Dynamic Energy Consumption Forecasting: Exploring the Performance of Machine Learning Algorithms in Dynamic Environment

Yahaya Abdullahi, Bowoo Kim, Dongjun Suh*

Kyungpook National University

{yahaya, kbw5913, *dongjunsuh}@knu.ac.kr

동적 에너지 소비 예측: 동적 환경에서 머신 러닝 알고리즘 성능 탐색

Yahaya Abdullahi, 김보우, 서동준*

*경북대학교

Abstract

Efficient energy management in buildings relies heavily on accurately predicting energy demand. However, the challenge lies in selecting the most suitable algorithm for energy forecasting. This study compares the performance of three machine learning models in predicting energy consumption under changing demand conditions by utilizing a real-world dataset greatly affected by the COVID-19 pandemic, resulting in notable alterations in consumption patterns.

Ⅰ. Introduction

The European Union (EU) aims to reduce greenhouse gas emissions and become climate-neutral by 2050. Among various contributors, the energy sector stands out as the most significant, accounting for 78% of emissions in the EU [1]. The buildings sector is the largest energy consumer within the energy sector, accounting for 40% of total energy usage in the EU [2]. As lifestyles and habits change, the energy demands of buildings evolve, making accurate energy demand forecasting essential.

Recent advancements in computing and cloud technologies have enabled researchers to use data-driven modeling techniques for building control and energy management. This study compares the performance of three data-driven models - RNN, LSTM, and a hybrid RNN-LSTM in predicting energy consumption based on past observations. The analysis centers on scenarios with notable energy consumption changes due to measures implemented in response to the COVID-19 pandemic.

Ⅱ. Method

2.1 Dataset

The dataset comprises energy consumption data from a seven-floor, mixed-use academic building at the University of Twente in the Netherlands, spanning three years (January 1st, 2018, to December 31st, 2021) [3]. It includes hourly records of electricity, heat, and water consumption. Figure 1 displays histograms of the selected environmental and energy consumption data, visually representing the data distribution. Meteorological data was collected from the local weather station in the Netherlands Portal [4].

2.2 Experimental Analysis

The models were implemented in Python 3 using TensorFlow with the Keras API. Two sets of input features were evaluated:

Set 1: Energy demand + water and heat consumption.

Set 2: Energy demand + weather conditions.

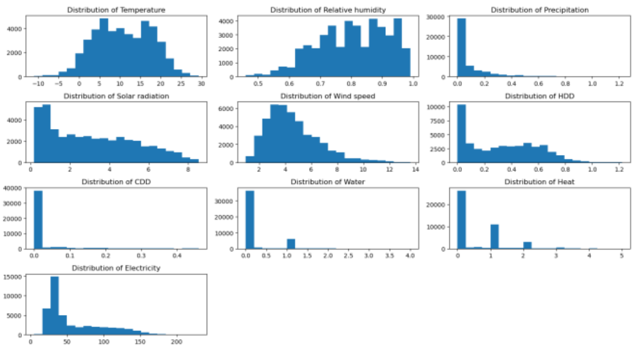
Fig. 1 Distribution of data

The Pearson correlation coefficient was used to select the most relevant input features, as shown in Figure 2. By using the Pearson correlation coefficient to select input features, the objective is to identify those exhibiting a strong linear relationship with the target variable. This approach aims to reduce model bias by concentrating on features significantly contributing to the prediction task.

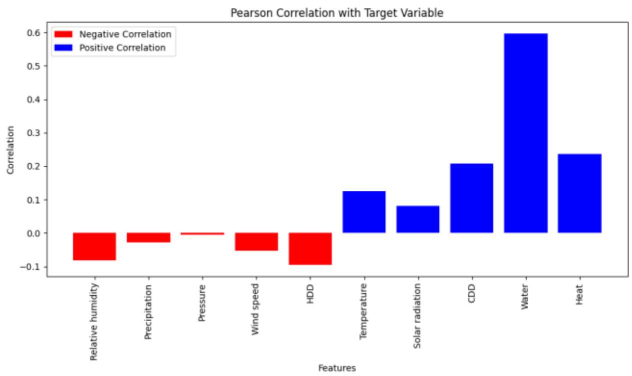

Fig. 2 Feature ranking selection

2.3 Result and Discussion

All models were trained using data recorded before the implementation of restriction measures (March 2020) and tested on data from periods when the restrictions were in place. The performance of each model was evaluated using three error metrics: CV-RMSE, MAPE, and R^2.

Table 1 and 2 present the results of the experiment considering the feature combination. The integration of water and heat consumption data (Set 1) significantly enhances the prediction performance of all models compared to Set 2, which considers weather features. This improvement was attributed to the strong correlation between the building's energy consumption patterns and water and heat consumption despite the lack of a direct correlation with weather patterns. Incorporating date and time features also boosted prediction performance across all models. Interestingly, LSTM and RNN models perform better when using Set 2 input features, while the hybrid approach models excel when incorporating Set 1 input features. Although the performance differences may seem small, they indicate notable improvements in accuracy.

Table 1 Model Result for Set 1 Features

Models	CV-RMSE	MAPE	R^2
RNN	0.259	0.253	0.890
LSTM	0.251	0.211	0.894
RNN-LSTM	0.189	0.197	0.902

Table 2 Model Result for Set 2 Features

Models	CV-RMSE	MAPE	R^2
RNN	0.304	0.382	0.788
LSTM	0.153	0.289	0.824
RNN-LSTM	0.351	0.421	0.774

Ⅲ. Conclusion

In conclusion, the findings of this study highlight the significance of carefully selecting input features and considering simple models before investing resources in elaborate hybrid approaches. Furthermore, the potential of water and heat consumption data in enhancing energy demand predictions warrants additional exploration. Future research should aim to provide more concrete answers regarding the practicality and effectiveness of integrating such data in energy prediction models.

ACKNOWLEDGMENT

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIT) (No. 2021R1I1A3049503) and (No. 2021R1A5A8033165).

REFERENCES

[1] European Commission, DG Climate ActionEuropean Environment Agency (2020).
[2] Owolabi AB, Yahaya A, Li HX, Suh D. Analysis of the Energy Performance of a Retrofitted Low-Rise Residential Building. *Sustainability*. 2023; 15(16):12129. https://doi.org/10.3390/su151612129
[3] U.o. Twente, University of Twente energy data, in, 2022.
[4] K.N.M. Insitituut, Daggegevens van het in Nederland 2022.