buildings

MDPI

Article

Anomaly Detection Based on LSTM Learning in IoT-Based Dormitory for Indoor Environment Control

Seol-Hyun Noh and Hyeun Jun Moon *

Department of Architectural Engineering, Dankook University, Yongin-si 16890, Republic of Korea; shnoh92@gmail.com

* Correspondence: hmoon@dankook.ac.kr

Abstract: This study focuses on gathering environmental data concerning the indoor climate within a dormitory, encompassing variables such as air temperature, relative humidity, CO_2 concentration, fine dust concentration, illuminance, and total volatile organic compounds. Subsequently, an anomaly detection long short-term memory model (LSTM) model, utilizing a two-stacked LSTM model, was developed and trained to enhance indoor environment control. The study demonstrated that the trained model effectively identified anomalies within eight environmental variables. Graphical representations illustrate the model's accuracy in anomaly detection. The trained model has the capacity to monitor indoor environmental data collected and transmitted using an Internet-of-Things sensor. In the event of an anomaly domain prediction, it proactively alerts the building manager, facilitating timely indoor environment control. Furthermore, the model can be seamlessly integrated into indoor environment control systems to actively detect anomalies, thereby contributing to the automation of control processes.

Keywords: indoor environment control; anomaly detection; LSTM; IoT sensors

Citation: Noh, S.-H.; Moon, H.J. Anomaly Detection Based on LSTM Learning in IoT-Based Dormitory for Indoor Environment Control. *Buildings* **2023**, *13*, 2886. https://doi.org/10.3390/buildings13112886

Academic Editors: Chong Zhang, Rui Guo, Yunran Min and Chaoqun Zhuang

Received: 22 October 2023
Revised: 11 November 2023
Accepted: 14 November 2023
Published: 19 November 2023

1. Introduction

The indoor climate in contemporary buildings, encompassing air temperature, relative humidity, CO_2 concentration, fine dust concentration, illuminance, and total volatile organic compounds (TVOCs), plays a pivotal role in human health, comfort, and task productivity [1–4]. The quest for techniques that can effectively conserve energy while upholding clean and pleasurable indoor environments is of paramount importance and has garnered significant attention from both society and academia [1,2,5]. The continuous evolution of computing and machine learning technologies has significantly contributed to optimizing human comfort and health, as well as reducing energy consumption [6,7].

In recent studies, machine learning has been leveraged for anomaly detection. X. Liu and Nielsen [8] pioneered an online anomaly detection system based on predictive analysis using smart meter data. Wang et al. [9] employed the k-nearest neighbor (k-NN) method to anticipate and identify unusual occurrences in relation to the thermal comfort of building occupants. Xin et al. [10] proposed an artificial neural network (ANN) approach, implementing an overlapped moving window, for predicting indoor air temperature and relative humidity. For detecting anomalies in semiconductor manufacturing processes, Kim et al. [11] developed a recurrent neural network model. For anomalies in Internet-of-Things (IoT) communication, Xu et al. [12] introduced an enhanced long short-term memory (LSTM) model. Taylor et al. [13] harnessed the potential of LSTM networks for anomaly detection in automobile control network data. Han et al. [14] investigated in-vehicle concentration levels of CO_2, comparing the accuracy of an autoregressive integrated moving average (ARIMA) model and LSTM model in predicting the change in CO_2 concentration. Ji et al. [15] proposed a hybrid neural network (HNN) prediction model (CNN-BiLSTM-Attention) based on deep learning (DL) for predicting the exhaust gas temperature (EGT) of marine diesel engines. In [14,15], graphs showing abnormal sections were not presented.

The research question of this study is as follows:

- What data should be collected and what methods should be used to develop technology that comfortably controls the indoor environment?

To address this research question, we utilized environmental data from eight variables, collected over a period, covering 20 bedrooms and 20 study rooms in a dormitory. Based on the collected data, we constructed and trained an LSTM model tailored for identifying anomalies surpassing a predefined threshold in indoor environments.

While our dataset represented big data, it adhered to a time series structure. Consequently, our anomaly detection LSTM model was designed based on two sequentially stacked LSTM models. The efficacy of this stacked LSTM model for detecting anomalies in time-series data has been corroborated by Malhotra et al. [16]. We have named this configuration the LSTM-based anomaly detection (LSTM-AD) model.

The trained model boasts the capability to monitor indoor environmental data, which are collected and transmitted by an IoT sensor, and to pre-emptively alert the monitoring center to potential or detected anomalies. This proactive approach empowers indoor environment control and safety management. Furthermore, these data may serve as a foundation for developing an indoor environment control system equipped with the trained model.

The machine learning-based anomaly detection model introduced in this study is engineered to be adaptable across various applications, enabling the active identification of anomalies in indoor environments and aiding in the automation of indoor environment control processes. Moreover, it is envisioned that this model will bolster the efforts of the monitoring center in maintaining and managing indoor environments.

The main contributions of this study are as follows:

- Based on big data corresponding to the eight environmental variables collected at one-minute intervals from 20 bedrooms and 20 study rooms in a dormitory between February 2022 and September 2023, we developed an LSTM-AD model for use in indoor environment control. The model was validated based on the measured performance metrics.
- To augment the performance of the developed LSTM-AD model for anomaly detection, we estimated the optimal threshold by comparing multiple thresholds derived through trial and error with the optimal threshold suggested by Noh (2023) [17]. Additionally, we produced graphical representations to compare predicted values against actual values in the test dataset, facilitating a thorough examination of the model's performance for anomaly detection. To reinforce the model's validation through visual representations of adequate anomaly detection, we have generated graphs illustrating the anomaly score and the anomaly domain indicated by predicted values. Importantly, the source code for this novel model is openly accessible in the public domain, facilitating its integration into indoor environment control systems.

This study is structured as follows: Section 2 presents the data acquired from 20 bedrooms and 20 study rooms in a dormitory, and provides an in-depth description of the LSTM-AD model employed for indoor environmental anomaly detection and outlines the model training methodologies. Section 3 highlights an exhaustive account of the performance metrics of the trained LSTM-AD model. Section 4 presents a summary of the performance results presented in Section 3. Finally, Section 5 offers a concise summary of the key findings of this research.

2. Materials and Methods

2.1. Data

In this study, we developed an LSTM-AD model for detecting anomalies within indoor environmental data. The model was trained using data from 20 bedrooms and 20 study rooms in a dormitory, encompassing eight environmental variables. Subsequently, we

evaluated the model's performance, and the data for these eight variables collected from the dormitory bedrooms and study rooms are presented in Table 1.

Table 1. The eight variables of the indoor environment measured in a dormitory.

Variable	Definition and Unit
Temperature	Indoor air temperature (°C)
Relative Humidity	Relative humidity (%)
CO_2	Carbon dioxide concentration (ppm)
Dust_pm_0.1	Ultrafine dust particles of $\leq$ 0.1 μm diameter $(\mu g/m^3)$
Dust_pm_1.0	Fine dust particles of $\leq$ 1.0 μm diameter $(\mu g/m^3)$
Dust_pm_2.5	Fine dust particles of $\leq$ 2.5 μm diameter $(\mu g/m^3)$
Illuminance	Illuminance (lux)
TVOC	Total Volatile Organic Compounds level (ppb)

The data for these eight environmental variables were collected at one-minute intervals from 20 bedrooms and 20 study rooms in a dormitory, spanning the period from 3 February 2022 to 5 September 2023 (Table 1). Figures 1 and 2 provide data in graphical representations of the data collected from bedroom #303 and study room #303, respectively.

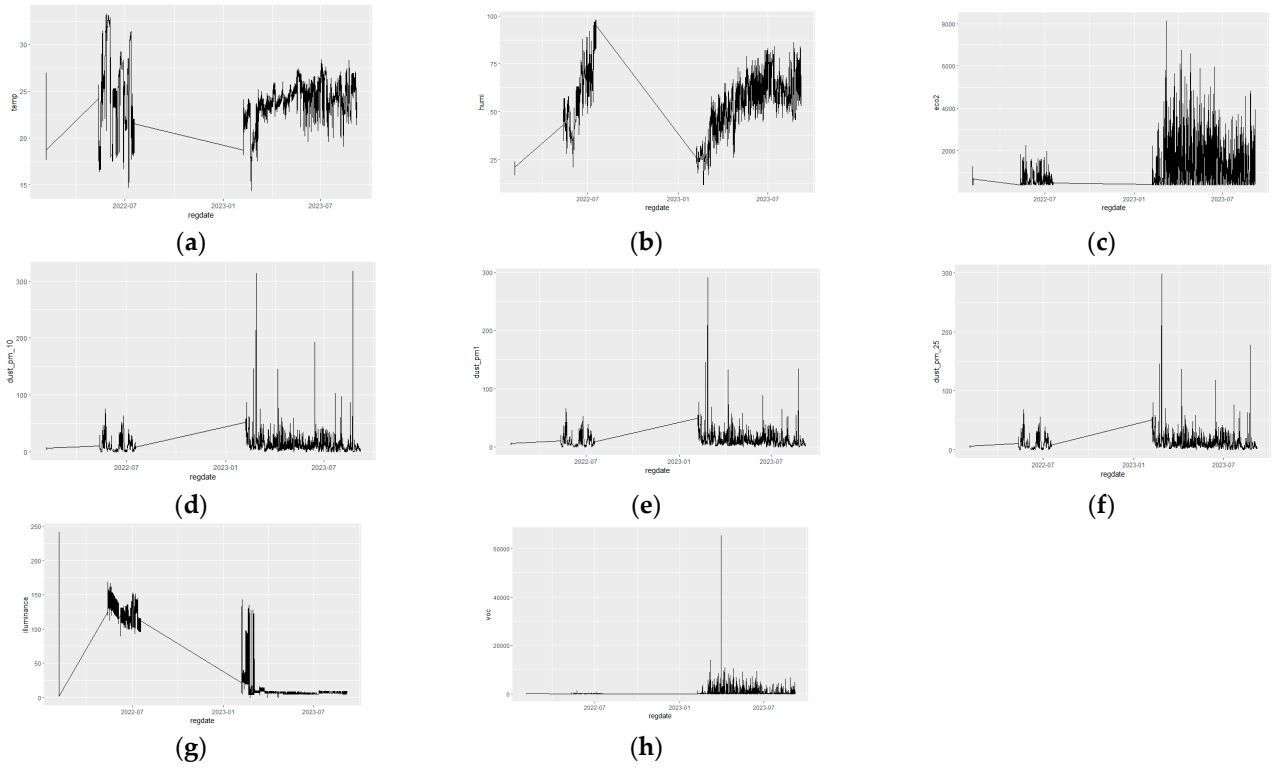


Figure 1. Time series graphs for the environmental data collected from bedroom #303 in the dormitory. (**a**) Temperature time series; (**b**) humidity time series; (**c**) CO_2 time series; (**d**) Dust_pm_0.1 time series; (**e**) Dust_pm_1.0 time series; (**f**) Dust_pm_2.5 time series; (**g**) illuminance time series; (**h**) TVOC time series.

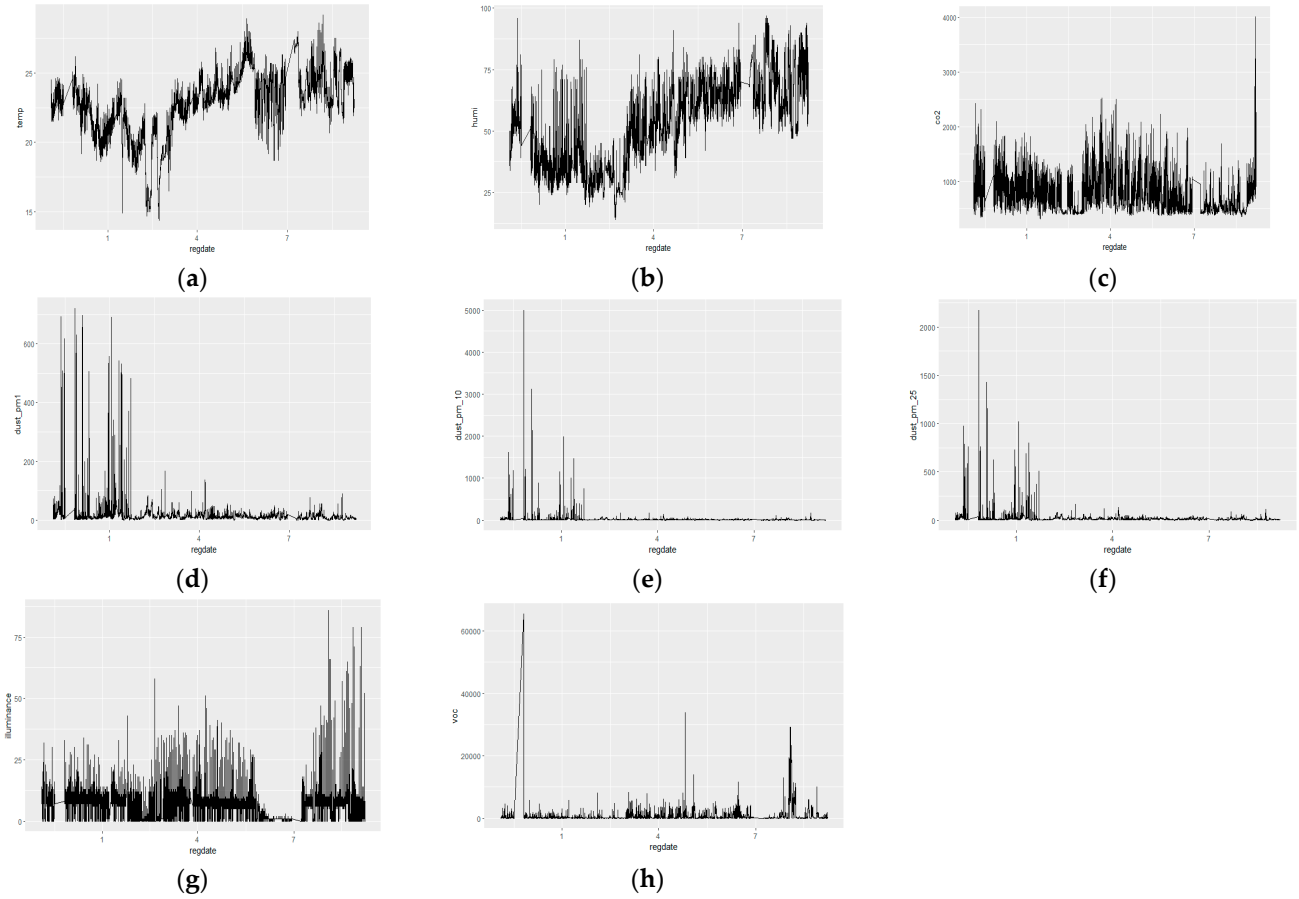


Figure 2. Time series graphs for the environmental data collected from study room #303 in the dormitory. (**a**) Temperature time series; (**b**) humidity time series; (**c**) CO_2 time series; (**d**) Dust_pm_0.1 time series; (**e**) Dust_pm_1.0 time series; (**f**) Dust_pm_2.5 time series; (**g**) illuminance time series; (**h**) TVOC time series.

The period of time where the graph progresses as a straight line in Figure 1 indicates a domain without measured values due to communication error.

As presented in Figures 1 and 2, anomalies outside the 95% CI based on the mean are evident in each time series graph. These measured values of environmental data within such domains exceed an adequate threshold level, signaling the need for environmental control.

Through unsupervised training, the LSTM-AD model learns to distinguish between normal and abnormal values. Subsequently, we verified the effectiveness of anomaly detection using data from 2023. As mentioned in the Introduction, the source code for this innovative model is available in the public domain. This accessibility enables its integration into various applications, facilitating the active detection of anomalies in indoor environments. It also contributes to the automation of indoor environment control and supports the maintenance and management of indoor environments by the monitoring center.

2.2. Model

As previously outlined in the Introduction, Malhotra et al. [16] provided empirical support for the LSTM-AD model's proficiency in detecting anomalies, with successful applications to the ECG dataset, Space Shuttle Marotta value time series, and Power demand dataset. In this study, the model underwent unsupervised training, utilizing the environmental data from 2022 as the training dataset. It was designed to flag an abnormal value when the discrepancy between the model-predicted value and the actual value surpassed a predetermined threshold. The model's performance was then assessed using

the indoor environmental data from 2023, serving as the test dataset, with a focus on key performance metrics: precision, recall, F1 score, and AUC.

The LSTM-AD model in this research is underpinned by two stacked LSTM models, as illustrated in Figure 3. For a detailed understanding of the model's architecture, an individual LSTM model is presented in Figure 4.

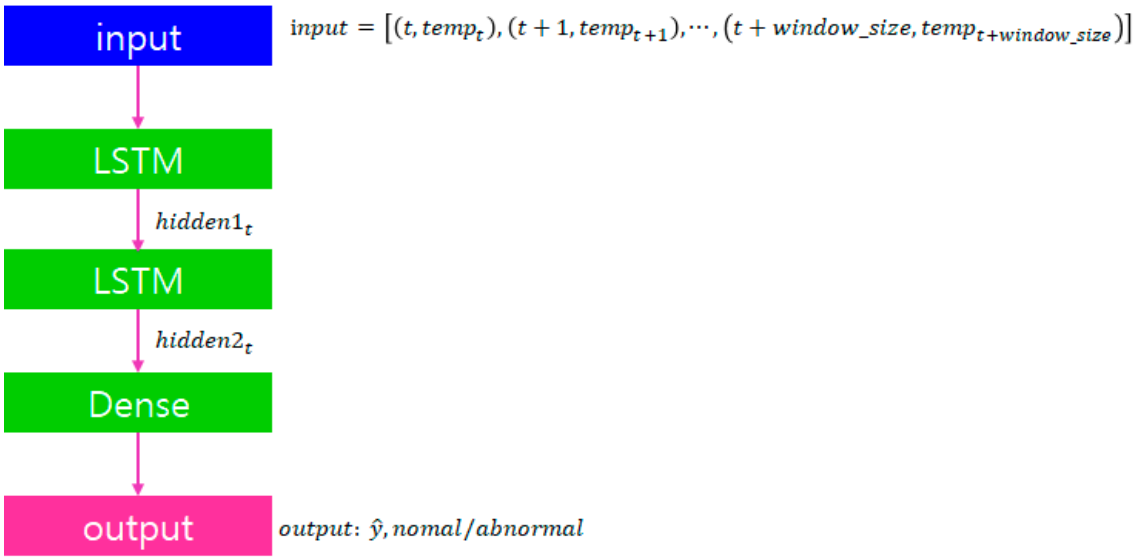


Figure 3. A stacked LSTM model (LSTM-AD).

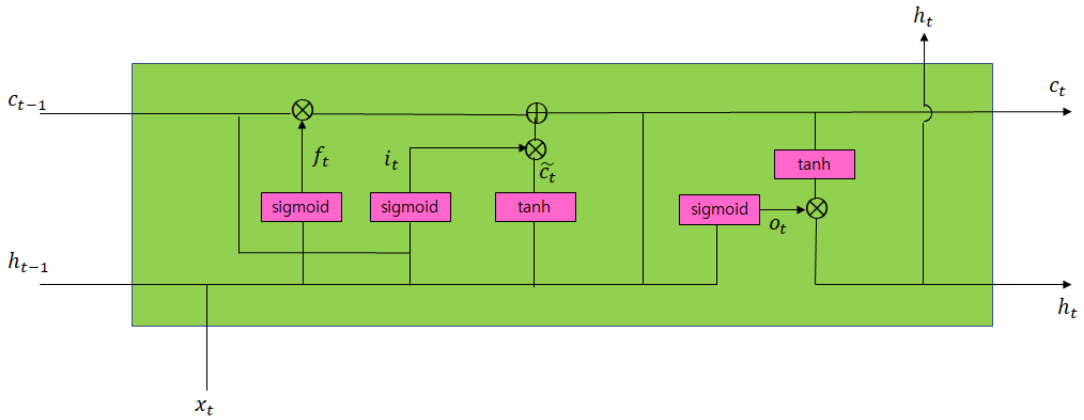


Figure 4. The LSTM model [18].

The model constructed in this study was based on the two stacked LSTM models, as depicted in Figure 5.

To facilitate the detection of abnormal values, the model systematically compares the difference between the predicted and actual values with the predetermined threshold, applied to the environmental data collected from each bedroom and study room. In the quest for optimal anomaly detection, various threshold values, including those previously recommended as optimal, are rigorously tested to ensure that the established threshold signifies exceptional performance.

```
class AnomalyLSTM(pl.LightningModule):
    def __init__(self, hparams: dict = None):
        super().__init__()
        if hparams is None:
            hparams = {}
        self.hparams.update(hparams)
        self.threshold = self.hparams["threshold"]
        self.lstm_encoder = torch.nn.LSTM(
            input_size=hparams["input_size"],
            hidden_size=128,
            batch_first=True,
        )
        self.lstm_decoder = torch.nn.LSTM(
            input_size=128,
            hidden_size=hparams["input_size"],
            batch_first=True,
        )
        self.mse = torch.nn.MSELoss(reduction='none')
```

Figure 5. The model used in this study to detect anomalies in indoor environmental data from the dormitory bedrooms and study rooms.

Table 2 presents the mean, standard deviation (SD), and 95% confidence interval (CI) for each environmental variable observed in the data collected from both bedroom #303 and study room #303. Table 3 presents the mean, SD, and 95% CI, averaging the data from 20 bedrooms and 20 study rooms within the dormitory.

As shown in Tables 2 and 3, the environmental data exhibit considerable variations across different bedrooms, while the study rooms exhibit negligible differences owing to relatively well-controlled conditions. Notably, the boundaries of the 95% CI in both bedrooms and study rooms significantly surpass the predetermined threshold, underscoring the necessity for environment control through effective anomaly detection.

Table 2. The mean, standard deviation, and 95% confidence interval for the environmental variables measured for bedroom #303 and study room #303.

Bedroom No.	Variable	Mean	Standard Deviation	95% Confidence Interval
#303	Temperature	24.3909	2.7604	[18.9801, 29.8012]
	Humidity	54.7534	15.1651	[25.0305, 84.4764]
	CO_2	845.1237	704.019	[0, 2224.9760]
	Dust_pm_0.1	10.7271	10.5005	[0, 31.3077]
	Dust_pm_1.0	11.1798	11.2659	[0, 33.2606]
	Dust_pm_2.5	10.7981	10.6723	[0, 31.7155]
	Illuminance	34.0503	49.4213	[0, 130.9143]
	TVOC	440.3392	1022.464	[0, 2444.332]
Study Room No.	**Variable**	**Mean**	**Standard Deviation**	**95% Confidence Interval**
#303	Temperature	22.7116	2.2795	[18.2437, 27.1794]
	Humidity	51.4847	17.2598	[17.6561, 85.3132]
	CO_2	739.9548	310.6303	[131.1307, 1348.7790]
	Dust_pm_0.1	12.4174	14.5863	[0, 41.0060]
	Dust_pm_1.0	12.771	23.0179	[0, 57.8853]
	Dust_pm_2.5	12.5012	16.598	[0, 45.0327]
	Illuminance	5.8397	5.5818	[0, 16.7799]
	TVOC	591.1457	2074.741	[0, 4657.5630]

Table 3. The mean, standard deviation, and 95% confidence interval averaged from the data of the environmental variables measured for 20 bedrooms and 20 study rooms in the dormitory.

Location	Variable	Mean	Standard Deviation	95% Confidence Interval
20 bedrooms on average	Temperature	24.4377	3.40808	[17.7579, 31.1174]
	Humidity	59.07272	17.38868	[24.9916, 93.1539]
	CO_2	958.64188	886.38118	[0, 2695.9170]
	Dust_pm_0.1	24.43332	120.88044	[0, 261.3546]
	Dust_pm_1.0	24.85702	121.24792	[0, 262.4986]
	Dust_pm_2.5	24.50702	120.96248	[0, 261.5891]
	Illuminance	50.68842	61.80158	[0, 171.8173]
	TVOC	613.73264	1390.1088	[0, 3338.295]
20 study rooms on average	Temperature	23.5726	7.16634	[9.5285, 37.6184]
	Humidity	52.02044	19.97088	[12.8782, 91.1626]
	CO_2	758.34338	541.98324	[64.9457, 1820.6110]
	Dust_pm_0.1	10.41324	12.14486	[0,34.2167]
	Dust_pm_1.0	10.65804	14.25156	[0, 38.5906]
	Dust_pm_2.5	10.45922	12.63612	[0, 35.2256]
	Illuminance	10.4937	11.41642	[0, 33.1015]
	TVOC	576.94022	1651.17796	[0, 3813.1890]

The model, trained on the 2022 dataset, is integrated into the system. Subsequently, the model is tested with the 2023 dataset to predict anomalies, and the results are presented in Section 4. The performance is evaluated through the comparison of the predicted outcomes with the actual data, focusing on key performance metrics, such as precision, recall, F1 score, and AUC.

3. Performance Results

The LSTM-AD model in this study defines the loss function based on the mean square error (MSE) between the predicted values, $\hat{y}$, and the actual values, y. The optimal threshold estimated through multiple experiments was 0.01.

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

The values surpassing the threshold of 0.01 were identified as anomalies. Key hyper-parameters are as follows:

window_size: 16; threshold: 0.01; tolerance: 3; learning rate: 0.001; batch_size: 128

The window_size indicates that 16 sets of continuous data are grouped into a single individual input. As the batch_size is 128, the LSTM-AD model is trained using a single input of data at $n = 128$. This allows the model to readily detect a general data pattern and produce up to 128 predicted values for the test dataset, suggesting that 128 predicted values can be obtained before the 128 time slots. Consequently, anomalies in the data of environmental variables can be detected pre-emptively to enable environmental control. Furthermore, the tolerance parameter stipulates that an anomaly score above the threshold must be detected at least three times in sequence to define an anomaly domain. Figure 6 presents graphs comparing the actual values and the values predicted by the trained model for the 2023 test dataset.

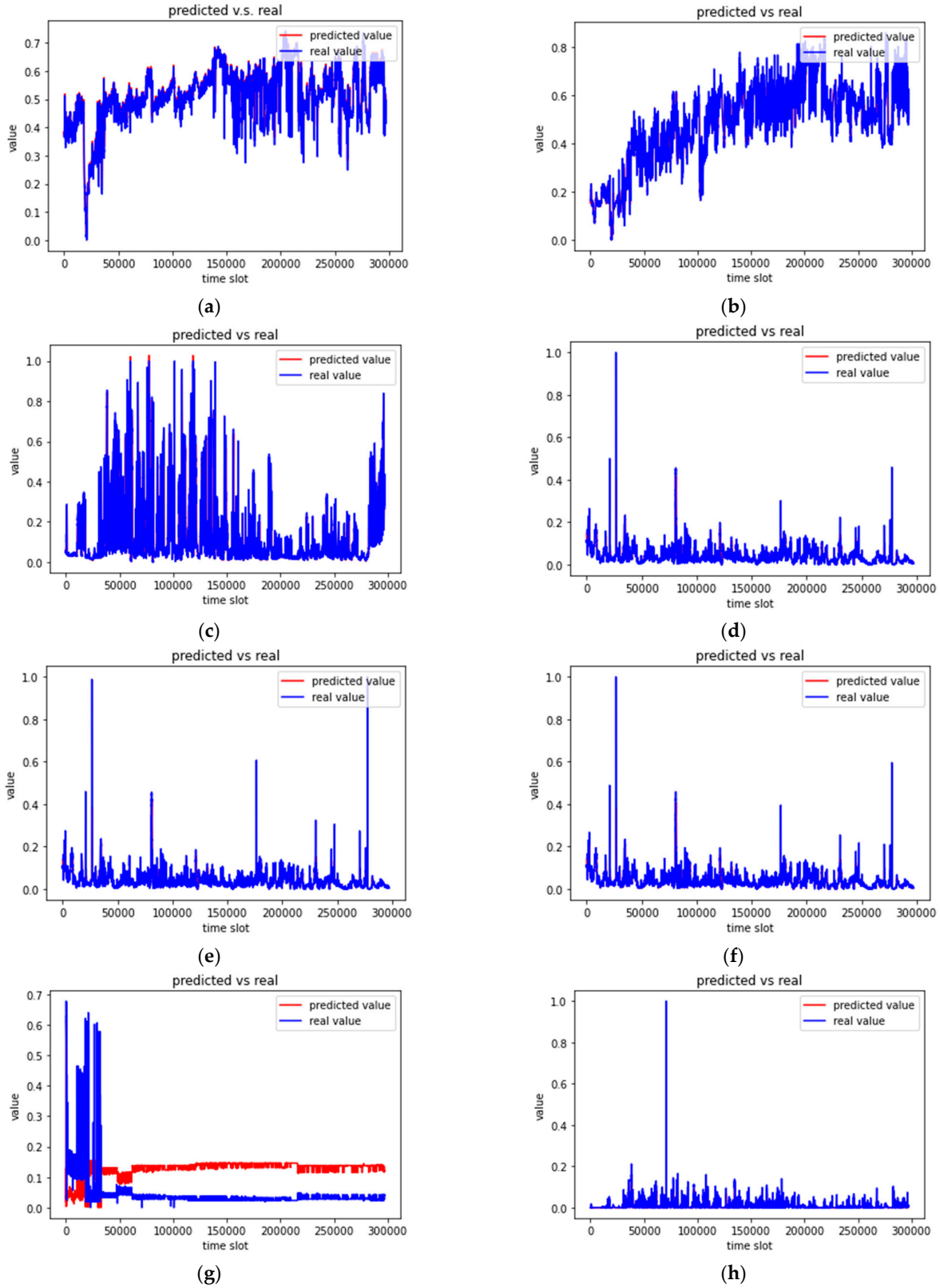
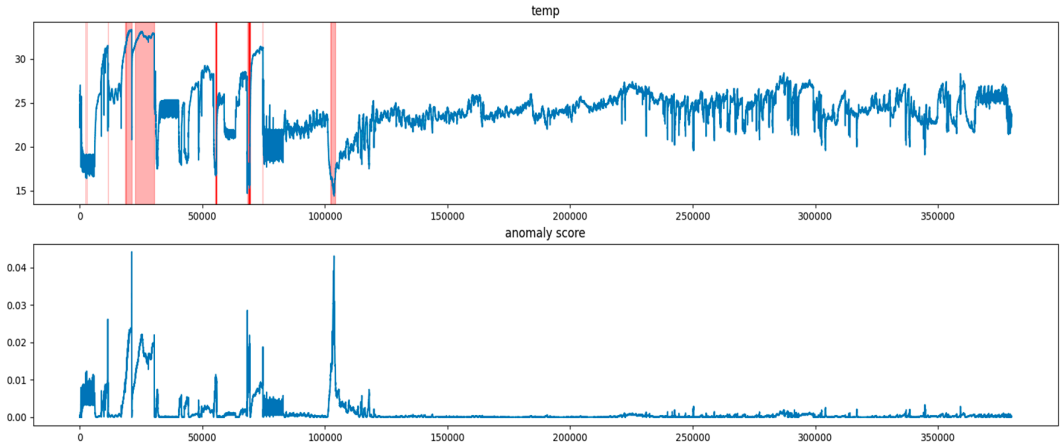


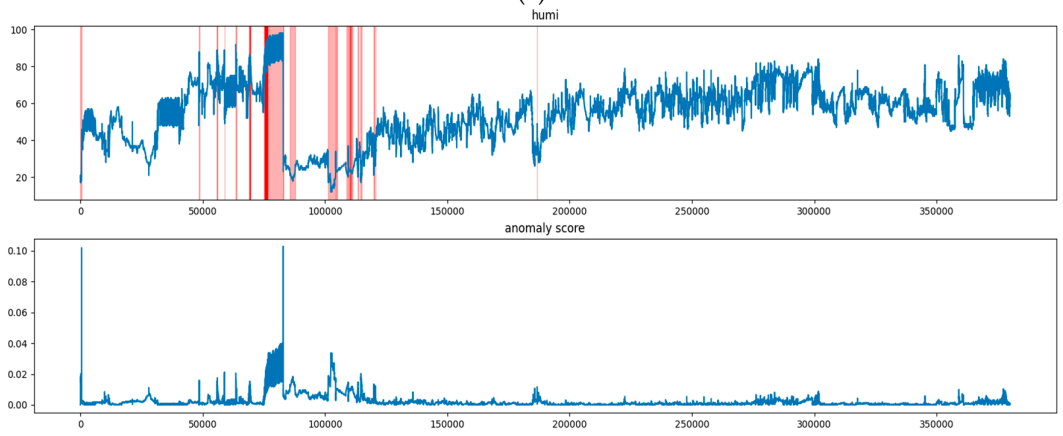
Figure 6. Graphs comparing the actual values and the values predicted by the LSTM-AD model for the data of 2023 as the test dataset for the bedroom #303. (**a**) Values predicted by the LSTM-AD model for the normalized temperature data and the real values. (**b**) Values predicted by the LSTM-AD model for the normalized humidity data and the real values. (**c**) Values predicted by the LSTM-AD

model for the normalized CO_2 data and the real values. (**d**) Values predicted by the LSTM-AD model for the normalized dust_pm_0.1 data and the real values. (**e**) Values predicted by the LSTM-AD model for the normalized dust_pm_1.0 data and the real values. (**f**) Values predicted by the LSTM-AD model for the normalized dust_pm_2.5 data and the real values. (**g**) Values predicted by the LSTM-AD model for the normalized illuminance data and the real values. (**h**) Values predicted by the LSTM-AD model for the normalized TVOCs data and the real values.

Figure 7 provides additional graphs that highlight the effective anomaly detection capabilities of the LSTM-AD model when trained with the optimal threshold. These graphs showcase the comparison between the model-predicted values and the anomaly scores. The red sections in the graphs indicate domains classified as anomalies by the model, with large anomaly scores aligning with the detection of anomaly domains when the measured values of the environmental variables exceed the set threshold.

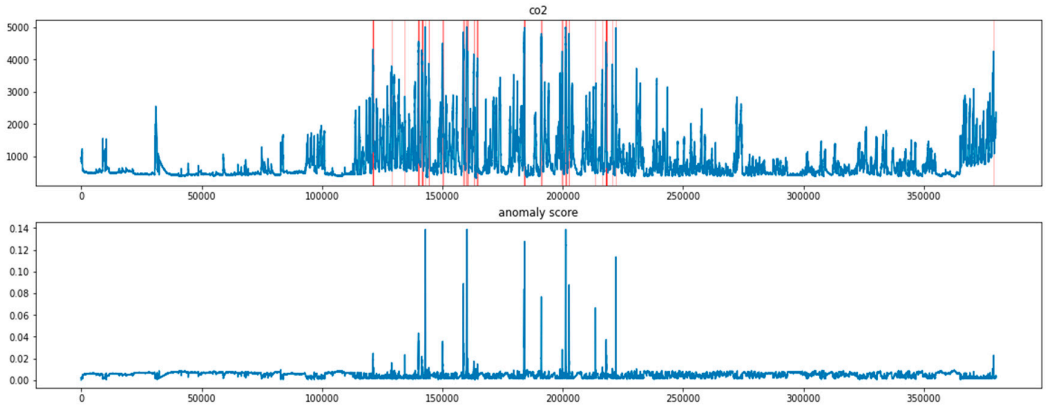


(**a**)

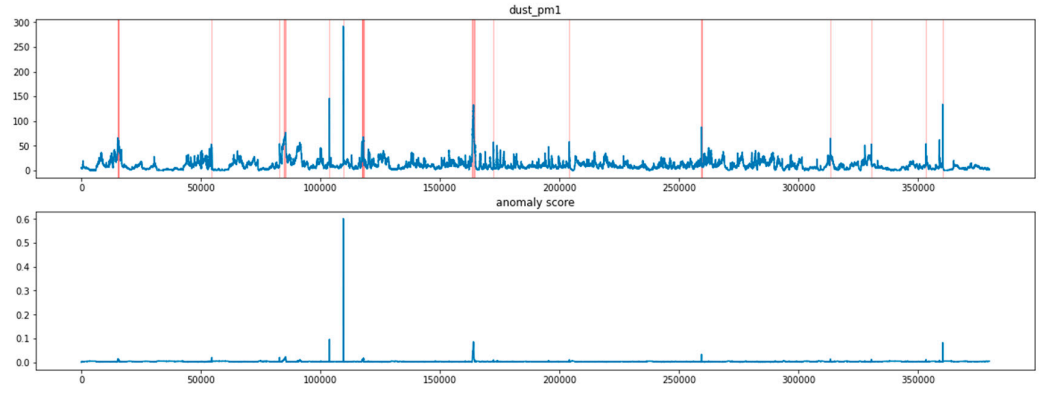


(**b**)

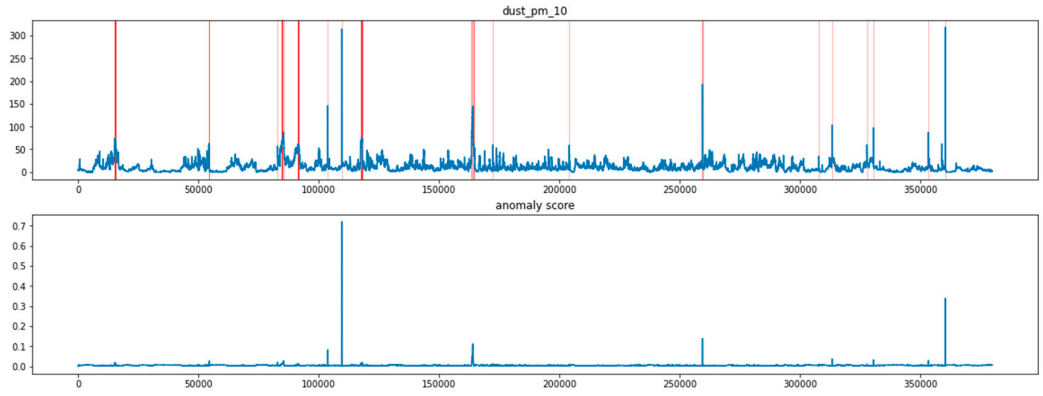
Figure 7. *Cont.*



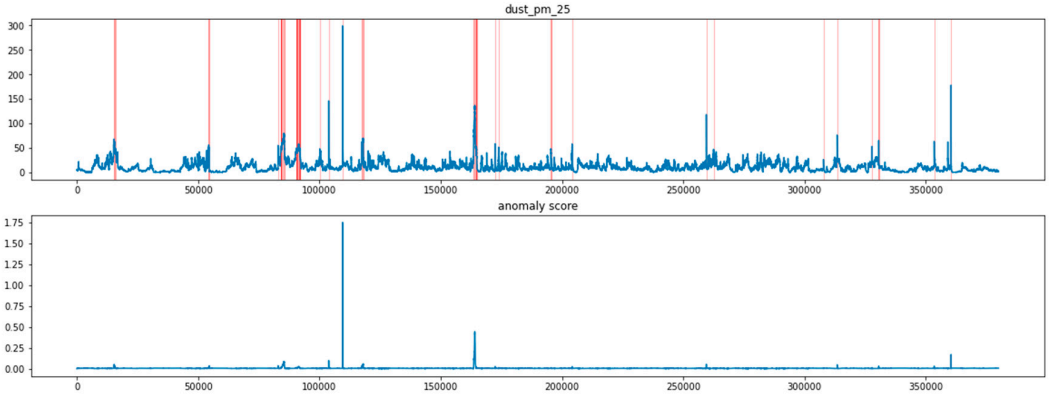
(**c**)



(**d**)

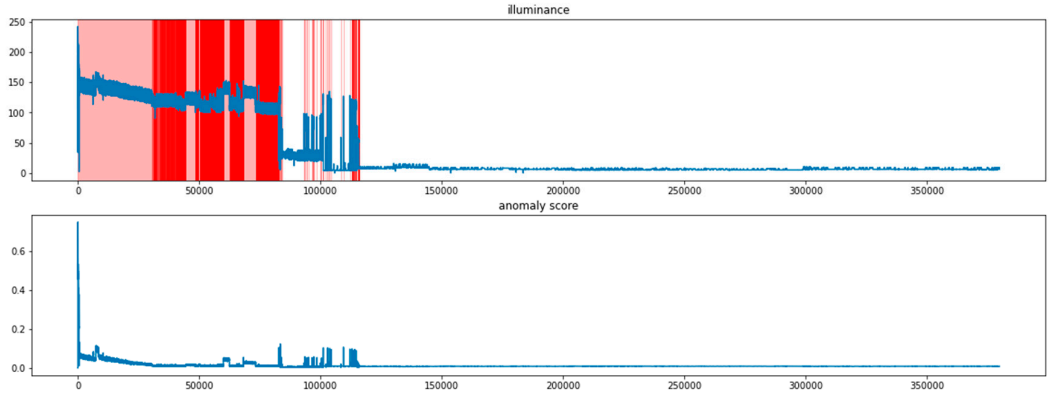


(**e**)

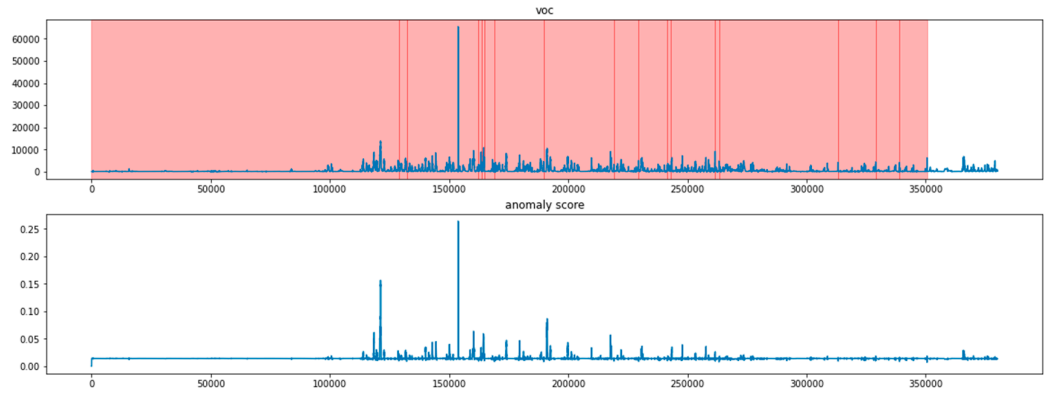


(**f**)

Figure 7. *Cont.*



(**g**)



(**h**)

Figure 7. Graphs showing the anomaly scores and the anomaly domains predicted by the LSTM-AD model for the eight environmental variables in the data from bedroom #303. (**a**) Graphs showing the anomaly domains predicted by the LSTM-AD model for the temperature data and the anomaly scores. (**b**) Graphs showing the anomaly domains predicted by the LSTM-AD model for the humidity data and the anomaly scores. (**c**) Graphs showing the anomaly domains predicted by the LSTM-AD model for the CO_2 data and the anomaly scores. (**d**) Graphs showing the anomaly domains predicted by the LSTM-AD model for the dust_pm_0.1 data and the anomaly scores. (**e**) Graphs showing the anomaly domains predicted by the LSTM-AD model for the dust_pm_1.0 data and the anomaly scores. (**f**) Graphs showing the anomaly domains predicted by the LSTM-AD model for the dust_pm_2.5 data and the anomaly scores. (**g**) Graphs showing the anomaly domains predicted by the LSTM-AD model for the illuminance data and the anomaly scores. (**h**) Graphs showing the anomaly domains predicted by the LSTM-AD model for the TVOCs data and the anomaly scores.

The LSTM-AD model employs past data to predict the state of the next time slot, distinguishing between normal and abnormal states. The model performance is quantified through various metrics, including normal state precision, recall, F1 score, and AUC. The definitions of these performance metrics are derived from the confusion matrix presented in Table 4.

$$Normal\ state\ Precision = \frac{TP}{TP + FP}$$

$$Normal\ state\ Recall = \frac{TP}{TP + FN}$$

$$Normal\ state\ F1\ Score = \frac{2 \times Normal\ state\ Precision \times Normal\ state\ Recall}{Normal\ state\ Precision + Normal\ state\ Recall}$$

Table 4. Confusion matrix.

		Prediction Outcome	
		Normal State	**Abnormal State**
Actual outcome	Normal state	True positive (TP)	False negative (FN)
	Abnormal state	False positive (FP)	True negative (TN)

The AUC represents the area between the ROC AUC curve and the x-axis, and an increase in AUC indicates an increase in the prediction power. Table 5 presents an overview of the estimated model performance based on the measured normal state precision, recall, *F*1 score, and AUC.

Table 5. The performance metrics of the LSTM-AD model for the data of eight environmental variables.

	Precision	**Recall**	**F1 Score**	**AUC**
Temperature	0.95	0.36	0.52	0.1789
Humidity	0.89	0.43	0.58	0.2156
CO_2	0.96	0.99	0.98	0.7850
Dust_pm_0.1	0.96	1.0	0.98	0.5
Dust_pm_1.0	0.96	1.0	0.98	0.5
Dust_pm_2.5	0.96	1.0	0.98	0.5
Illuminance	0.99	0.99	0.99	0.84
TVOC	0.94	1.0	0.97	0.5

Most of the data predominantly fall under the normal state, resulting in higher performance metrics for the normal state. Nonetheless, as evident from Figures 6 and 7, the model exhibits commendable predictive power for the abnormal state.

4. Discussion

Figure 6 presents graphs comparing the actual values and the values predicted by the trained model for the 2023 test dataset. As illustrated in Figure 6, the actual and model-predicted values exhibit significant similarity. Figure 7 showcases the comparison between the model-predicted values and the anomaly scores. The red sections in the graphs indicate the domains classified as anomalies by the model, with large anomaly scores aligning with the detection of anomaly domains when the measured values of environmental variables exceed the set threshold.

The LSTM-AD model, trained on the 2022 dataset, is integrated into the system. Subsequently, the model is tested with the 2023 dataset to predict anomalies, and the results are presented in Table 5. The performance is evaluated through the comparison of the predicted outcomes with the actual data, focusing on key performance metrics, normal state precision, recall, F1 score, and AUC. The estimated mean values for normal state precision, recall, F1 score, and AUC were found to be 0.95, 0.85, 0.87, and 0.82, respectively. These metrics underscore the remarkable performance of the model when dealing with normal-state scenarios.

5. Conclusions

This study involved the collection of environmental data pertaining to the indoor climate in a dormitory, encompassing parameters such as air temperature, relative humidity, CO_2 concentration, fine dust concentration, illuminance, and TVOCs. Subsequently, an LSTM-AD model was meticulously constructed and trained for the purpose of indoor

environment control. The effectiveness of this trained model in detecting anomaly domains, spanning eight environmental variables, is well-demonstrated through the graphical representations in Figures 6 and 7.

For the indoor environment control system, the LSTM-AD model, trained using the 2022 data, was seamlessly integrated. It was put to the test for predicting anomalies in the 2023 test dataset. The comparison between the predicted and actual outcomes allowed for the measurement of various performance metrics, including precision, recall, F1 score, and AUC. These metrics underscore the remarkable performance of the model when dealing with normal-state scenarios.

The trained model offers the capability to continuously monitor indoor environmental data collected and transmitted by IoT sensors. In the event of predicting an anomaly domain, the system can proactively notify the building manager, thereby enabling prompt and effective indoor environment control. Furthermore, the model is adaptable for integration into indoor environment control systems, facilitating the automation of anomaly detection processes.

In future studies, we envisage an expansion in the scope of indoor environmental data collection including the kitchen, coupled with the rigorous evaluation of various machine learning models to determine the optimal performance configuration. This ongoing exploration will further enhance the precision and applicability of anomaly detection in indoor environments.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/buildings13112886/s1.

Author Contributions: Conceptualization, H.J.M.; Methodology, S.-H.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Research Foundation of Korea (NRF) grant, funded by the Korean government (MSIT) (no. 2021R1A2B5B02002699). This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (no. 20212020800120).

Data Availability Statement: The code and datasets used and analyzed during this study can be obtained from the online resource in Supplementary Materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, Y.; Pang, Z.; Karlsson, M.; Gong, S. Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. *Build. Environ.* **2020**, *183*, 107212. [CrossRef]
2. Kim, J.; Kong, M.; Hong, T.; Jeong, K.; Lee, M. Physiological response of building occupants based on their activity and the indoor environmental quality condition changes. *Build. Environ.* **2018**, *145*, 96–103. [CrossRef]
3. Kim, J.; Hong, T.; Jeong, J.; Koo, C.; Kong, M. An integrated psychological response score of the occupants based on their activities and the indoor environmental quality condition changes. *Build. Environ.* **2017**, *123*, 66–77. [CrossRef]
4. Kim, J.; Hong, T.; Kong, M.; Jeong, K. Building occupants' psychophysiological response to indoor climate and CO_2 concentration changes in office buildings. *Build. Environ.* **2020**, *169*, 106596. [CrossRef]
5. IoT Analytics. *Predictive Maintenance Report 2019–2024*; IoT Analytics: Hamburg, Germany, 2019; Available online: https://iot-analytics.com/product/predictive-maintenance-report-2019-2024/ (accessed on 30 September 2023).
6. Dai, X.; Liu, J.; Zhang, X. A review of studies applying machine learning models to predict occupancy and window-opening behaviors in smart buildings. *Energy Build.* **2020**, *223*, 110159. [CrossRef]
7. Mason, K.; Grijalva, S. A review of reinforcement learning for autonomous building energy management. *Comput. Electr. Eng.* **2019**, *78*, 300–312. [CrossRef]
8. Liu, X.; Nielsen, P.S. Scalable prediction-based online anomaly detection for smart meter data. *Inf. Syst.* **2018**, *77*, 34–47. [CrossRef]
9. Wang, Z.; Parkinson, T.; Li, P.; Lin, B.; Hong, T. The Squeaky wheel: Machine learning for anomaly detection in subjective thermal comfort votes. *Build. Environ.* **2019**, *151*, 219–227. [CrossRef]
10. Shi, X.; Lu, W.; Zhao, Y.; Qin, P. Prediction of indoor temperature and relative humidity based on cloud database by using an improved BP neural network in Chongqing. *IEEE Access* **2018**, *6*, 30559–30566. [CrossRef]
11. Kim, C.; Lee, J.; Kim, R.; Park, Y.; Kang, J. DeepNAP: Deep neural anomaly pre-detection in a semiconductor fab. *Inform. Sci.* **2018**, *457–458*, 1–11. [CrossRef]

12. Xu, R.; Cheng, Y.; Liu, Z.; Xie, Y.; Yang, Y. Improved Long Short-Term Memory based anomaly detection with concept drift adaptive method for supporting IoT services. *Future Gener. Comput. Syst.* **2020**, *112*, 228–242. [CrossRef]
13. Taylor, A.; Leblanc, S.; Japkowicz, N. Anomaly detection in automobile control network data with long short-term memory networks. In Proceedings of the 2016 IEEE International Conference on Data Science and Advanced Analytics, DSAA, Montreal, QC, Canada, 17–19 October 2016; pp. 130–139.
14. Han, J.; Lin, H.; Qin, Z. Prediction and Comparison of In-Vehicle CO_2 Concentration Based on ARIMA and LSTM Models. *Appl. Sci.* **2023**, *13*, 10858. [CrossRef]
15. Ji, Z.; Gan, H.; Liu, B. A Deep Learning-Based Fault Warning Model for Exhaust Temperature Prediction and Fault Warning of Marine Diesel Engine. *Mar. Sci. Eng.* **2023**, *11*, 1509. [CrossRef]
16. Malhotra, P.; Vig, L.; Shroff, G.; Agarwal, P. Long Short Term Memory Networks for Anomaly Detection in Time Series. In Proceedings of the European Symposium on Artificial Neural Networks, Bruges, Belgium, 22–23 April 2015; pp. 89–94.
17. Noh, S.-H. Comparing the Performance of Corporate Bankruptcy Prediction Models Based on Imbalanced Financial Data. *Sustainability* **2023**, *15*, 4794. [CrossRef]
18. Noh, S.-H. Analysis of Gradient Vanishing of RNNs and Performance Comparison. *Information* **2022**, *12*, 442. [CrossRef]