



RESEARCH ARTICLE

실내 이산화탄소 농도 예측을 위한 기계학습 모델 검증

김효준¹ · 조영훈² · 류성룡^{3*}

¹영남대학교 일반대학원 건축학과 박사과정, ²영남대학교 건축학부 부교수,

³국립금오공과대학교 건축학부 부교수

Verification of Machine Learning Algorithm for CO₂ Prediction in Building

Kim, Hyo-Jun¹ · Cho, Young-Hum² · Ryu, Seong-Ryong^{3*}

¹Ph.D. student, Department of Architectural Engineering, Graduate School of Yeungnam University, Gyeongsan, Korea

²Associate Professor, School of Architecture, Yeungnam University, Gyeongsan, Korea

³Associate Professor, Department of Architectural Engineering, Kumoh National Institute of Technology, Gumi, Korea

*Corresponding author: Ryu, Seong-Ryong, Tel: +82-54-478-7592, E-mail: archiryu@kumoh.ac.kr

ABSTRACT

The objective of this study is to develop prediction model of indoor carbon dioxide (CO₂) concentration using machine learning algorithm. Indoor CO₂ concentration is one of the indicators of indoor ventilation standard, and indoor air quality and ventilation performance can be checked through CO₂ concentration. The machine learning model is a method of analyzing the relationship between measured input/output data and does not require a high level of theoretical knowledge about the output value to be predicted, making it easy to develop a prediction model. In this study, a CO₂ prediction model was developed using an artificial neural network, a support vector machine, a random forest, and a K-nearest neighbor algorithm based on the existing HVAC system operation data. When comparing the performance of the developed CO₂ prediction model, the ANN model showed high performance. As a result of analyzing the time series data using the developed model, the measured indoor CO₂ concentration and the CO₂ concentration of the prediction model were similar, but on average, a relative error of less than about 5% occurred.

주요어 : 이산화탄소 농도, 예측모델, 기계학습

Keywords: Carbon dioxide concentration, Prediction model, Machine learning

서론

건물에서 사용되는 에너지 소비 절감을 위해 기밀성능이 향상됨에 따라 실내의 자연환기가 줄어들고, 기계적 환기에 많이 의존하고 있으며, 에너지의 많은 부분이 환기에 이용된다 (Wang et al., 2017). 실내 이산화탄소 농도는 실내 환기 기준의 지표 중 하나이며, 이산화탄소 농도를 통해 실내 공기질 및 환기성능을 확인 할 수 있다(Taylor, 2006; Cali et al., 2015).

OPEN ACCESS

Journal of KIAEBS 2020 December, 14(6): 699-706
<https://doi.org/10.22696/jkiaeb.20200059>

pISSN : 1976-6483
eISSN : 2586-0666

Received: October 28, 2020
Revised: November 20, 2020
Accepted: November 23, 2020

© 2020 Korean Institute of Architectural Sustainable Environment and Building Systems.



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

실내 이산화탄소 농도를 측정하기 위해서는 물리적 센서를 직접 설치하는 방법과 이론적 방법(수학적 모델, 기계학습 모델 등)을 이용한 예측모델 개발 방법이 있다.

이산화탄소 농도 측정 센서를 이용할 경우, 측정 장비를 위한 공간 확보 및 설치 비용이 필요하며, 측정센서의 정확도 확보를 위한 주기적인 교정 및 교체가 필요하다(Shin et al., 2016). 예측 모델 개발에는 수학적 모델을 이용한 방법과 기계학습 알고리즘을 이용한 모델 개발 방법이 있다. 수학적 모델은 고도의 전문지식과 경험이 요구되고, 예측하고자 하는 모델에 따라 많은 입력 변수를 요구한다(Ahn et al., 2016). 데이터 기반의 기계학습 모델은 측정된 입출력 데이터의 관계를 분석하는 방법으로 예측하고자 하는 출력값에 대한 고도의 이론적 지식이 필요하지 않아 예측 모델 개발에 용이하다(Shin et al., 2017). 따라서 본 연구에서는 기존의 공조 시스템의 운영 데이터를 바탕으로 기계학습 알고리즘을 이용하여 실내 이산화탄소 농도를 예측모델을 개발 및 비교 하였다.

기계학습 알고리즘

기계학습 알고리즘을 이용한 예측 모델 개발을 위하여 R 프로그램을 이용하였다. R은 통계, 데이터마이닝, 기계학습 등을 위한 언어로 주로 연구 및 산업 응용 프로그램으로 많이 사용되고 있다. 본 연구에서는 기계학습 알고리즘 중 인공신경망, 서포트벡터머신, 랜덤 포레스트, K-최근접이웃 알고리즘을 이용하여 예측모델을 개발하였다.

인공신경망(Artificial Neural Network, ANN)

인공신경망은 입력변수와 출력변수 사이 입력층, 은닉층, 출력층으로 구성된다(Figure 1(a)). 사용자는 은닉층 및 은닉 노드의 수를 결정할 수 있으며, 각 노드 사이의 가중치는 역전파 학습 알고리즘에 의해 결정된다. 본 연구에서는, 은닉층 수는 1개, 노드 수는 9개로 선정했으며, 활성화함수는 시그모이드 함수를 사용하였다.

서포트벡터머신(Support Vector Machine, SVM)

서포트벡터머신은 마진(margin)을 최대화하는 초평면을 찾는 방법이다. 초평면은 비선형의 훈련 데이터가 사상(mapping)된 n차원의 공간을 의미하며, 훈련 데이터는 초평면 상에서 선형 패턴으로 변환된다(Figure 1(b)). 본 연구에서는 최적화 문제를 선형 시스템으로 단순화시킨 Least Squares SVM을 사용하였다(Ra et al., 2017).

랜덤 포레스트(Random Forest, RF)

앙상블 방법(ensemble method) 중 하나인 랜덤 포레스트는 의사 결정나무들의 집합으로 구성된다. 각 의사결정나무는 원본 훈련 데이터에서 무작위 복원 추출된 샘플에 대해 학습되며, 각 노드에서는 무작위로 선택된 입력변수에 대해 최적의 노드 분할 기준으로 탐색한다(Figure 1(c)). 노드 분할 기준은 불순도 함수(impurity function)에 따라 계산된다. 본 연구에

서는 MSE (Mean Square Error)를 불순도함수로 선정하였으며, 의사결정나무의 개수는 500개, 무작위 선택 입력변수의 개수는 전체 입력변수 개수의 1/3으로 선정하였다(Ra et al., 2017).

K-최근접 이웃(K-Nearest Neighbor, KNN)

K-최근접 이웃 알고리즘은 현행상태벡터와 유사한 과거의 입력상태 벡터의 유사성을 이용하여 추출된 과거 상태의 자료 중 K개의 최근접 이웃을 선정하고, 선정된 이웃을 이용하여 미래의 상태를 예측하는 방법이다(Figure 1(d)). 따라서 K-최근접 이웃 알고리즘을 적용하기 위해서는 상태간 거리(Distance metric) 방정식이 정의되어야 하며, 본 연구에서는 거리방정식은 유클리디안 거리를 이용하였으며, k 근접 이웃수는 3으로 선정하였다.

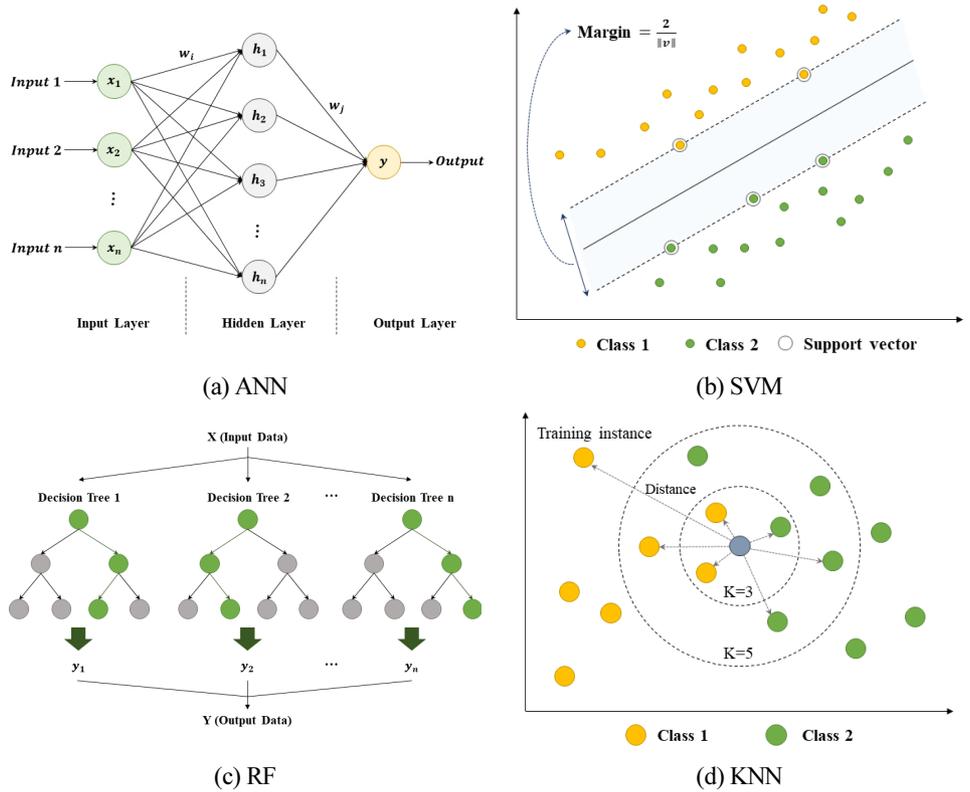


Figure 1. Schematic of machine learning

예측모델 개발 및 검증

대상건물

본 연구에서는 실측 데이터를 바탕으로 실내 이산화탄소 농도 모델을 개발하였다. 실측데이터 수집을 위하여 Y대학교의 HVAC (Heating, Ventilation & Air conditioning)시스템이 설치된 실험실을 대상공간으로 선정하였다. 대상공간에 설치된 HVAC 시스템은 변풍량 터

미널 유닛, 공조기, 공냉식 히트펌프, 펌프로 구성되어 있으며, Figure 2는 대상공간의 개략도를 나타낸다. 변풍량 터미널 유닛은 재열코일이 설치되어 있으며, 설계풍량은 1360CMH, 재열 코일 난방용량은 4000kcal이다. 공조기는 냉온수 코일과 공급,환기팬이 설치되어 있으며, VFD (Variable frequency drive)를 통한 팬 속도 제어가 가능하다. 또한 BAS (Building Automation system)을 통하여 모니터링 및 자동제어되며, Table 1은 대상 시스템의 사양, Figure 2는 대상 공간의 개략도를 나타낸다.

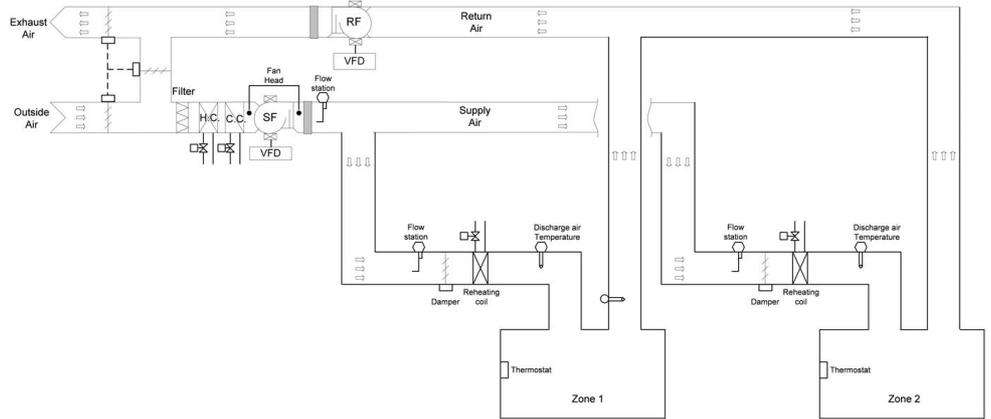


Figure 2. Schematic of test-bed

Table 1. Summary of test system

Category	Specification		
AHU	Air Flow (CMH)	12,000	
	Supply fan	Static Pressure (mmAq)	92
		Power (kW)	5.5
	Return fan	Air Flow (CMH)	9,600
		Static Pressure (mmAq)	35
VAV Terminal unit		Power (kW)	3.7
		Rated air flow rate (CMH)	1,360
		Capacity of reheating coil (kcal)	4,000

예측 모델 개발을 위한 데이터 입력변수 구성

기계학습을 이용한 예측 모델의 정확성은 입력 변수에 큰 영향을 받기 때문에(Jovanovic et al., 2015), 이산화탄소 농도 예측을 위해 적절한 입력 변수를 선정하여야 한다. 식 (1)은 실내 이산화탄소 농도를 예측하는 대표적인 이론식이므로써, 대상 공간의 체적, 실내 이산화탄소 발생량, 초기 이산화탄소 농도, 공급 이산화탄소 농도, 공급 풍량을 변수로 이산화탄소 농도를 산정한다(Costanzo et al., 2011).

$$C(t) = C_o + \frac{G}{Q} + (C(0) - C_o - \frac{G}{Q})e^{-It} \quad (1)$$

이 때, V : 대상공간의 부피,

$C(t)$: 시간 t 에 따른 실내 이산화탄소 농도,

Q : 실내 공급 풍량,

$C_o(t)$: 시간 t 에 따라 실내로 공급되는 이산화탄소 농도,

$G(t)$: 시간 t 에 따라 공간의 발생되는 이산화탄소 농도

$C_{(0)}$: 시간 $t=0$ 일 때 실내 이산화탄소 농도,

$I(= Q/V)$: 실내 환기율

본 연구에서는 수학적 모델을 바탕으로 BAS에서 수집되는 운영 데이터를 비교하여 기계 학습 입력 변수를 선정하였다. 입력변수로는 공급 풍량과 관계가 있는 공기조화기 공급 팬 스피드와 터미널 유닛 댐퍼 개도율, 공급 이산화탄소 농도와 관계가 있는 외기댐퍼 개도율, 실내 이산화탄소 발생량과 관계가 있는 재실인원수를 선정하였다.

데이터 수집

이산화탄소 농도 예측을 위해 공조시스템의 운영 데이터를 수집하였다. 다양한 운영 조건에 따라 이산화탄소의 변화를 예측하기 위하여 재실인원을 변화하며 데이터를 수집하였다. 재실인원은 10일간 1명씩, 최대 5명까지 재실인원을 변화시키며 데이터를 수집하였다. 수집 데이터는 변풍량 터미널 유닛 제어 및 모니터링에 사용되는 5개의 지점(공급팬 스피드, 외기 댐퍼 개도율, 터미널 유닛 댐퍼 개도율, 재실인원, 실내 이산화탄소 농도)을 선정하였다. 실내 이산화탄소 농도의 경우 이산화탄소 측정 장치를 환기 덕트에 설치하여 실내 이산화탄소 농도를 측정하였다. 그 외 3개의 데이터는 BAS를 통해 실시간 데이터를 수집하였고, 수집된 데이터는 총 3678 set이다. 수집된 데이터는 R 프로그램의 createDataPartition 함수를 이용하여 임의 분할하였으며, 85% (2719 set)는 학습용 데이터로 사용되었고, 학습용 데이터로 활용되지 않은 15% (959 set)는 검증용 데이터로 사용하였다. Table 2는 데이터 수집에 사용된 계측 센서의 사양 및 수집된 데이터의 범위를 나타내고 있다.

Table 2. Summary of Measured data and equipment

Measurement point	Measuring equipment	Accuracy	Full scale	Measurement data range
Fan speed	VFD	0.1 Hz	60 Hz	15~40 Hz
OA damper position	Damper actuator	1%	100%	10~90%
VAV terminal unit damper position	Damper actuator	1%	100%	6~100%
CO ₂ concentration	CO ₂ sensor	1 ppm	5000 ppm	414~1873 ppm
Occupancy	-	-	-	1~5 person

예측 모델 검증

이산화탄소 농도 예측 모델을 평가하기 위해 CvRMSE (Coefficient of Variation of Root Mean Square Error)와 MBE (Mean Bias Error)를 이용하여 정확도를 평가하였다. MBE는 예측값의 총량적인 오차를 의미하며, CvRMSE는 분산 정도를 통해 오차를 분석하는 방법이다. Table 2에서 ASHRAE Guideline 14에서 제공하는 기준을 사용하여 다음 식 (2), 식 (4)와 같이 CvRMSE (Coefficient of Variation of Root Mean Square Error)와 MBE (Mean Bias Error)를 이용하여 예측모델의 정확도를 평가하였다(ASHRAE, 2014).

$$MBE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \times 100 \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (3)$$

$$C_vRMSE = \frac{RMSE}{\bar{y}} \times 100 \quad (4)$$

예측모델 결과 분석

기계학습 모델을 바탕으로 실내 이산화탄소 예측 기계학습 모델을 개발하였다. 학습 데이터를 이용하여 개발된 예측 모델을 검증용 데이터를 이용하여 MBE, CvRMSE를 비교한 결과는 Table 3과 같다. 각 기계학습 알고리즘별 평균 CvRMSE는 KNN (8.1%) → RF (6.9%) → SVM (6.0%) → ANN (5.4%)로서 ANN 모델의 성능이 전반적으로 우수함을 알 수 있었으며, 남은 세가지 기계학습 모델(SVM, RF, KNN) 모두 우수하게 실내 이산화탄소 농도를 예측하는 것을 확인 할 수 있다. 또한, Figure 3은 검증 데이터를 바탕으로 제시한 모델의 유효성을 검토하기 위하여 실측을 통해 얻은 실내 이산화탄소 농도와 예측모델을 통해 얻은 예측값을 비교하였다. 비교 결과, 실측한 실내 이산화탄소 농도와 예측모델의 이산화탄소 농도가 유사하였으나 평균적으로 약 5% 이내의 상대오차가 발생하였다.

Table 3. Summary of Measured data and equipment

Machine learning model	MBE (%)	CvRMSE (%)	R ²
ANN	-3.2	5.4	0.97
SVM	-3.9	6.0	0.96
RF	-4.7	6.9	0.95
KNN	-5.1	8.1	0.93

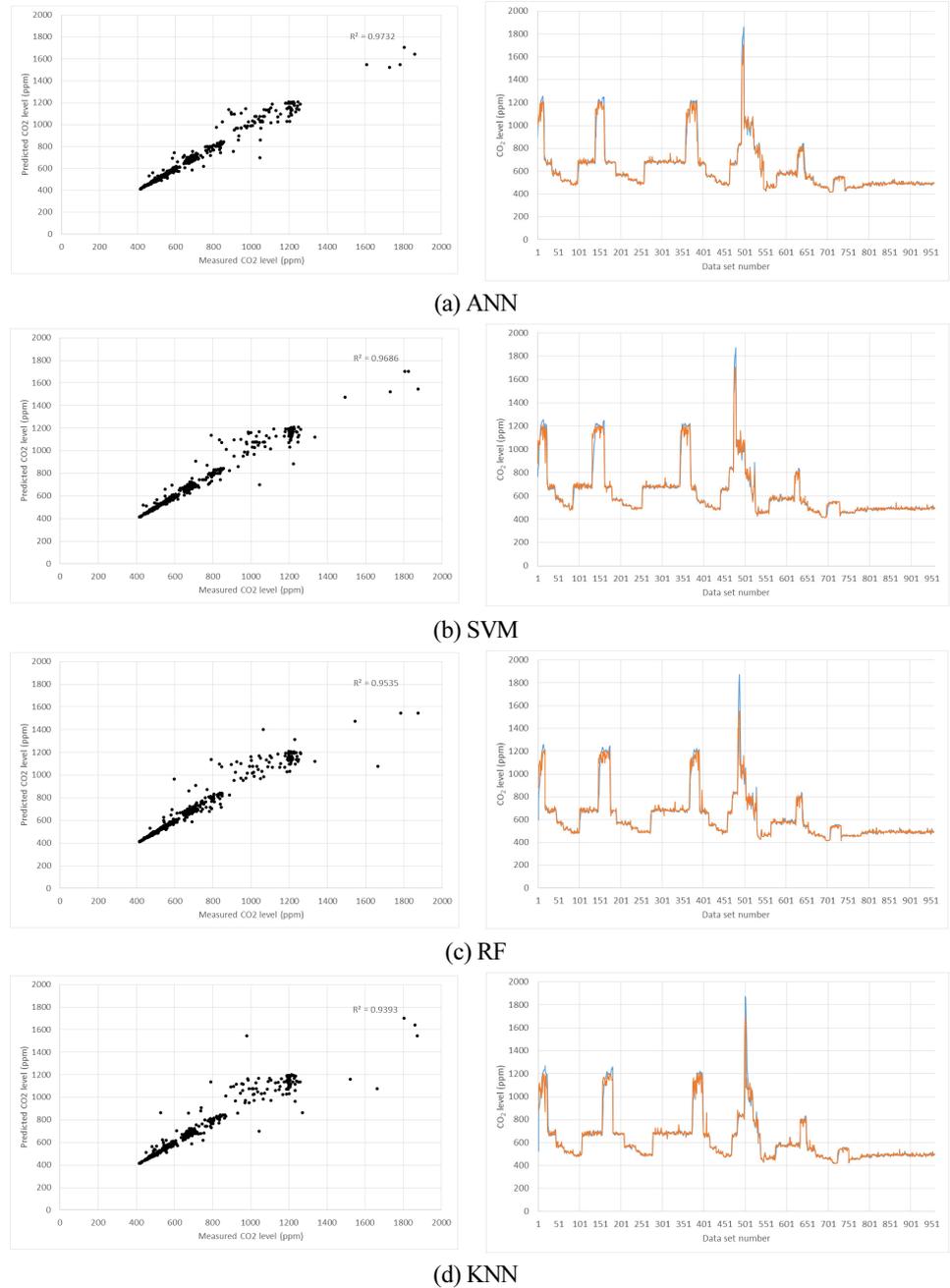


Figure 3. Comparison of measured and predicted CO₂ concentration

결론

본 연구에서는 실내 정보와 기존의 공조 시스템의 정보를 바탕으로 기계학습 알고리즘을 이용하여 실내 이산화탄소 농도를 예측모델을 개발하였다. 예측에 이용된 기계학습 방법은 인공신경망, 랜덤 포레스트, 서포트 벡터 머신, K-최근접 이웃을 이용하여 예측 모델 개발 및 검증을 하였다. 각 기계학습 알고리즘별 평균 CvRMSE는 KNN (8.1%) → RF (6.9%) → SVM (6.0%) → ANN (5.4%)로서 ANN 모델이 전반적으로 우수함을 알 수 있다. 도출한 모

델을 이용하여 시계열 데이터를 분석한 결과 실측한 실내 이산화탄소 농도와 예측모델의 이산화탄소 농도가 유사하였으나 평균적으로 약 5% 이내의 상대오차가 발생하였다.

본 연구를 통해 실내 공기질의 유지·관리를 위해 기계학습 기반의 실내 이산화탄소 농도 예측 모델을 개발 하였고, 향후 연구에서는 실제 건물을 대상으로 개발된 예측 모델 기반 HVAC 시스템 자동제어 및 고장감지진단 등의 활용방안에 관한 연구를 진행할 것이다.

후기

본 연구는 금오공과대학교학술연구비(2018-104-119)로 지원되었습니다.

References

1. Ahn, K.U., Kim, D.W., Kim, Y.J., Yoon, S.H., Park, C.S. (2016). Issues to Be Solved for Energy Simulation of An Existing Office Building. *Sustainability*, 8(4), 345.
2. Cali, D., Matthes, P., Huchtemann, K., Streblov, R., Müller, D. (2015). CO₂ based occupancy detection algorithm: Experimental analysis and validation for office and residential buildings. *Building and Environment*, 86, 39-49.
3. Costanzo, S., Cusumano A., Giaconia, C. (2011). Ventilation rates and unsatisfied percentage from indoor CO₂ concentration. *Indoor Built Environment*, 20(2), 232-245.
4. Jovanović, R.Ž., Sretenović, A.A., Živković, B.D. (2015). Ensemble of various neural networks for prediction of heating energy consumption. *Energy and Buildings*, 94, 189-199.
5. Ra, S.J., Shin, H.S., Suh, W.J., Chu, H.G., Park, C.S. (2017). Five Machine Learning Models for HVAC Systems in an Existing Office Building. *Journal of The Architectural Institute of Korea Structure & Construction*, 33(10), 69-77.
6. Shin, H.S., Suh, W.J., Chu, H.G., Ra, S.J., Park, C.S. (2017). Machine Learning Models for Prediction and Control of an Ice Thermal Storage System in an Existing Building. *Journal of The Architectural Institute of Korea Structure & Construction*, 33(11), 63-70.
7. Shin, K.J., Kim, H.J., Cho, Y.H. (2016). A study of the Prediction Model of Indoor Carbon Dioxide Concentrations. *Journal of Korean Institute of Architectural Sustainable Environment and Building Systems*, 10(1), 83-89.
8. Taylor, S.T. (2006). CO₂-based DCV using 62.1-2004. *ASHRAE Journal*, 48(5), 67-75.
9. Wang, Y., Kuckelkorn, J., Zhao, F.Y., Spliethoff, H., Lang, W. (2017). A state of art of review on interactions between energy performance and indoor environment quality in Passive House buildings. *Renewable and Sustainable Energy Reviews*, 72, 1303-1319.
10. ASHRAE, Guideline 14-2014. (2014). Measurement of energy, demand, and water savings, American Society of Heating, Refrigerating, and Air Conditioning Engineers, Atlanta, GA.