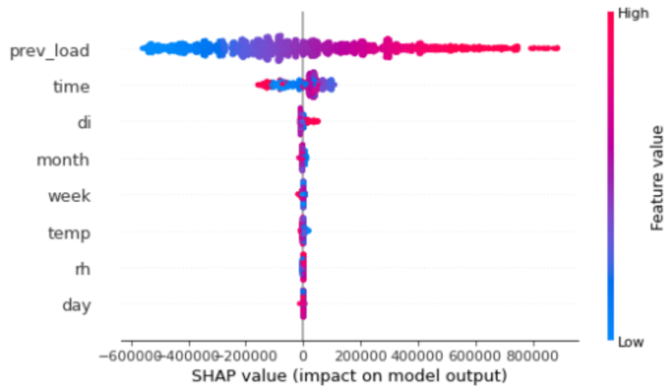


Figure 7. SHAP value of LSTM power load

The results of PV power generation forecasting and load forecasting using LGBM are shown in Figure 8 and 9. The blue line indicates the actual value and the orange line indicates the predicted results. The RMSE of the generation amount was 0.0973 and the RMSE of the load was 0.0881, indicating an excellent measurement value.

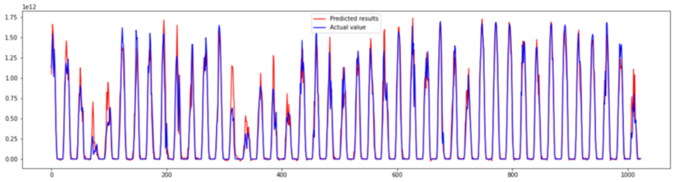


Figure 8. PV power generation forecasting result using LGBM

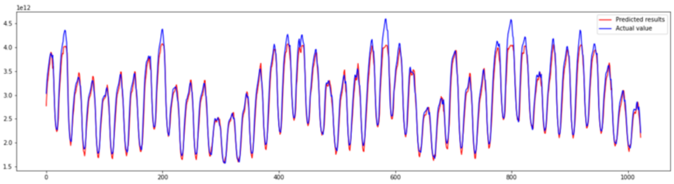


Figure 9. Power load forecasting using LGBM

Figures 10 and 11 show the values expressed using SHAP to measure how much the input value affected the result value and show the SHAP value of power generation and load, respectively.

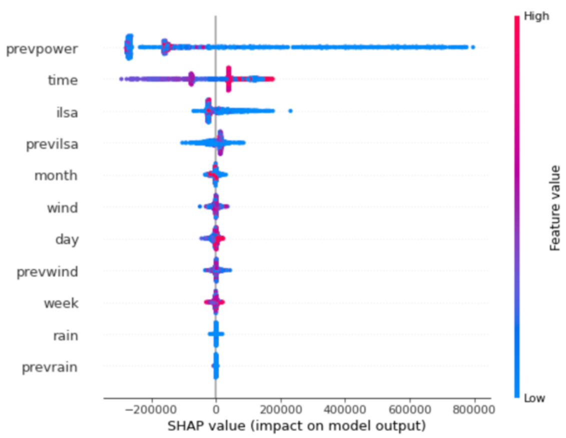


Figure 10. SHAP value of LGBM PV power generation

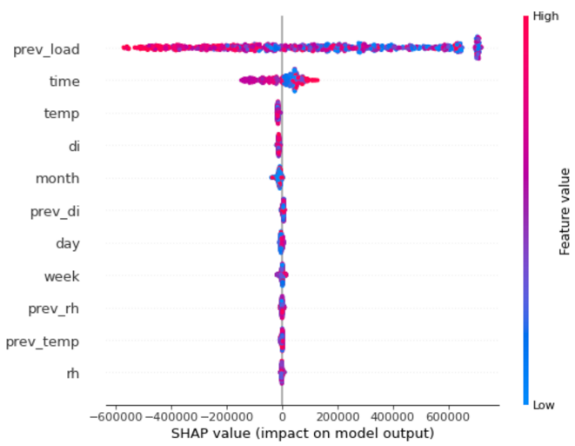


Figure 11. SHAP value of LGBM power load

Through the results, the two prediction models showed excellent performance in predicting PV power generation and load volume, and the validity of the results was shown.

V. CONCLUSIONS

We developed PV power generation and power load prediction models using deep learning and machine learning that can increase renewable energy efficiency to help companies convert renewable energy from existing power, and provided the validity of the results using explainable artificial intelligence. The excellent performance of the prediction model was derived by substituting appropriate variables for the time when the power generation could not be measured, and which values influenced the result value was derived through the SHAP algorithm. Research to increase energy efficiency in a small power grid by applying a prediction model developed in the future to a building or home unit and scheduling it based on the prediction results is expected to be possible.

ACKNOWLEDGMENT

This work was supported by the Technology Development Program to Solve Climate Changes through the National Research Foundation of Korea (NRF) by the Ministry of Science and ICT under Grant NRF-2021M1A2A2065447 and by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2022-0-00106).

REFERENCES

[1] FELLOUS, Jean-Marc, *Explainable artificial intelligence for neuroscience: behavioral neurostimulation,* Frontiers in neuroscience, 2019.

[2] SARRIETA, Alejandro Barredo, *Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI*, Information fusion, pp. 82-115, 2020.

[3] Lundberg, Scott M., and Su-In Lee, *A unified approach to interpreting model predictions*, Advances in neural information processing systems, 2017.

[4] Tjoa, Erico, and Cuntai Guan, *A survey on explainable artificial intelligence (xai): Toward medical xai*, IEEE transactions on neural networks and learning systems 32.11, pp. 4793-4813, 2020.

[5] Gómez-Gil, Pilar, *A neural network scheme for long-term forecasting of chaotic time series*, Neural Processing Letters 33.3 pp. 215-233, 2011.

[6] Abdel-Nasser, Mohamed, and Karar Mahmoud, *Accurate photovoltaic power forecasting models using deep LSTM-RNN*, Neural Computing and Applications, pp. 2727-2740, 2019.

[7] Le, Tuong, *Improving electric energy consumption prediction using CNN and Bi-LSTM*, Applied Sciences, 2019.

[8] Muralitharan, K., Rathinasamy Sakthivel, and R. Vishnuvarthan, *Neural network based optimization approach for energy demand prediction in smart grid*, Neurocomputing pp. 199-208, 2018.

[9] ZAREMBA, Wojciech, SUTSKEVER, Ilya, VINYALS, Oriol, *Recurrent neural network regularization*, arXiv preprint arXiv:1409.2329, 2014.

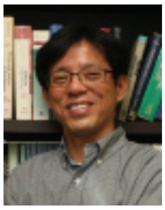
[10] Al-Dahidi, Sameer, *Extreme learning machines for solar photovoltaic power predictions*, Energies, 2018.

[11] MADONDO, Malvern, GIBBONS, Thomas, *Learning and modeling chaos using lstm recurrent neural networks*, In: MICS 2018 proceedings. 2018.

[12] Voyant, Cyril, *Machine learning methods for solar radiation forecasting: A review*, Renewable Energy pp. 569-582, 2017.

[13] Yu, Yunjun, Junfei Cao, and Jianyong Zhu, *An LSTM short-term solar irradiance forecasting under complicated weather conditions*, IEEE Access 7 pp. 145651-145666, 2019.
[14] Ke, Guolin, *Lightgbm: A highly efficient gradient boosting decision tree*, Advances in neural information processing systems 30, 2017.
[15] Wang, Zhen, and Li Jia, *Short-Term Photovoltaic Power Generation Prediction Based on LightGBM-LSTM Model*, 2020 5th International Conference on Power and Renewable Energy (ICPRE). IEEE, 2020.
[16] Jeong, Jin Hwa, and Young-Tae Chae, *Improvement for forecasting of photovoltaic power output using real time weather data based on machine learning*, J. Korean Soc. Living Environ. Sys 25, pp. 119-125, 2018.
[17] Kraemer, Frank Alexander, *Solar energy prediction for constrained IoT nodes based on public weather forecasts,* Proceedings of the Seventh International Conference on the Internet of Things, 2017.
[18] Alanazi, Mohana, Abdulaziz Alanazi, and Amin Khodaei. *Long-term solar generation forecasting*, 2016 IEEE/PES transmission and distribution conference and exposition (T&D). IEEE, 2016.

SooHyun Shim received the B.E. and M.S. degrees in computer engineering from Sejong University, Seoul, South Korea, in 2019 and 2021, where he is currently working toward the Ph.D. degree in computer engineering.



Hyo-Sik Yang received the B.E. degree in information and communication engineering from Myongji University, Yongin, South Korea, in 1998, and the M.S. and Ph.D. degrees in electrical engineering from Arizona State University, Tempe, AZ, USA, in 2000 and 2005, respectively. He joined Sejong University, Seoul, Korea, in 2006, where he is currently an Associate Professor with the Department of Computer Science and Engineering. Before joining Sejong University, he was an Assistant Professor with Kyungnam University, Masan, Korea. He served as a Faculty Research Associate with Arizona State University until 2005. His research interests include wavelength-division multiplexing (WDM) all-optical networks, mobile ad hoc networks, and smart grid, in particular, on IEC 61850 and communication architecture.