www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

MACHINE LEARNING-BASED ENERGY USE PREDICTION FOR THE SMART BUILDING ENERGY MANAGEMENT SYSTEM

SUBMITTED: April 2023
REVISED: June 2023
PUBLISHED: September 2023
EDITORS: Chansik Park, Nashwan Dawood, Farzad Pour Rahimian, Akeem Pedro
DOI: 10.36680/j.itcon.2023.033

Mustika Sari, Ph.D.
Center for Sustainable Infrastructure Development, Universitas Indonesia, Indonesia;
mustika.sari01@ui.ac.id

Mohammed Ali Berawi, Professor
Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia, Indonesia;
maberawi@eng.ui.ac.id

Teuku Yuri Zagloel, Professor
Department of Industrial Engineering, Faculty of Engineering, Universitas Indonesia, Indonesia;
yuri@ie.ui.ac.id

Nunik Madyaningarum, Ms.
National Research and Innovation Agency (BRIN), Indonesia;
nuni005@brin.go.id

Perdana Miraj, Ph.D. Student
Center for Sustainable Infrastructure Development, Universitas Indonesia, Indonesia;
perdana.miraj@ui.ac.id

Ardiansyah Ramadhan Pranoto, Mr.
Center for Sustainable Infrastructure Development, Universitas Indonesia, Indonesia;
ardiansyahpranoto@gmail.com

Bambang Susantono, Ph.D.
Center for Sustainable Infrastructure Development, Universitas Indonesia, Indonesia;
bsusantono@gmail.com

Roy Woodhead, Ph.D.
Sheffield Business School, Sheffield Hallam University, United Kingdom;
R.M.Woodhead@shu.ac.uk

SUMMARY: *Smart building is a building development approach utilizing digital and communication technology to improve occupants' comfort inside the building and help increase energy usage efficiency in building operations. Despite its benefits, the smart building concept is still slowly adopted, particularly in developing countries. The advancement of computational techniques such as machine learning (ML) has helped building owners simulate and optimize various building performances in the building design process more accurately. Therefore, this study aims to assist energy efficiency design strategies in a building by identifying the features of the smart building characteristics that can potentially foster building energy efficiency. Furthermore, an ML model based on the features identified is then developed to predict the level of energy use. K-Nearest Neighbor (k-NN) algorithm is employed to develop the model with the openly accessible smart building energy usage datasets from Chulalongkorn University Building Energy Management System (CU-BEMS) as the training and testing datasets. The validation result shows that the predictive model has an average relative error value of 17.76%. The energy efficiency levels obtained from applying identified features range from 34.5% to 45.3%, depending on the reviewed floor. This paper also proposed the dashboard interface design for ML-based smart building energy management.*

KEYWORDS: *smart building, energy management system, energy use prediction, machine learning*

REFERENCE: *Mustika Sari, Mohammed Ali Berawi, Teuku Yuri Zagloel, Nunik Madyaningarum, Perdana Miraj, Ardiansyah Ramadhan Pranoto, Bambang Susantono, Roy Woodhead (2023). Machine learning-based energy use prediction for the smart building energy management system. Journal of Information Technology in Construction (ITcon), Special issue: 'The future of construction in the context of digital transformation (CONVR 2022)', Vol. 28, pg. 622-645, DOI: 10.36680/j.itcon.2023.033*

1. INTRODUCTION

One of the issues faced by the building sector in various countries is the number of energy consumption that continued to increase drastically over the past decade (Yas and Jaafer, 2020). Countries in Asia are no exception, as more than half of the world's new buildings are built yearly, increasing building energy use accounting for approximately 25% of Asia's total energy consumption (USAID, 2018). The rise in building energy use is attributed to several factors, including rapid urbanization (Cao et al., 2016), improved living standards (Kadoshin et al., 2000), and global climate change (Nematchoua et al., 2015).

Practitioners, professionals, and academics in the construction and building industries defined alternative approaches to anticipating the growth of building energy demand, one of which is by implementing the smart building concept (Parisi et al., 2021). A smart building is perceived as a building development equipped with an automated process that monitors and controls its operation by utilizing technological advances as instrumentation measures (Qolomany et al., 2019). The primary goal of smart buildings is to solve problems of sustainability and intelligence in buildings by developing an integrated management system, resulting in the optimal combination of building energy consumption and human comfort (Benavente-Peces, 2019; Wang et al., 2012).

The automation system in smart buildings can result in more energy-efficient building design, construction, and operation (Wang et al., 2012). With this automation, the building management system can collect information about the occupants' behaviors to provide intelligent data related to energy use. These data are then analyzed further in the smart building energy management system (SBEMS), a crucial feature of a smart building that can help the decision-making processes to increase energy efficiency. (Buckman et al., 2014).

Despite its benefits in building energy efficiency and occupant comfort, the literature indicates that the implementation of smart buildings in developing countries is not optimal. Some of the most reported barriers to the adoption of smart buildings to include the high cost of initial construction, the knowledge gap between the users' perception and the actual impact of smart buildings, the resistance to change from the use of traditional technologies, and the lack of a framework representing the features, design guidelines, and parameters of smart buildings (El-Motasem et al., 2021; Kadoshin et al., 2000; Omar, 2018).

Advanced technologies have helped improve the construction industry's value creation and decision-making processes (Hossain and Nadeem, 2019). Relatedly, the performance of SBEMS can also be improved by using artificial intelligence (AI) applications, such as machine learning (ML) techniques. Recent studies on the energy efficiency in smart buildings conducted these past few years show that ML has been widely used to predict the building energy demand and optimize energy usage (Panchalingam and Chan, 2021). According to Djenouri et al. (2019), studies focused on ML-based energy-related solutions for smart buildings fall into three categories: energy demand estimation, equipment profiling and failure detection, and sensor decisions.

As previously stated, the slow adoption of smart buildings is due to a lack of experience and knowledge about technologies, features, components, and energy efficiency strategies. Moreover, according to Amasyali and El-Gohary (2018), there was a relative lack of research on long-term building energy consumption prediction attributed to the data insufficiency and the complexity of occupant energy use behavior in the building. Therefore, the objectives of this study are twofold. First, this paper investigates smart building features that can aid in building energy efficiency, addressing the slow adoption of smart buildings by closing the knowledge gap existed regarding the benefits that smart buildings yield. Subsequently, an ML model is developed to obtain data-driven energy use predictions that will be used as the basis for decision-making for the energy efficiency strategies in smart buildings, utilizing publicly available data on smart building energy use. Nevertheless, this research also develops the interface design of the ML-based SBEMS dashboard, which represents a dependable and user-friendly instrument improving the decision-making process in energy management.

This study is expected to provide a new approach to foster the adoption of smart building concepts through the development of a user-friendly smart building energy management that can predict building energy use using data analytics and ML model. The ML model was developed using open-source energy datasets from the Building Energy Management System in Thailand designed by Smart Grid Research Unit (SGRU). The K-Nearest Neighbor (k-NN) algorithm was used in the model development due to its eminence as a prediction technique that aids in building energy management efficiency by minimizing energy consumption (Mazlan et al., 2020; Shapi et al., 2021; Wahid and Kim, 2016). Furthermore, k-NN algorithm is also renown for simplicity, versatility, adaptability to complex data, as well as robustness against outliers, making it the most suitable method for ML predictive model development, particularly with the energy datasets used in this study.

There are three sections to the paper. The theoretical framework first presents pertinent aspects of smart building, energy management systems, and machine learning. The subsequent section describes the research methods, including dataset analysis, data preprocessing, model development, and dashboard design development. Afterward, the findings are presented and discussed in the results and discussion section. Furthermore, the study's limitations are discussed, and recommendations for future research are provided. The final section concludes with an assessment of all previous sections.

2. LITERATURE REVIEW

2.1 Smart Building

A smart building can be defined by its ability to provide judgment, problem-solving, operational speed, creativity, general knowledge, and encouragement (Ghaffarianhoseini et al., 2016). It acknowledges and reflects the technological advances and convergence of building systems, the shared aspects of the systems, and the added functionality provided by interconnected systems (Sinopoli, 2010). In a smart building, intelligence and control as a fully automated building system are incorporated by utilizing advanced information and communication technology (ICT) as an integrated approach to generate a more efficient building design, construction, and operation (Berawi et al., 2017). With the constant advancement of technology comes a new chance to improve people's lives through pervasive computing technology in smart buildings that facilitate enhanced communications, awareness, and functionality. (Edwards and Grinter, 2001).

ICT is regarded as a critical intelligence aspect in smart buildings to achieve various objectives, such as providing smart and comfortable services to building occupants, monitoring user safety within the building, managing the building behavior to reduce energy consumption, and ensuring the health and durability of building's equipment (Abdennadher et al., 2016). In addition, the social aspect of smart buildings also takes into account the occupants' quality of life and comfort.

Though the notion of a smart building is being developed at all stages of a building's life cycle, the operation phase of a smart building is known as the most advanced automated process that enables building services to interact with one another. Its automatic devices can regulate the functioning of multiple systems, including heating, air conditioning, ventilation, lighting, security, water, waste, etc., hence minimizing energy consumption, enhancing safety and security, and improving the comfort of the residents inside the building (Apanaviciene et al., 2020).

To obtain its objectives of ensuring building efficiencies and delivering occupants' comforts, the elements making up smart buildings involve hardware, software, and network (Batov, 2015). Hardware such as sensors and meters are used in smart buildings to record and monitor conditions in a building and the surrounding (Grindvoll et al., 2012). The building systems can then determine the necessary follow-up actions for a variety of building aspects, such as light intensity, room temperature, and gas or water leaks. On the other hand, the software is used to receive the data provided by the hardware, learn from it, and process and make decisions that affect the building's performance and the occupants. The network that connects these two elements must incorporate them as an integrated smart building energy management system, in which decisions can be made to optimize building energy efficiency and occupants' comfort by collecting and analyzing data on occupants' behaviors using BEMS.

2.2 Smart Building Energy Management System

Smart building energy management systems (SBEMS) are used to observe energy use patterns and determine the potential savings in building energy use while maintaining occupant comfort. It changes a conventional building into an energy-aware environment that enables automated real-time building management and operations for controlling various aspects of smart buildings, providing greater energy savings on building appliances and enhancing indoor comfort (Aliero et al., 2022). The general scheme of an SBEMS is simply a predictive and optimization series of inputs from the external data (e.g., weather), sensors (e.g., temperature, lighting, and energy usage levels), and information from occupants inside via the human-machine interface (HMI). The SBEMS algorithms will process these parameters, resulting in outputs as commands for the actuator to recognize the conditions of both the systems and the occupants.

Latest developments around the BEMS have focused on providing occupants with an interface to monitor, schedule, and modify building energy consumption profiles and enabling utilities to participate in a communication grid via demand response control and automatic self-report function (Aliero et al., 2022; Eini et al., 2021). Research related to BEMS can be classified into two complementary categories: predictive energy management and real-time direct control (Missaoui et al., 2014). In the first category, control is executed by applying a predictive model to measured data to forecast the most optimal control strategy that can be performed. In the second category, predictive control is employed alongside a real-time control that offers more accuracy. Other than smart metering, additional solutions for BEMS include statistical models, cloud and fog computing, and several approaches, such as Big Data (BD), data science (SC), Internet of Things (IoT), and other areas (Mir et al., 2021).

Previous studies have investigated the application of data science techniques, such as regression, clustering, and sequence exploration, to energy management in buildings for energy prediction, failure prevention, load balancing, and fault detection (Molina-Solana et al., 2017). Fan et al. (2021) argued that utilizing advanced data analytics methods, such as advanced data mining and machine learning algorithms, provides significant technical support for big data analytics for enhancing building performance modeling and analysis. Furthermore, another study showed numerous data science techniques for predicting building energy, such as the grey-and-white box approaches based on the physics law, time-series solutions, and ML techniques (Bourdeau et al., 2019).

2.3 Machine Learning

Various research on ML development has been extensively conducted in the building sector, focusing on energy, building performance, and occupant comfort (D'Amico et al., 2020; Djenouri et al., 2019). In the aspect of energy, Ascione et al. (2017) built an ML model for predicting energy performance and the energy retrofitting scenario for building elements. The study stated that the conventional tools for simulating building performance might be replaced with the ML method due to its high reliability and low computational times. Moreover, Singh et al. (2020) have developed energy predictions utilizing four schematic predictive approaches, one of which is a long short-term memory (LSTM)-based ML method that claims to have a faster computation time than the EnergyPlus simulation. On the other hand, Westermann et al. (2020) used the convolutional neural network (CNN) algorithm to produce an ML model to displace a building energy modeling tool that can accurately predict thermal energy with varied building designs and climate variations as inputs.

Similarly, a study by Seyedzadeh et al. (2020) concluded that a gradient-boosted regression tree (GBRT) based ML model predicting energy performance could be a robust tool to explore the expanding solution space that supports decision-making in energy retrofit scenarios. The recently published research by Liu et al. (2022) proposed a multi-step predictive deep reinforcement learning algorithm based on an LSTM neural network with a generalized cross entropy (GC-LSTM) for smart building' HVAC control systems to reduce power consumption costs while maintaining user comfort. The algorithmic framework analyzes the current inside and the expected outdoor ambient temperature to improve user comfort. Furthermore, the suggested temperature prediction model was paired with the Deep Deterministic Policy Gradient (DDPG) reinforcement learning algorithm to flexibly modify the HVAC system's output power in response to dynamically fluctuating electricity costs. It claimed that power consumption costs could be reduced by up to 12.79% compared to other approaches.

A deep understanding of the energy consumption pattern in a smart building is crucial in improving energy efficiency. With the vast data and sensors available, data-driven smart technologies have beaten the performance of conventional energy simulation techniques. Along with various methods used to analyze building energy consumption, ML has been proven to be one of the reliable techniques used to plan and estimate the imminent improvement of smart building energy efficiency. Consequently, smart models are required in the smart building best practice for the industry to bring convenience and effectiveness to building management teams performing energy efficiency strategies.

Other than some algorithms mentioned above, there are several different ML methods utilized to develop a building energy prediction, such as support vector machine (SVM), decision tree (DT), k-nearest neighbor (k-NN), and many more (Deng et al., 2018; Kontokosta and Tull, 2017; Wahid and Kim, 2016). The k-NN algorithm is known as one of the prediction methods considered to be a simple classification technique used for prediction. In predicting building energy, kNN is a non-parametric and straightforward learning algorithm that uses a database to classify new data (Mazlan et al., 2020).

Based on the literature discussed above, this study believes that the adoption of smart buildings can be increased through the implementation of an ML-based SBEMS that bridges the existing knowledge gap by effectively showcasing the multitude of benefits that smart buildings offer. By leveraging ML algorithms and data-driven decision-making, the proposed SBEMS is expected to enhance operational efficiency and reduces overall environmental impact. Through the integration of advanced technologies and data-driven decision-making, the proposed system empowers building owners, operators, and stakeholders to make informed choices, leading to greater acceptance and widespread adoption of smart building solutions.

3. METHODS

This research was conducted in two stages to answer the three research objectives (RO), which include: (RO1) identifying the energy efficiency efforts gained through varied energy components in smart buildings; (RO2) developing an ML model for the energy use prediction of the energy-efficient smart buildings; (RO3) designing the dashboard interface for the ML-based SBEMS. To achieve the first objective, the energy efficiency features in smart buildings were identified by conducting a research desk study to analysis related archive and literature. Following that, expert validations were done to determine the relevancy and feasibility of the energy-efficient smart building components based on the experience and knowledge of the experts. As for the second objective, a series of activities included data collection, preprocessing, ML model development, and model validation. Consequently, the third objective was obtained by conducting a benchmark study of previously developed ML-based BEMS and in-depth interviews with experts and practitioners in the smart building operations field. The flowchart of the research method is shown in Fig. 1. The detailed research methods for RQ2 are described in the following subsections below.

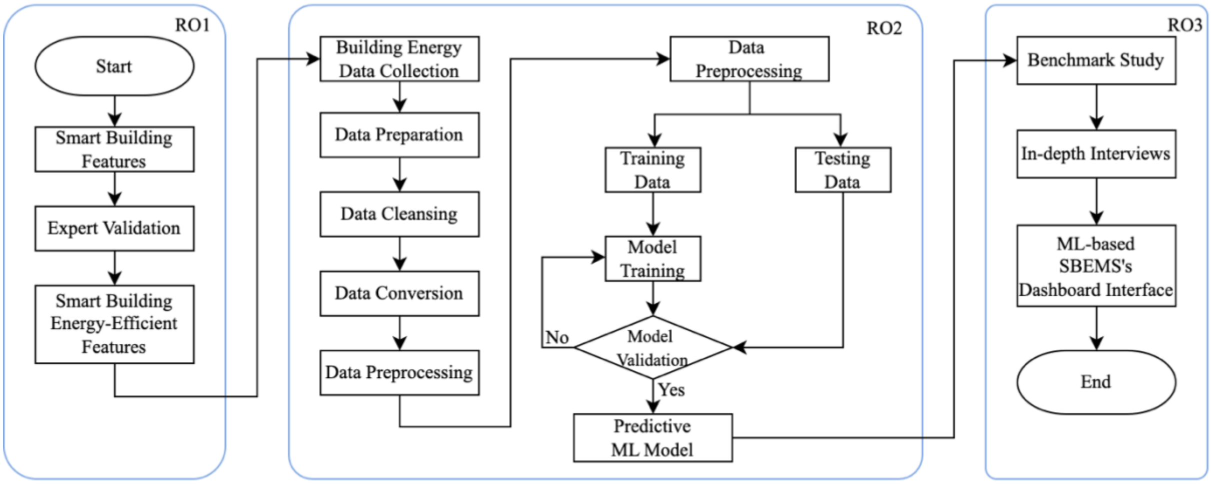

FIG. 1: Research workflow

3.1 Dataset

The object of research data utilized in this study is the publicly available building data of the Chamchuri 5 building, a seven-story office with smart building concept in the Bangkok area, Thailand, with a total floor area of about 11,700 square meters and 33 zones for energy data collection (Pipattanasomporn et al., 2020). The datasets recorded at one-minute intervals from July 1, 2018, to December 31, 2019, include energy level readings for lighting (Li), plug (Pl), and cooling components/air conditioner (AC), as well as sensor readings in the form of lighting level (lux), relative humidity (%), and temperature (°C) for each zone on each floor of the building. Nevertheless, this study only used the data from January 1, 2019, to December 31, 2019, to reflect one year of data collection. Table 1 displays a portion of the collected building data recording.

This data was then descriptively examined to identify the data structure and assist in determining the data processing strategies; as a result, data preparation activities can be optimized. These are the outcomes of the descriptive analysis of the data:

1. The average value of the room temperature varies from 23.2 °C to 27.3 °C, with a minimum value in the range of 17.9°C to 24.3°C.

2. The average relative humidity value is between 33.1% and 45.5%, with a maximum value between 76.9% and 89.9%.
3. The value of the illumination level for each zone is quite varied, with the maximum value ranging from 16 lux to 92 lux.
4. In the same zone, the components of energy consumption differ significantly from one another. Thus, zones with several AC units are dominated by high energy-consuming ones.

TABLE 1: Collected Building Energy Dataset (Source: http://www.sgru.eng.chula.ac.th/cu-bems/)

Date	z1_Li (kW)	z1_Pl (kW)	z2_AC1 (kW)	z2_AC2 (kW)	z2_AC3 (kW)	z2_AC4 (kW)	z2_Li (kW)	z2_Pl (kW)	z3_Li (kW)	z3_Pl (kW)	z4_Li (kW)
2019-01-01 00:00:00	0.03	0.58	2.31	21.15	0.01	0.02	17.37	15.2	10.01	0.38	14.58
2019-01-01 00:01:00	0.0	0.58	2.31	35.07	0.0	0.02	17.34	19.16	9.98	0.37	14.57
2019-01-01 00:02:00	0.02	0.58	30.96	34.37	0.01	0.03	17.31	19.02	9.98	0.38	14.62
2019-01-01 00:03:00	0.0	0.57	51.32	18.91	0.01	0.01	17.39	18.85	10.01	0.37	14.6
2019-01-01 00:04:00	0.01	0.56	48.87	1.35	0.01	0.01	17.48	18.57	10.05	0.38	14.6

3.2 Data Preparation

ML model training demands vast volumes of data, and the quantity of data availability determines how to handle missing data and data processing procedures later in the modeling process. Therefore, examining the available data in the initial stage is crucial. In Jupyter Notebook, the original CSV-formatted datasets were then examined for missing data using the module command in Pandas, an open-source data analysis tool. Table 2 displays the proportion of missing data determined based on the number of observation columns on each floor. The investigation of the missing data revealed that the degree of data availability is enough training material for the ML model, with an overall loss rate of less than 15%, with the total amount of available observed data remaining rather high. Moreover, the location of the missing data will be examined for further review based on the recording date and the missing components. This examination will inform the selection of components and time utilized as a model trainer on each floor, as well as a processing method and model preparation applicable to all floors.

TABLE 2: Total missing data from each floor dataset

Floor	Total Missing Data	Total Expected Data	Total Available Data	Missing Data (%)
1st Floor	7,865	6,307,200	6,299,335	0.12%
2nd Floor	1,352,388	19,447,200	18,094,812	6.95%
3rd Floor	1,688,708	15,768,000	14,079,292	10.71%
4th Floor	2,000,069	15,768,000	13,767,931	12.68%
5th Floor	2,096,193	15,768,000	13,671,807	13.29%
6th Floor	1,318,076	15,768,000	14,449,924	8.36%
7th Floor	1,455,410	15,768,000	14,312,590	9.23%

The data generates a DataFrame with columns providing information on the building's zones, components, and sensors. The information was then arranged with a dictionary containing a column of energy use in units of kW and sensor readings with information such as lux (lighting), degC (temperature), and RH% (relative humidity). In addition, there was also a column under 'Date', that carried textual information on the observation time. Fig. 2 illustrates the data structure for each CSV file, where each floor consists of several zones with their energy meters and sensor reading components. The energy measurement components were then categorized into AC readings, lighting, plugs, and electronic devices. Meanwhile, the sensor captured the room temperature, humidity, and illumination levels in the zone concerned.

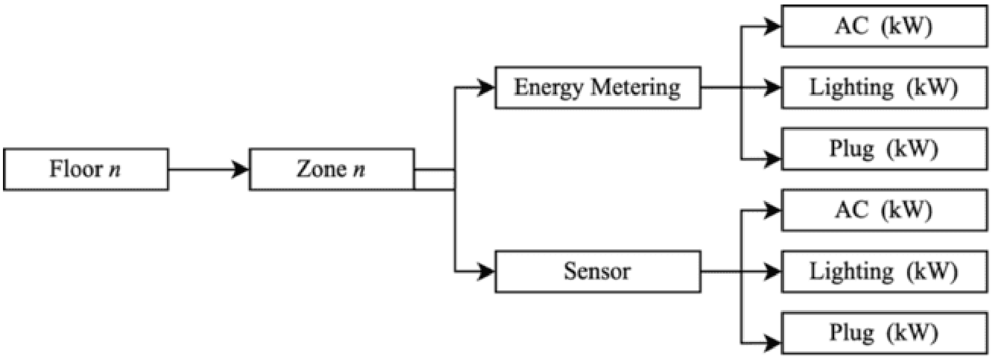

FIG. 2: The data structure for each floor and each zone

During the data analysis and modeling phases, automating the Jupyter Notebook data was necessary for quick and structured access. The code in the Jupyter Notebook was prepared by splitting the reading columns in the reading data file, which transformed the previously built data form based on the recording column in the CSV file into a dictionary form that can be retrieved depending on usage requirements. This method was conducted by defining a column divider function in the floor data files, a dictionary consisting of zones on the floor, with sub-dictionary recording components for each zone, including AC, lighting, and plugs.

3.3 Data Cleansing

Unfitting data, such as incomplete recording data and numbers much beyond the range value, should be handled using a data cleansing procedure. This process is vital since complete logging data with missing information on recording time will be unusable and only be confounding when given input to ML algorithms. The initial cleaning procedure was performed on data whose recording time did not conform to the standard recording format.

3.4 Data Conversion

Data conversion is the process of converting the time data format, which was previously a string type, into a DateTime type. This process aims to enable various data processing with the DateTime format, including incorporating the average hourly value, retrieval of hour information for model training needs, and other processing carried out in model development.

The next step is to change the 'Date' column to be an index on the DataFrame to facilitate the calculation and data slicing processes. This step replaced the index component previously marked by a number string from 0 to the addition of 1 according to the amount of data into an index containing the datetime64 data type, which describes the time and day of data recording.

The previously described data has a recording interval of one minute. However, the data needs to be used in units of hours for energy use analysis to determine an average value for each hour of observation. Therefore, the energy use data in units of a kilowatt-hour (kWh) with no significant statistical change in data distribution before and after the change (see Table 3). After obtaining energy data in kWh units, the missing data is cleaned again if there is no observation at a particular hour. The data totaling 8,760 hours of observation will be cleaned to generate the final data for the next stage.

TABLE 3: Total data before and after the data conversion

Floor	Initial Available Data	Cleaned and Conversed Data
1st Floor	6,299,335	8,754
2nd Floor	18,094,812	6,871
3rd Floor	14,079,292	5,516
4th Floor	13,767,931	4,255
5th Floor	13,671,807	4,621
6th Floor	14,449,924	4,794
7th Floor	14,312,590	6,697

3.5 Data Preprocessing

The first step in data preprocessing is to ask for input in the form of zones reviewed along with parts to be trained and then accessed in an organized dictionary. The input is the numbers from the zone and the category text, either 'AC', 'Light', or 'Plug'.

The hour and day data originally stored in the 'Date' column of the index were subsequently retrieved and converted from string format to DateTime. This index's hour and day description features were presented in binary form, where a value of 1 indicates that the observation data is present at the relevant point. Fig. 3 depicts daily and weekly usage patterns of energy data in a binary format that can be read by the ML algorithm. For instance, the number 1 in the fourteenth column of the first row (index 0) shows that the energy record at index 0 was created at 2:00 p.m. The same applied to data on energy use days, which consisted of seven columns (Monday – Sunday).

	0	1	2	3	4	5	6	7	8	9	...	14	15	16	17	18	19	20	21	22	23
0	0	0	0	0	0	0	0	0	0	0	...	1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	...	0	0	1	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	1	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	1	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

FIG. 3: Binary format of the hours of energy usage

The energy usage data from January to November was utilized as training data for the creation of the ML model, while the data from December was used as testing data. However, due to an incomplete dataset, the sixth floor only utilized data from January to August. Meanwhile, September data was used as the testing data that measured the accuracy of predictions by the ML model for the sixth floor. For the division between the months used on the sixth floor and other floors, the IF-ELSE algorithm was employed to divide the various month components. The ratio between data training and testing data approaches 80 to 20%.

The last step in data preprocessing is the scaling process, adjusting the feature size to be on par with other compilers. It is a crucial process since this study used k-NN as the algorithm for the ML model development. In this study, the k-NN algorithm was chosen based on several factors, which include: (1) The type of algorithm based on the type of data that yields results for the provided features; (2) the relatively basic and easy-to-implement algorithm; (3) a sufficient quantity of data, and (4) research objectives do not concentrate on determining the significance of the given features but rather on the outcomes of processing algorithms.

In the k-NN algorithm, the user determines the number of k values, with the optimal k values varying between cases. Since this algorithm is distance-dependent, normalizing the training data can considerably increase the algorithm's accuracy if the scales of the features are vastly different (Piryonesi and El-Diraby, 2020). Using the following Equation 1, the style of distance calculation will be used to calculate the distance between points in this algorithm.

$$\left(\sum_{i=1}^{k} \left(\left| x_i - y_i \right| \right)^q \right)^{1/q} \quad \text{...(Equation 1)}$$

Where *k* is the dimensional components at point, *x* is the point 1, *y* is the point 2, and *p* is the distance matrix

Based on the similarity measure, the k-NN algorithm predicts a numerical target using input data (distance). In regression form, the algorithm's output is the object's value, which is the average of various values of its *k*-nearest neighbors. The number of *k* values of the algorithm can be determined by finding the case-specific optimal *k* value. Although this approach depends on distance, balancing the training data using feature scaling can dramatically improve the accuracy of the ML model if one feature utilizes a very different scale.

The k-NN regression process consists of the following steps:

1. Load the necessary data,

2. Determine the desired value of k,
3. Calculate the distance between the examples in the query and those in the data
4. Add the distance and index of the instances and collected data to the collection,
5. Sort by distance from the smallest to the greatest.
6. Choose the first k-number in the collection,
7. Returns the average of the chosen k integers.

The third through seventh stages are the algorithm's work phases, which typically do not necessitate special consideration. According to the preferences of the user, the first and second phases can be modified. However, loading data before passing it to an algorithm necessitates special treatment. It is necessary to examine the data to ensure that it is worthy of input in terms of completeness and distribution.

Afterward, it has to be determined whether the input features have significance for the model's output. Because if the feature does not correlate with the outcomes, using the feature in question will corrupt the algorithm's output. In addition, for algorithms that exploit the distance between data points, it is necessary to ensure that the values of the features fall within a relatively narrow range. If the data distribution is excessively large or substantially larger or smaller than the other features, feature normalization or scaling must be performed. Only then can optimal *k* values for the algorithm be sought.

Using the correct value of *k* to accomplish the objective of employing the most effective algorithm is a crucial aspect of this algorithm's application. The optimal use of *k* values depends on the data used; the higher the value of *k* reduces the influence of noise (inconsistent data) on classification but causes the boundaries between classes difficult to distinguish (Everitt et al., 2011). The selection of optimal k values will be based on the iteration of various *k* values derived from the employed data. The algorithm's accuracy will be evaluated at each iteration, and the k value will be derived from the experiment with the lowest error rate.

The ML model development used Python programming language and its open-source libraries, including Scikit-learn, Matplotlib, Datetime, Pandas, and Numpy. In addition, the Mean Absolute Percentage Error (MAPE) metric was employed to assess the ML model accuracy. Due to its scale independence and interpretability, it is one of the most widely used prediction accuracies for evaluating the performance of time series, demand forecasting, and other predictive models (Kim and Kim, 2016). It considers the prediction by comparing the actual (A) and predicted values of the desired output (P), as calculated using Equation 2.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{A_i - P_i}{P_i} \times 100 \qquad \text{……………………………………………….}(Equation\ 2)$$

Fig. 4 illustrates the data preparation, cleansing, conversion, and preprocessing workflow and the machine learning model development and evaluation processes.

3.6 Dashboard Design

European Commission (2010) stated that there are three important characteristics for an indicator of intelligence in buildings, which include the technological readiness of a building to (1) efficiently manage itself, (2) interact with its users, and (3) interact with the larger energy environment. Key components of a building energy management system include structure-wide and appliance-specific control. The user interface is a crucial component of the dashboard of SBEMS, allowing users to see, analyze, and control the building energy use (Moseley, 2017).

Innovative user interface design can be obtained by combining advanced design ideas with practical techniques. According to Hong (2010), the user interface design principles that are useful for developing new solutions include: (1) Examining promising alternatives from the widest range of possible alternatives to provide the best user experience through the integration of various features, including hardware, software, artistic, mathematical, and intuitive aspects; (2) Involve subject matter experts in all phases of the design; (3) Utilize object-oriented design concepts throughout the development process.

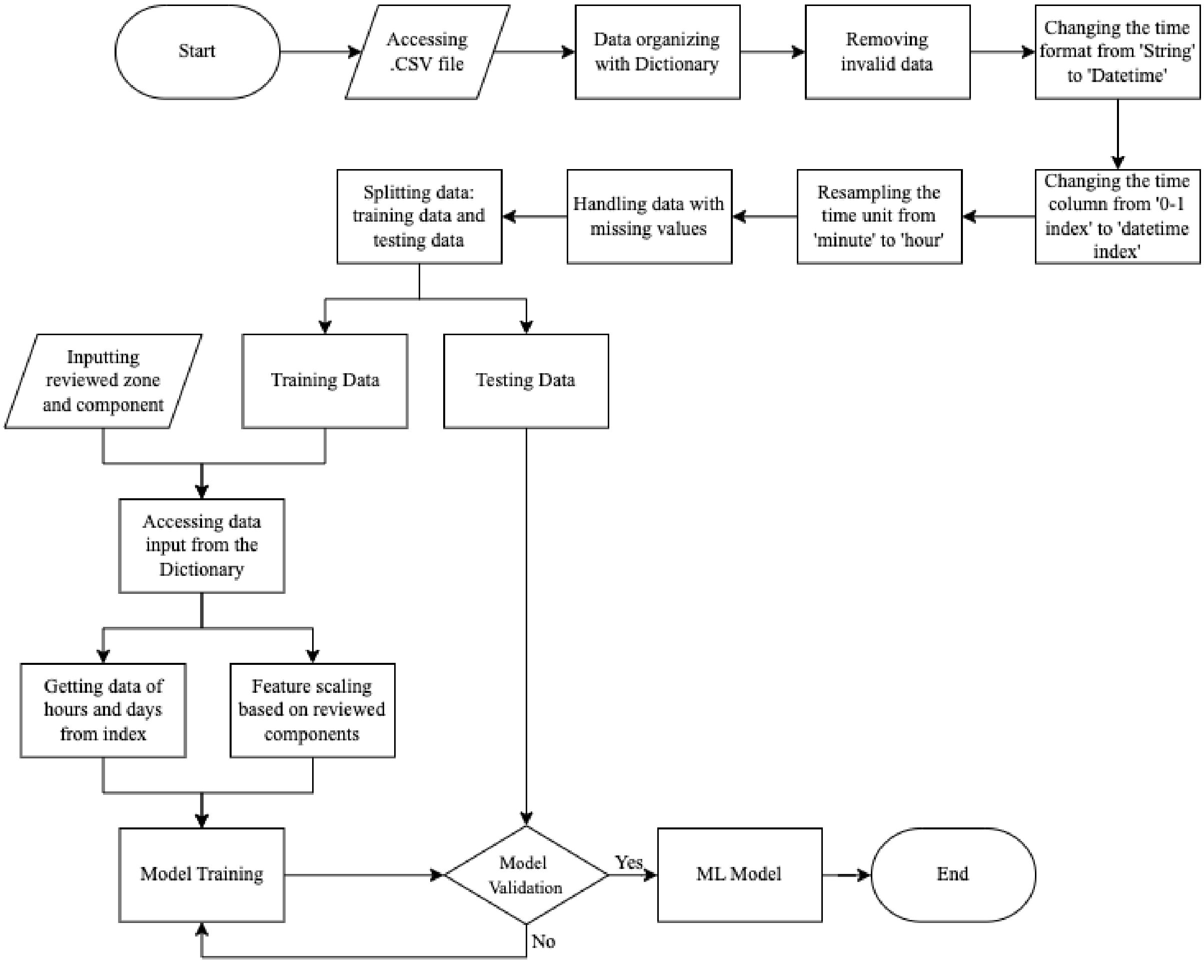

FIG. 4: Feature inputs for each energy component in ML model development

This research conducted a benchmark study on some existing dashboard development of SBEMS to identify the features and activities needed for the SBEMS before the design development of the dashboard's user interface. The initial UI design of the ML-based SBEMS dashboard would incorporate energy efficiency-related parameters for smart buildings. Afterward, a selection of floors and zones may be accessible for the energy consumption of each component, depending on the model algorithm developed from data processing. In addition, the choice of intervention type to was also carried out. Subsequently, in-depth interviews with experts from the field of smart building management and the development of user interface/user experience were performed to evaluate the developed UI design of the SBEMS dashboard. The expert validation was carried out based on established usability principles and guidelines.

4. RESULTS AND DISCUSSION

4.1 Energy Efficiency Strategies for Smart Buildings

A literature study was conducted to identify the energy efficiency efforts in the cooling, plug, and lighting components that can be implemented in the smart building concept. Archive analysis was performed on the studies that have experiments tested in obtaining a particular amount of efficiency in energy use. In addition, a review was also carried out to identify the supporting features in the smart building energy system that are currently used to improve building energy efficiencies from various sides, such as regulation, supervision, and control. Energy efficiency efforts in smart buildings are a collection of components of both management and technology that play a role in the energy system of a smart building, seeking to optimize the efficiency of energy use. According to expert interviews, this study grouped these features and components into AC, lighting, plug, and electronic devices.

Energy efficiency efforts in AC include inverter air conditioners, energy-friendly operating systems, dehumidification systems, efficient compressors, ventilation devices, and systems. Inverter air conditioners fundamentally differ from conventional AC systems in the compressor's speed which can be adjusted depending on the cooling needs, maintaining a more stable temperature, increasing comfort faster, and hence more energy efficient. In addition, the energy-friendly operation can adjust the temperature setting in the building optimally by optimizing comfort with energy efficiency. The increase in energy load has been empirically proven that for every room air temperature decrease of 1℃ in office buildings, an observed 9% is required.

The energy efficiency efforts identified in the lighting components comprise effective placement and lighting automation based on occupancy and movement sensors. Lighting can consume approximately 20%-50% of the electricity consumption; therefore, managing the correct number of luminaries needed in a specific place inside a building is crucial in improving energy efficiency without reducing the lighting quality. According to Muhammad et al. (2010), effective lighting placement can help minimize energy consumption. Meanwhile, lighting control based on occupancy is one of the strategies that has been proven effective in saving energy in various buildings within the range of 20%-60% (De Bakker et al., 2018).

An energy efficiency strategy through the setting of an electricity plug aims to minimize energy use when the equipment is not in use. The intervention efforts can vary from the detection method and the sensor-based treatment to the electrical plugs. Acker and Duarte (2012) compared the level of energy consumption from the two types of sensors with conventional plugs. The first type of sensor is an occupancy sensor using passive infrared (PIR) technology, which can turn off all equipment connected to the outlet when no occupancy is detected. The second sensor type is a usage load detector that can turn off the plug based on the load-pulling device condition at the control outlet. These interventions can save energy by 19.75% and 16.97%, respectively.

Another energy efficiency strategy can be intervening in the occupants' behavior. According to Acker (2012), it can reduce energy use in the vacant period with a value of 4.14%. In addition, the same study's results show that compared to computers not meeting the energy requirements of ENERGY STAR®, energy-friendly electronic devices passed the standard contribute to considerable energy savings at an average of 14.49% on working days.

The initial identification of smart building energy efficiency strategies was then presented to two Indonesian experts: (1) an experienced engineer in smart building technology and (2) a senior researcher in energy-efficient mechanical air conditioning systems for buildings. The experts stated that implementing smart building technology must consider several factors, including the automated building area, human input controlling preferences, supervision forms, the desired automation system, and sensor types used. According to Expert 1's experience, all the proposed efficiency measures are consistent with the implementation of smart building development. However, according to Expert 2, several variable combinations have the same category coverage; hence, variable simplification is needed. Table 4 summarizes the energy efficiency strategies and their respective efficiencies validated by the expert.

TABLE 4: Energy efficiency efforts in the smart building components

No.	Category		Energy Efficiency Strategies	Efficiency (%)	Reference
I	AC	1.	AC and ventilation system	9% every 1 ℃, efficiency 5.34% – >70%	(Andarini, 2014; Chua et al., 2013; Siriwardhana and Namal, 2017)
		2.	Dedicated operating system	27.9% - 32.4%	(Lee and Kim, 2020)
II	Lighting	3.	Lighting components with effective placement	30%	(Muhamad et al., 2010)
		4.	Lighting automation control based on occupancy and movement sensors	25-30%	(De Bakker et al., 2018; Ringel et al., 2019)
III	Plug and Electronics	5.	Utilization of sensors for energy use intervention at sockets / plug-strips	16.97% - 19.75%	(Acker et al., 2012)
		6.	Intervention in the behavior of using electrical devices in buildings (Personal Computer, laptop, printer)	4.14%	(Acker et al., 2012; Kwong et al., 2014)
		7.	Use and control of energy-friendly electronic devices	14.49%	(Acker et al., 2012; Kwong et al., 2014)
		8.	Power usage/power management system settings	10.15%	(Kwong et al., 2014)

4.2 KNN-Based ML Model Development

The KNN algorithm employs the distance between points as a reference for predicting values. The features must have the same value scale (0 and 1) to ensure no imbalance when calculating the distance between constituent points. Therefore, the scaling procedure, which involves adjusting the scale of the feature to make it comparable to other compilers, must be performed before the development of the ML model. Fig. 5 depicts the feature before the scaling process (left) and after the scaling process (right).

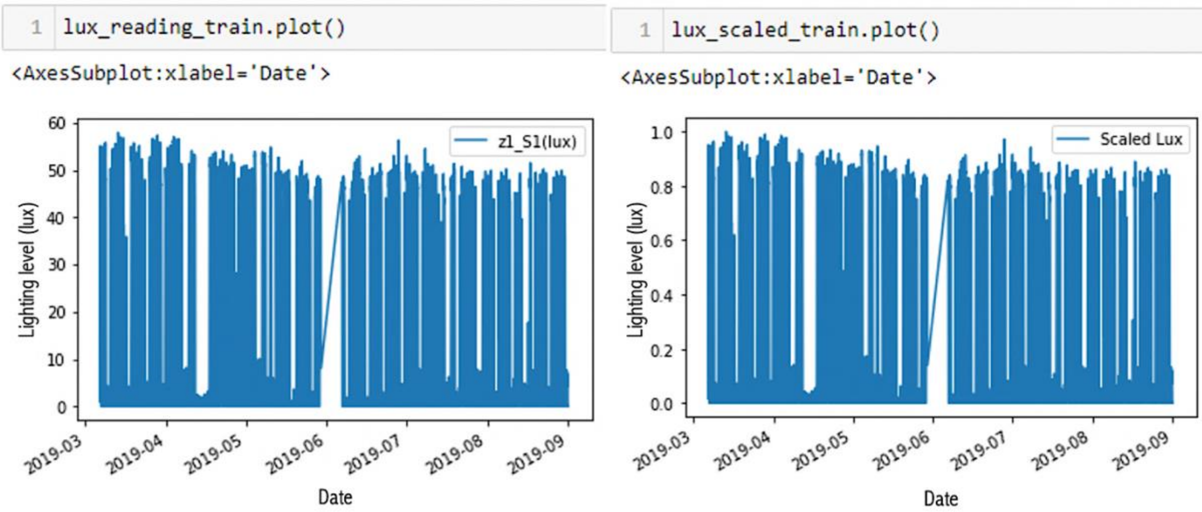

FIG. 5: Features before (left) and after the scaling process (right)

ML model training combines energy usage data from energy meters and sensors as inputs to produce a model predicting energy consumption. In addition to the features scaled, one of the inputs for training the model is to assign an energy use value to the component reviewed. This process was carried out by accessing the value of the energy use component in the dictionary based on the zone and section reviewed. The components featured in the training model and additional features used to predict and its relationships are illustrated in Fig. 6.

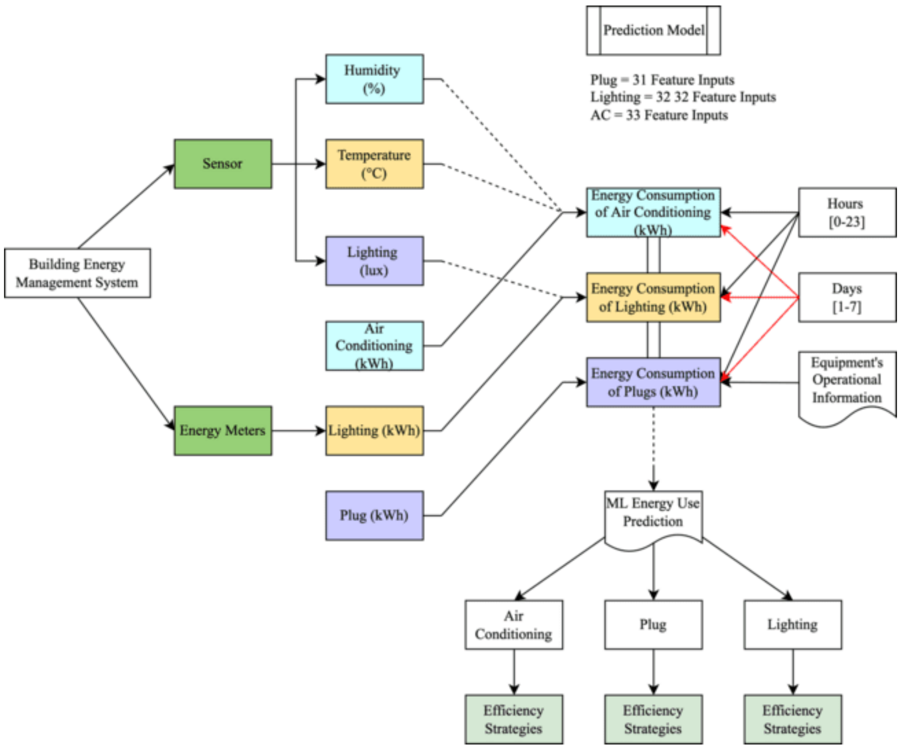

FIG. 6: Feature inputs for each energy component in ML model development

This prediction process included testing data as a model feature. The predictions result from recognizing patterns in the data based on daily and weekly usage patterns and features that have been scaled. The predicted values generated by the ML model were then compiled to facilitate the visualization of the comparison between the actual values and the predicted values. The visualization in Fig. 7 demonstrates the model performance with a more in-depth calculation by showing the magnitude of the error between points and the total energy prediction for the components in Zone 1 of the sixth floor as the reviewed zone and floor.

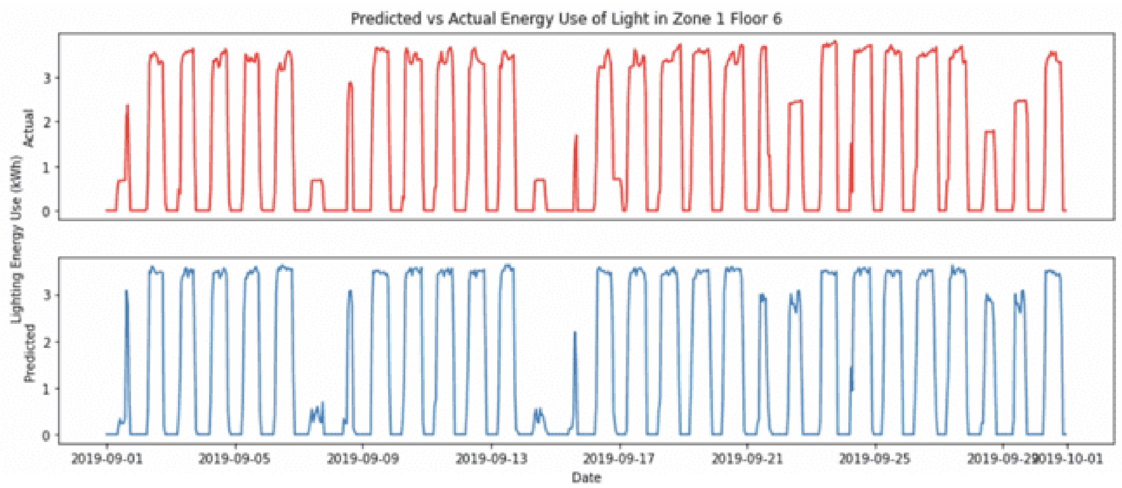

FIG. 7: The comparison between actual and predicted data for lighting energy in zone 1 of the 6th floor

The processed feature options are then examined to determine whether the inclusion of features as a model trainer parameter provides a correlation with the model's predicted results. The influence of these various features will be divided based on the prediction components, which include air conditioning, illumination, and switches/plugs. The reference for the standard of the obtained correlation values by Miyamoto et al. (2017) was used, which classifies ranges of correlation coefficient values and describes the range values (see Table 5).

TABLE 5: Correlation value standard

Correlation Value	**Correlation Degree**
0.0 - ±0.2	Little Correlation
±0.2 - ±0.4	Weak Correlation
±0.4 - ±0.7	Correlated
±0.7 - ±0.9	Strong Correlation
±0.9 - ±1.0	Very Strong Correlation

The correlation value for the quantity of AC energy consumption derived from the hours of use is 0.12, which is in the range of little correlation. This value was determined by calculating the average usage hours of all features. The average correlation value for days of use is 0.13, which places it in the same category. Furthermore, the correlation value for the relative humidity feature is above 0.37, indicating that the feature plays an adequate role in assisting the model in predicting the amount of AC energy consumption. Meanwhile, the temperature feature has a little correlation with a 0.16 value, reflecting the level of AC energy use caused by a decrease in the correlation between temperature and the level of energy use for the duration outside the use of buildings. The aggregate correlation value between AC features is depicted in Fig. 8(a), along with a color legend indicating the magnitude of the correlation between the features comprising the AC energy use level.

The correlation value of the lighting component's features to the energy used for lighting in buildings is nearly the same as that of the AC component for hours of use and days of use, which is in the little correlation category. This suggests that the information provided about specific hours and days can create improved forecasts of the value of energy use due to the discovery of a modest link. Meanwhile, the light level feature detected by the sensor is also in the weak correlation range with a 0.2 value. This can be understood as variations in the illumination level read by the sensor helping forecast the degree of energy use but are not accurate without the support of other features.

The correlation values between light component features are shown in Fig. 8(b), with a color legend indicating the magnitude of the correlation between the associated features.

Furthermore, the magnitude of the correlation value between plug component characteristics and the level of energy consumption for the requirements of building electronic devices, which consists of hours of use and days of use, also falls within the range of little correlation. It indicates that energy consumption value varies between hours and days. Since no sensor readings can aid the model in predicting the plug's energy consumption, this prediction is solely based on the energy consumption patterns of the previous model's trainer durations. In general, correlation values for plugs are highest between 10 p.m. and midnight and on days of use other than Saturday. Therefore, a relatively low correlation value is derived, as the average use of hours and days does not correlate significantly with the level of plug energy consumption, as shown in Fig. 8(c).

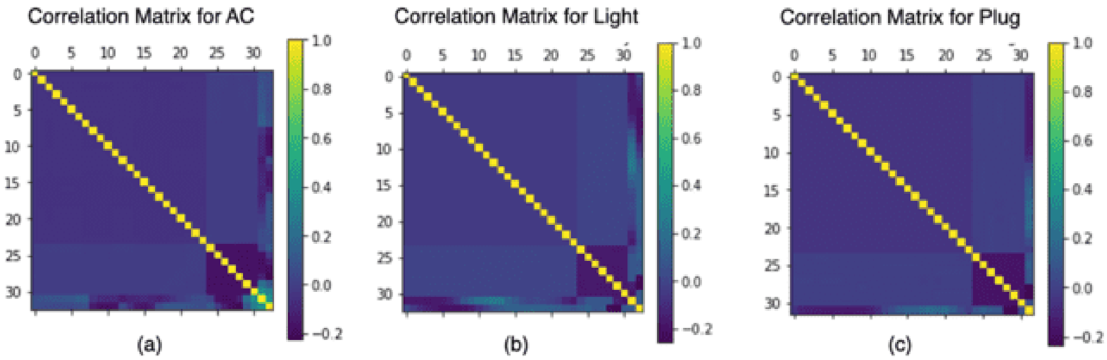

FIG. 8: The correlation matrix for the AC, Light, and Plug

After the training data features were entered into the model, a model predicting the data with the appropriate number of feature inputs was developed. The model accuracy was evaluated by using the testing data that was previously prepared. The accuracy evaluation was carried out by comparing the predicted energy value to the amount of actual energy consumption.

The magnitude of the relative error was known through the difference between the total predicted energy use and the actual values divided by the amount of the original energy, reflecting the performance of the overall prediction. On the other hand, the MAPE value obtained using the average error value for each observation reflects how far each prediction is from the original value of energy use. The relative error and MAPE value from the prediction results of each floor are summarized in Table 6. MAPE values ranging between 10% and 20% denote that the models are good predictors, while models with MAPE ranging between 20% and 50% are considered acceptable (Elmousalami, 2021). Therefore, predictive models developed for each floor of the Chamchuri 5 building are fit based on the accuracy values.

TABLE 6: MAPE value of each floor's prediction results

Floor No.	Average Relative Error (%)	MAPE (%)
1	13.31	22.49
2	17.75	20.74
3	18.9	23.26
4	18.13	22.15
5	26.78	22.62
6	5.63	15.37
7	23.85	27.45

The energy consumption for each component on each floor was then calculated based on the results of the model's energy use predictions on each floor and presented in Table 7. An assessment was then conducted based on this value to determine which components are significant to the overall level of energy use on each floor and for the entire building.

TABLE 7: *Total Predicted Energy Usage*

Floor	Total Predicted Energy Usage (kWh)		
	AC	**Light**	**Plug**
1	0	35,362.7	11,862.01
2	16,765.32	3,173.05	981.77
3	4.134.39	1,646.01	433.22
4	11.682.55	3,290.08	1,744.86
5	3,480.66	1,066.44	9,547.81
6	5,391.95	1,187.65	10,264.02
7	4,269.55	571.41	16,392.98

Since the level of energy consumption on the 1st floor is relatively high compared to the other floors, the 1st-floor component is omitted from the bar chart in Fig. 9 because its inclusion would distort the overall graphic scale and make it difficult to compare other floors. The 1st floor has the highest energy consumption due to its use as a lobby, necessitating a higher lighting level and longer operating hours. Air conditioning uses the most energy on other floors, followed by illumination, with plug components using the least energy. This must be considered when determining the priority of building components for energy efficiency initiatives.

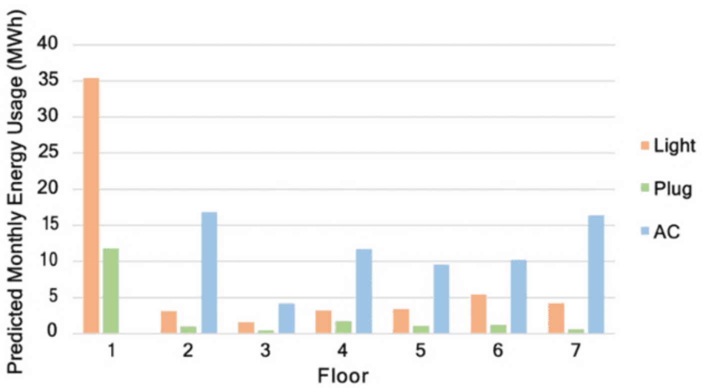

FIG. 9: The generated monthly energy use prediction of each floor

Based on the identified strategies and the efficiency level shown in Table 3, the lower limit value from the table was then used to determine the combined efficiency value if all energy efficiency efforts are applied. The percentages for strategies from X1 to X8 are 5.34%, 27.9%, 30%, 25%, 16.97%, 4.14%, 14.49%, and 10.15%, respectively. The efficiency percentages of those strategies were then combined and used to attain the total possible energy efficiency of each component on all floors achieved due to the application of all efficiency strategies. The combined efficiency percentages for AC, lighting, and plug are 31.75%, 47.5%, and 38.85%, respectively. The calculation for the 1st floor was obtained based on estimates with the highest efficiency percentage of 45.3%, implementing all strategies from the lighting and plug aspects. Meanwhile, the efficiency on other floors ranges from 34.5% to 37.5%, using efficiency measures in all three parts of AC, lighting, and plug. Table 8 shows the results of the potential energy efficiency calculations.

The results of this study have revealed significant discoveries in terms of the k-NN method's performance in predicting smart buildings' energy consumption. The findings are aligned with previous studies that used the k-NN algorithm to develop a predictive ML model. Research conducted by Wahid and Kim (2016) predicted the energy consumption type of apartment buildings based on the classification according to their energy usage. The k-NN classifier model in that study can predict whether an apartment has a low or high-power consumption with the accuracy performance ranging from 95.1-95.9% for the data training and testing ratios of 80-20%, 75-25%, and 70-30%. Furthermore, the result of this study follows the research that developed a k-NN predictive model with a 70-30% ratio of data training and testing, implying a low difference error in predicting the energy demand for the upcoming period of a smart commercial building (Mazlan et al., 2020).

TABLE 8: Total potential energy efficiency of each floor

Floor	Combined Energy Efficiency of Building Components			Total Energy Efficiency	
No.	AC	Light	Plug	kWh	%
1	-	16,797.3	4,608.4	21,405.7	45.3%
2	5,323.0	1,507.2	381.4	7,211.6	34.5%
3	1,312.7	781.9	168.3	2,262.8	36.4%
4	3,709.2	1,562.8	677.9	5,949.9	35.6%
5	3,031.4	1,653.3	414.3	5,099.1	36.2%
6	3,258.8	2,561.2	461.4	6,281.4	37.3%
7	5,204.8	2,028.0	222.0	7,454.8	35.1%

On the other hand, Shapi et al. (2021) indicated that the predictive model using k-NN has a slightly lower accuracy in predicting energy demand than the SVM algorithm. However, regarding the training time needed to predict the monthly average energy consumption, the ML model with the k-NN algorithm has the fastest to finish the training process compared to SVM and ANN. The ML model predicting school building designs' annual thermal and electricity consumption by Paterson et al. (2017) implied that the ANN method could predict energy usage if the features are within the range of collected design and briefing parameters used as the training data. In addition, to tailor the ANN-based model with better accuracy, the datasets fed for the model training and testing should be big enough for the complex learning process (Bourhnane et al., 2020).

Moreover, various studies have been conducted on the data-driven prediction for building energy consumption using Long Short-Term Memory (LSTM) method. It is a distinctive form of the recurrent neural network (RNN) that provides a more robust approach to dealing with sequential data in forecasting and time-series prediction than traditional RNN (Mahjoub et al., 2022). Panagiotou and Dounis (2022) investigated eight ML models for a hospital building's short-term load forecasting. Their study showed that LSTM outperformed the other algorithms regarding MAPE, most likely due to its ability to learn long-term dependencies from time series. Based on the comparison to previous studies, the present study developed a predictive ML model utilizing the k-NN method with public timely data series of a smart building's energy consumption. This study particularly emphasizes the energy consumption from the features identified as the smart building characteristics that can significantly increase building energy efficiency.

4.3 SBEMS Dashboard Interface Design

SBEMS is crucial in optimizing building energy consumption, reducing waste, and promoting sustainability. The user interface (UI) design is an essential aspect of SBEMS because it facilitates effective energy management by building occupants or facility administrators. To understand more about the UI design of SBEMS, this research conducted a benchmark study of six selected papers, such as Hamidifar et al. (2009), Elbeltagi et al. (2021), Santos et al. (2020), Bonilla et al. (2018), Xu et al. (2018), and Utami et al. (2018).

The study conducted by Elbeltagi et al. (2021) concentrates on using artificial neural networks (ANN) to predict energy consumption in residential buildings through parametric modeling based on building characteristics, meteorological data, and occupant behavior. This article discusses using a user interface (UI) to input parameters, configure ANN, and visualize prediction results. The authors highlighted the need for a reliable, user-friendly, and interactive UI that facilitates data entry, visualization, and interpretation as a decision support tool for users to conduct energy consumption prediction without any experience in modeling and simulation tools. On the other hand, Hamidifar et al. (2009), which developed a graphical user interface (GUI) for a PHEV energy management system, emphasizes the significance of an aesthetically pleasing and user-friendly graphical user interface in enhancing the user experience and facilitating effective control of energy consumption. The GUI provides real-time information on the PHEV's energy consumption and performance, enabling the driver to make informed decisions.

The paper by Santos et al. (2020) introduces BRICKS, Building's Reasoning for Intelligent Control Knowledge-based System, a modular, interoperable building management system supported by a semantic context-aware rule-based system, allowing for changes to be made without the need for reprogramming the system. BRICKS also has an intuitive graphical user interface (GUI) that abstracts the user from all the semantic configurations. The proposed GUI enables users to define and modify control strategies following building characteristics, energy policies, and user preferences. It makes it easier for users to interact with the system without understanding all its underlying technical details. Additionally, the GUI visually represents energy consumption, comfort levels, and control actions. Bonilla et al. (2018) developed a practical and low-cost monitoring tool for BEMS utilizing virtual instrumentation called Virtual Energy Management System (VMS). It displays and records the information on electrical parameters, electricity costs, and carbon footprint on a computer. The GUI of the VMS was developed on the LabVIEW 2013® virtual programming platform to enable users to track energy consumption, view historical data, and configure alarm settings. Moreover, the user can visualize the monitoring display and access records remotely through the internet. This study emphasized the significance of a simple and intuitive UI that provides real-time feedback, actionable insights, and simple customization options to facilitate effective energy management to be used by employees, building operators, and facility managers.

Furthermore, a generic UI for building operating systems was developed by Xu et al. (2018) through design, implementation, and evaluation, enabling users to monitor and control energy consumption in various appliances, set energy-saving preferences, and receive feedback regarding energy usage. A series of generic data models independent of any building operating system was designed to ensure that the UI could be flexibly adapted to various buildings. The study also introduced three roles with different permissions and several functional components to support a wide range of use cases relating to smart homes. After designing the UI, the authors implemented a prototype called Building Operating System User Interface (BOS UI), which was evaluated in terms of design and usability, and conducted using a combination of qualitative and quantitative methods, including expert reviews, user testing, and surveys. On the other hand, a study by Utami et al. (2018) discussed an energy monitoring system (EMS) for extant structures in Indonesia. Using the concept of Engineering, Measurement, and Testing (EM&T), the authors developed a web-based GUI that provides real-time and historical energy consumption data and suggestions for energy-saving measures. The paper highlights the significance of a user-friendly and accessible GUI that provides exhaustive information, actionable insights, and customizable settings to assist building owners and facility administrators make informed decisions.

Based on the discussion of these previous studies, it can be concluded that the output of the KNN-based ML model developed in this study can be implemented into SBEMS and be made more user-friendly with the assistance of an interface. The UI should be designed to facilitate communication between building occupants and management to utilize available information. Before UI design, a data processing infrastructure was developed to convert the CSV-formatted energy and sensor recording data into outputs supporting dashboard use. These outputs include energy prediction values, efficiency calculations, and the division of energy readings for zones and floors.

In the dashboard, the data of energy use reports are displayed in graphs and charts to be further utilized as statistical analysis. Building components highly connected to the building's energy use, such as AC, lighting, plugs, and devices, will be presented. Fig. 10 shows the schematic of data logic for the dashboard, divided into three different colored zones. The stages on the green zone are the logic for accessing the data in the dashboard display related to the selection of floors and zones. When a floor is selected, the program accesses floor energy recording data in .CSV format and executes code to detect components in each zone. Then, when a zone is selected, access to the dictionary based on the selected zone is performed, followed by the execution of the developed k-NN-based ML model.

Subsequently, the selected floors and zones will be accessed for energy use for each building component (AC, Plug, Light) based on the developed ML model. The ML-based energy use will then be the initial energy requirement level shown on the SBEMS dashboard. Afterward, the stages on the yellow zone show the data access logic for each building component based on the ML model output that predicts the energy use level when the SBEMS dashboard is used for the desired duration. Furthermore, the blue zone shows the potential interventions of the energy use level, the identified smart building energy efficiency strategies, with their respective efficiency levels, are loaded for further calculation. The result shows the amount of energy and total efficiency based on the application of any intervention.

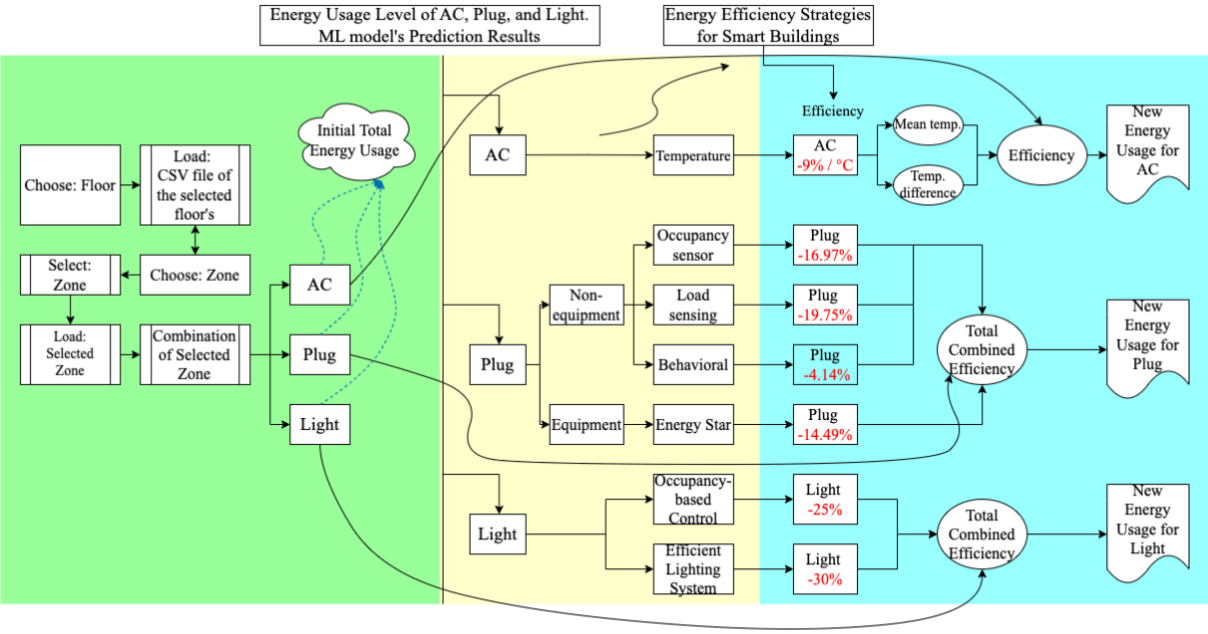

FIG. 10: The logic of accessing energy data in the SBEMS dashboard

The usage of this SBEMS dashboard comprises several activities. Started with the selection of floors and zones in the initial phase. Clicking the floor and zone options will load the recording data on the selected floor and zone and display the monthly energy demand based on the ML model's prediction of energy observation data and building sensors from previous months. Fig. 11 shows the UI of the SBEMS dashboard after the zones and floors have been selected, particularly the second floor and zone 3.

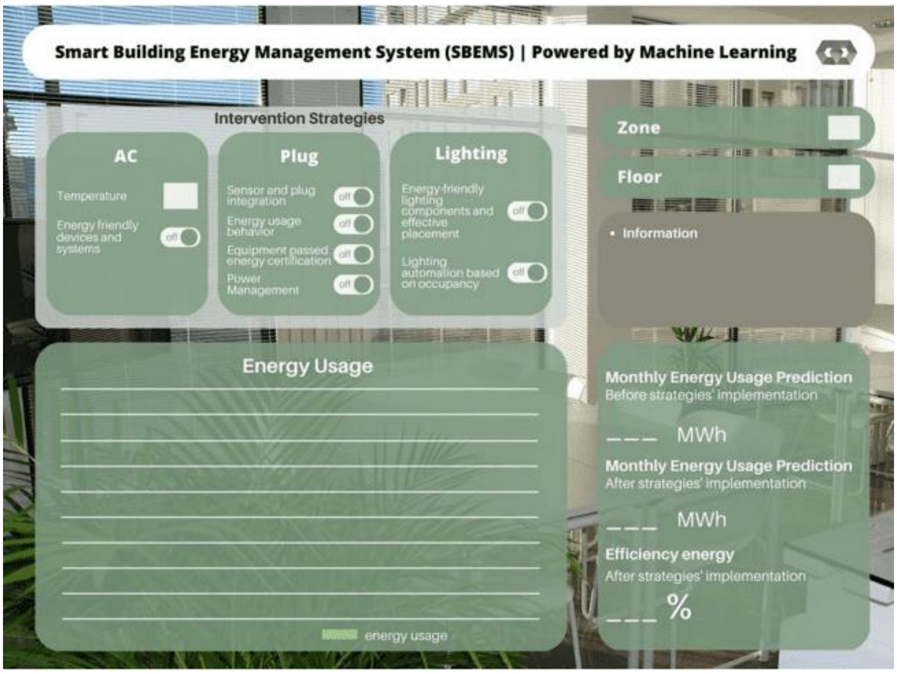

FIG. 11: The UI of the Initial display of the dashboard in the SBEMS Dashboard

When the floors and zones have been selected, the recording data for the selected floors and zones will be loaded, and the monthly energy demand will be displayed based on the predicted model of energy observation data and sensor buildings in the previous months. Fig. 12 presents the display after selecting zone and floor.

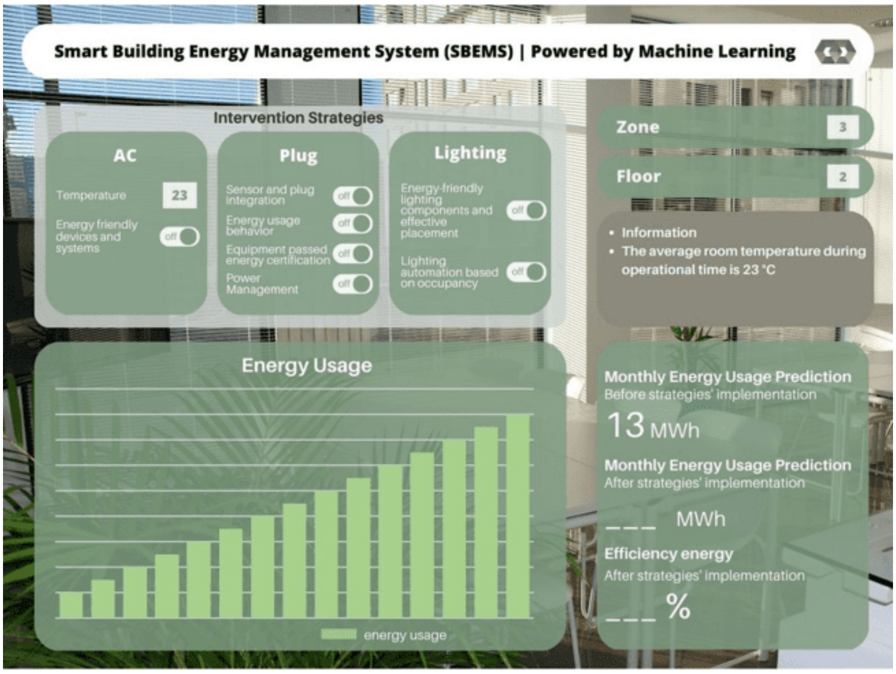

FIG. 12: The UI of Selection of Zone and Floor in the SBEMS Dashboard

The subsequent phase involves selecting energy-saving measures, in which the required energy value is presented. In addition, the amount of efficiency when the intervention is applied to the specified zone and floor is also shown, accompanied by a graphical representation. A brief description of the employed intervention in the information column is shown at the foot of the page pertaining to the section's effectiveness (see Fig. 13).

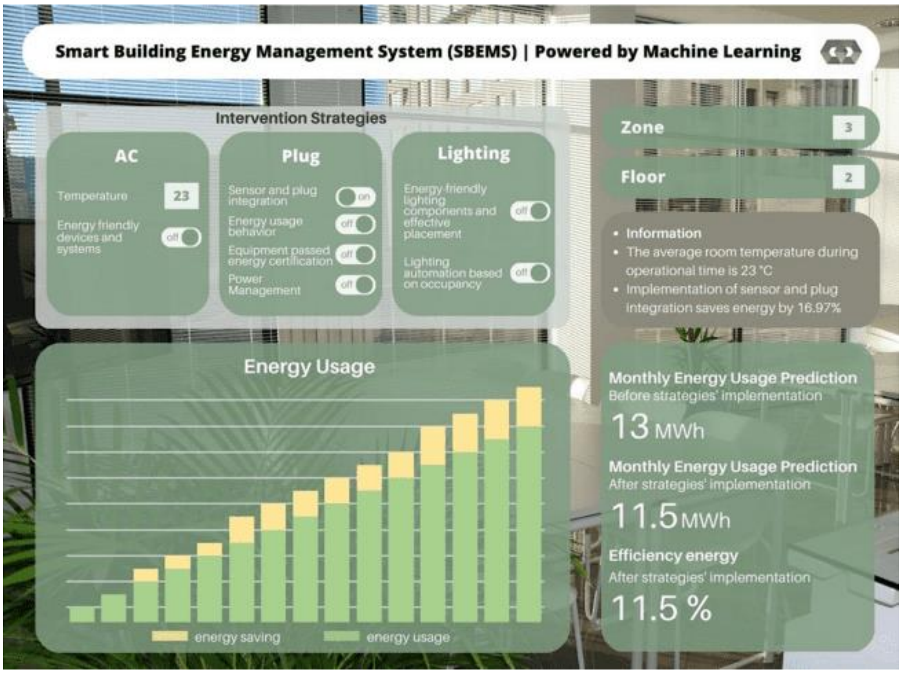

FIG. 13: The UI of Selection of Energy Efficient Intervention in the SBEMS Dashboard

The ML-based SBEMS proposed in this study has the potential to significantly increase the adoption of smart building concepts in developing countries, by identifying the cost-effective energy management strategies applicable to real-world situations and developing the user-friendly dashboard design to convey the outputs of the ML model, allowing building owners and operators to easily interpret and utilize the information provided by the model and facilitating effective decision-making and energy conservation initiatives.

5. CONCLUSION

Despite its benefits, the adoption of smart buildings is slow in developing countries, particularly due to the high initial investment cost and the knowledge gap between the users' perception and the actual impact of smart buildings. To address the lack of knowledge on smart building's impact, this study investigated the design strategies in smart that promotes energy efficiency. Furthermore, the study acknowledges the relative lack of research on long-term building energy consumption prediction, attributed to data insufficiency and the complexity of occupant energy use behavior. By developing an ML model and utilizing publicly available data, the study addresses this research gap and provides insights into long-term energy consumption prediction in smart buildings. The findings of this study showed that energy efficiency initiatives applied to implement the smart building concept are categorized into AC, lighting, and plug/electronic devices, where each initiatice can generate an energy efficiency of 31.75%, 47.5%, and 38.85%, respectively. The ML model developed with the k-NN algorithm to predict building energy consumption was proposed to help improve the energy efficiency level in the operation of smart buildings. This ML model has an average prediction relative error rate of 17.76% and a MAPE value of 22.01%, considered a good predictor. However, due to some limitations on the algorithm used, this study encourages future research to develop the ML model with other algorithms that can be compared, and the most excellent predictive model can be proposed. Moreover, the dashboard concept design was created to convey the outputs of the developed ML model in a user-friendly form so that it may be used as an assessment tool for building energy conservation initiatives.

ACKNOWLEDGEMENT

This is an extended and updated version of one of the best papers presented initially at CONVR 2022 organized and led by Prof. Chansik Park of Chung Ang University in Seoul, Korea. The authors would like to like to thank Indonesian National Research and Innovation Agency (BRIN) and Universitas Indonesia for the grant given to this research.

REFERENCES

Abdennadher, I., Khabou, N., Rodriguez, I. B., and Jmaiel, M. (2016). Designing energy efficient Smart Buildings in ubiquitous environments. *International Conference on Intelligent Systems Design and Applications, ISDA*, *2016-June*, 122–127. https://doi.org/10.1109/ISDA.2015.7489212

Acker, B., Duarte, C., and Wymelenberg, K. van den. (2012). Office Space Plug Load Profiles and Energy Saving Interventions. *ACEEE Summer Study on Energy Efficiency in Buildings*.

Aliero, M. S., Asif, M., Ghani, I., Pasha, M. F., and Jeong, S. R. (2022). Systematic Review Analysis on Smart Building: Challenges and Opportunities. In *Sustainability (Switzerland)* (Vol. 14, Issue 5). https://doi.org/10.3390/su14053009

Amasyali, K., and El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. In *Renewable and Sustainable Energy Reviews* (Vol. 81). https://doi.org/10.1016/j.rser.2017.04.095

Andarini, R. (2014). The role of building thermal simulation for energy efficient building design. *Energy Procedia*, *47*. https://doi.org/10.1016/j.egypro.2014.01.217

Apanaviciene, R., Vanagas, A., and Fokaides, P. A. (2020). Smart building integration into a smart city (SBISC): Development of a new evaluation framework. *Energies*, *13*(9). https://doi.org/10.3390/en13092190

Ascione, F., Bianco, N., de Stasio, C., Mauro, G. M., and Vanoli, G. P. (2017). Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach. *Energy*. https://doi.org/10.1016/j.energy.2016.10.126

Batov, E. I. (2015). The distinctive features of "smart" buildings. *Procedia Engineering*, *111*. https://doi.org/10.1016/j.proeng.2015.07.061

Benavente-Peces, C. (2019). On the energy efficiency in the next generation of smart buildings—supporting technologies and techniques. *Energies*, *12*(22). https://doi.org/10.3390/en12224399

Berawi, M. A., Miraj, P., Sayuti, M. S., and Berawi, A. R. B. (2017). Improving building performance using smart building concept: Benefit cost ratio comparison. *AIP Conference Proceedings*, *1903*(1), 030001. https://doi.org/10.1063/1.5011508

Bonilla, D., Samaniego, M. G., Ramos, R., and Campbell, H. (2018). Practical and low-cost monitoring tool for building energy management systems using virtual instrumentation. *Sustainable Cities and Society*, *39*. https://doi.org/10.1016/j.scs.2018.02.009

Bourdeau, M., Zhai, X. qiang, Nefzaoui, E., Guo, X., and Chatellier, P. (2019). Modeling and forecasting building energy consumption: A review of data-driven techniques. In *Sustainable Cities and Society* (Vol. 48, p. 101533). Elsevier Ltd. https://doi.org/10.1016/j.scs.2019.101533

Bourhnane, S., Abid, M. R., Lghoul, R., Zine-Dine, K., Elkamoun, N., and Benhaddou, D. (2020). Machine learning for energy consumption prediction and scheduling in smart buildings. *SN Applied Sciences*, *2*(2). https://doi.org/10.1007/s42452-020-2024-9

Buckman, A. H., Mayfield, M., and Beck, S. B. M. (2014). What is a smart building? *Smart and Sustainable Built Environment*, *3*(2). https://doi.org/10.1108/SASBE-01-2014-0003

Cao, X., Dai, X., and Liu, J. (2016). Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy and Buildings*. https://doi.org/10.1016/j.enbuild.2016.06.089

Chua, K. J., Chou, S. K., Yang, W. M., and Yan, J. (2013). Achieving better energy-efficient air conditioning - A review of technologies and strategies. In *Applied Energy* (Vol. 104). https://doi.org/10.1016/j.apenergy.2012.10.037

D'Amico, A., Bergonzoni, G., Pini, A., and Currà, E. (2020). BIM for healthy buildings: An integrated approach of architectural design based on IAQ prediction. *Sustainability (Switzerland)*, *12*(24). https://doi.org/10.3390/su122410417

De Bakker, C., Van De Voort, T., and Rosemann, A. (2018). The energy saving potential of occupancy-based lighting control strategies in open-plan offices: The influence of occupancy patterns. *Energies*, *11*(1). https://doi.org/10.3390/en11010002

Deng, H., Fannon, D., and Eckelman, M. J. (2018). Predictive modeling for US commercial building energy use: A comparison of existing statistical and machine learning algorithms using CBECS microdata. *Energy and Buildings*, *163*. https://doi.org/10.1016/j.enbuild.2017.12.031

Djenouri, D., Laidi, R., Djenouri, Y., and Balasingham, I. (2019). Machine learning for smart building applications: Review and taxonomy. *ACM Computing Surveys*. https://doi.org/10.1145/3311950

Edwards, W. K., and Grinter, R. E. (2001). At home with ubiquitous computing: Seven challenges. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *2201*, 256–272. https://doi.org/10.1007/3-540-45427-6_22

Eini, R., Linkous, L., Zohrabi, N., and Abdelwahed, S. (2021). Smart building management system: Performance specifications and design requirements. *Journal of Building Engineering*, *39*. https://doi.org/10.1016/j.jobe.2021.102222

Elbeltagi, E., and Wefki, H. (2021). Predicting energy consumption for residential buildings using ANN through parametric modeling. *Energy Reports*, *7*, 2534–2545. https://doi.org/10.1016/j.egyr.2021.04.053

El-Motasem, S., Khodeir, L. M., and Fathy Eid, A. (2021). Analysis of challenges facing smart buildings projects in Egypt. *Ain Shams Engineering Journal*, *12*(3). https://doi.org/10.1016/j.asej.2020.09.028

Elmousalami, H. H. (2021). Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction: A Case Study and Comparative Analysis. In *IEEE Transactions on Engineering Management* (Vol. 68, Issue 1, pp. 183–196). https://doi.org/10.1109/TEM.2020.2972078

European Commission. (2010). *Proposal for a Directive of the European Parliament and of the Council on the Energy Performance of Buildings*. Compendium of Sustainable Energy Laws. https://doi.org/10.1017/cbo9780511664885.046

Everitt, B. S., Landau, S., Leese, M., and Stahl, D. (2011). Cluster analysis: Fifth edition. In *Cluster Analysis: Fifth Edition*. wiley. https://doi.org/10.1002/9780470977811

Ghaffarianhoseini, A., Berardi, U., AlWaer, H., Chang, S., Halawa, E., Ghaffarianhoseini, A., and Clements-Croome, D. (2016). What is an intelligent building? Analysis of recent interpretations from an international perspective. *Architectural Science Review*, *59*(5), 338–357. https://doi.org/10.1080/00038628.2015.1079164

Grindvoll, H., Vermesan, O., Crosbie, T., Bahr, R., Dawood, N., and Revel, G. M. (2012). A wireless sensor network for intelligent building energy managementbased on multi communication standards-A case study. *Electronic Journal of Information Technology in Construction*, *17*.

Hamidifar, S., and Kar, N. C. (2009). Energy based graphical user interface modeling for PHEV energy management system. *2009 IEEE Electrical Power and Energy Conference, EPEC 2009*. https://doi.org/10.1109/EPEC.2009.5420922

Hong, J. (2010). *Why Is Great Design So Hard (Part Two)? | blog@CACM | Communications of the ACM*. Communications of the ACM. https://cacm.acm.org/blogs/blog-cacm/97958-why-is-great-design-so-hard-part-two/fulltext

Hossain, M. A., and Nadeem, A. (2019). Towards digitizing the construction industry: State of the art of construction 4.0. *ISEC 2019 - 10th International Structural Engineering and Construction Conference*. https://doi.org/10.14455/isec.res.2019.184

Kadoshin, S., Nishiyama, T., and Ito, T. (2000). The trend in current and near future energy consumption from a statistical perspective. *Applied Energy*, *67*(4), 407–417. https://doi.org/10.1016/S0306-2619(00)00033-7

Kim, S., and Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, *32*(3). https://doi.org/10.1016/j.ijforecast.2015.12.003

Kontokosta, C. E., and Tull, C. (2017). A data-driven predictive model of city-scale energy use in buildings. *Applied Energy*, *197*, 303–317. https://doi.org/10.1016/j.apenergy.2017.04.005

Kwong, Q. J., Goh, S. H., Adam, N. M., and Raghavan, V. R. (2014). A study on energy efficiency improvement opportunities for plug loads in buildings in the equatorial region. *Energy Procedia*, *56*(C). https://doi.org/10.1016/j.egypro.2014.07.201

Lee, J. W., and Kim, Y. il. (2020). Energy saving of a university building using a motion detection sensor and room management system. *Sustainability (Switzerland)*, *12*(22), 1–17. https://doi.org/10.3390/su12229471

Liu, X., Ren, M., Yang, Z., Yan, G., Guo, Y., Cheng, L., and Wu, C. (2022). A multi-step predictive deep reinforcement learning algorithm for HVAC control systems in smart buildings. *Energy*, *259*. https://doi.org/10.1016/j.energy.2022.124857

Mahjoub, S., Chrifi-Alaoui, L., Marhic, B., and Delahoche, L. (2022). Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks. *Sensors*, *22*(11). https://doi.org/10.3390/s22114062

Mazlan, N. L., Ramli, N. A., Awalin, L. J., Ismail, M. B., Kassim, A., and Menon, A. (2020). A Smart Building Energy Management using Internet of Things (IoT) and Machine Learning. *Test Engineering and Management*, *83*(8083–8090).

Mir, U., Abbasi, U., Mir, T., Kanwal, S., and Alamri, S. (2021). Energy Management in Smart Buildings and Homes: Current Approaches, a Hypothetical Solution, and Open Issues and Challenges. *IEEE Access*, *9*. https://doi.org/10.1109/ACCESS.2021.3092304

Missaoui, R., Joumaa, H., Ploix, S., and Bacha, S. (2014). Managing energy Smart Homes according to energy prices: Analysis of a Building Energy Management System. *Energy and Buildings*, *71*, 155–167. https://doi.org/10.1016/j.enbuild.2013.12.018

Molina-Solana, M., Ros, M., Ruiz, M. D., Gómez-Romero, J., and Martin-Bautista, M. J. (2017). Data science for building energy management: A review. In *Renewable and Sustainable Energy Reviews* (Vol. 70). https://doi.org/10.1016/j.rser.2016.11.132

Moseley, P. (2017). EU support for innovation and market uptake in smart buildings under the Horizon 2020 framework programme. *Buildings*, *7*(4). https://doi.org/10.3390/buildings7040105

Muhamad, W. N. W., Zain, M. Y. M., Wahab, N., Aziz, N. H. A., and Kadir, R. A. (2010). Energy efficient lighting system design for building. *ISMS 2010 - UKSim/AMSS 1st International Conference on Intelligent Systems, Modelling and Simulation*, 282–286. https://doi.org/10.1109/ISMS.2010.59

Nematchoua, M. K., Roshan, G. R., Tchinda, R., Nasrabadi, T., and Ricciardi, P. (2015). Climate change and its role in forecasting energy demand in buildings: A case study of Douala City, Cameroon. *Journal of Earth System Science*, *124*(1). https://doi.org/10.1007/s12040-014-0534-9

Omar, O. (2018). Intelligent building, definitions, factors and evaluation criteria of selection. *Alexandria Engineering Journal*, *57*(4). https://doi.org/10.1016/j.aej.2018.07.004

Panagiotou, D. K., and Dounis, A. I. (2022). Comparison of Hospital Building's Energy Consumption Prediction Using Artificial Neural Networks, ANFIS, and LSTM Network. *Energies*, *15*(17). https://doi.org/10.3390/en15176453

Panchalingam, R., and Chan, K. C. (2021). A state-of-the-art review on artificial intelligence for Smart Buildings. In *Intelligent Buildings International* (Vol. 13, Issue 4, pp. 203–226). Taylor and Francis Ltd. https://doi.org/10.1080/17508975.2019.1613219

Parisi, F., Fanti, M. P., and Mangini, A. M. (2021). Information and communication technologies applied to intelligent buildings: A review. *Journal of Information Technology in Construction*, *26*, 458–488. https://doi.org/10.36680/j.itcon.2021.025

Paterson, G., Mumovic, D., Das, P., and Kimpian, J. (2017). Energy use predictions with machine learning during architectural concept design. *Science and Technology for the Built Environment*. https://doi.org/10.1080/23744731.2017.1319176

Pipattanasomporn, M., Chitalia, G., Songsiri, J., Aswakul, C., Pora, W., Suwankawin, S., Audomvongseree, K., and Hoonchareon, N. (2020). CU-BEMS, smart building electricity consumption and indoor environmental sensor datasets. *Scientific Data 2020 7:1*, *7*(1), 1–14. https://doi.org/10.1038/s41597-020-00582-3

Piryonesi, S. M., and El-Diraby, T. E. (2020). Role of Data Analytics in Infrastructure Asset Management: Overcoming Data Size and Quality Problems. *Journal of Transportation Engineering, Part B: Pavements*, *146*(2). https://doi.org/10.1061/jpeodx.0000175

Qolomany, B., Al-Fuqaha, A., Gupta, A., Benhaddou, D., Alwajidi, S., Qadir, J., and Fong, A. C. (2019). Leveraging Machine Learning and Big Data for Smart Buildings: A Comprehensive Survey. *IEEE Access*. https://doi.org/10.1109/ACCESS.2019.2926642

Ringel, M., Laidi, R., and Djenouri, D. (2019). Multiple benefits through smart home energy management solutions—a simulation-based case study of a single-family-house in Algeria and Germany. *Energies*, *12*(8). https://doi.org/10.3390/en12081537

Santos, G., Vale, Z., Faria, P., and Gomes, L. (2020). BRICKS: Building's reasoning for intelligent control knowledge-based system. *Sustainable Cities and Society*, *52*. https://doi.org/10.1016/j.scs.2019.101832

Seyedzadeh, S., Pour Rahimian, F., Oliver, S., Rodriguez, S., and Glesk, I. (2020). Machine learning modelling for predicting non-domestic buildings energy performance: A model to support deep energy retrofit decision-making. *Applied Energy*, *279*. https://doi.org/10.1016/j.apenergy.2020.115908

Shapi, M. K. M., Ramli, N. A., and Awalin, L. J. (2021). Energy consumption prediction by using machine learning for smart building: Case study in Malaysia. *Developments in the Built Environment*, *5*. https://doi.org/10.1016/j.dibe.2020.100037

Singh, M. M., Singaravel, S., Klein, R., and Geyer, P. (2020). Quick energy prediction and comparison of options at the early design stage. *Advanced Engineering Informatics*, *46*. https://doi.org/10.1016/j.aei.2020.101185

Sinopoli, J. (2010). Smart Building Systems for Architects, Owners and Builders. In *Smart Building Systems for Architects, Owners and Builders*. https://doi.org/10.1016/c2009-0-20023-7

Siriwardhana, M., and Namal, A. (2017). Comparison of Energy Consumption Between a Standard Air Conditioner and an Inverter Type Air Conditioner in an Office Space. *SLEMA Journal*, *20*(1). https://doi.org/DOI: http://doi.org/10.4038/slemaj.v20i1-2.5

USAID. (2018). *Energy Performance of Buildings in Southeast Asia*. https://www.usaid.gov/energy/efficiency/examples/southeast-asia

Utami, S. S., Faridah, Azizi, N. A., Kencanawati, E., Tanjung, M. A., and Achmad, B. (2018). Energy Monitoring System for Existing Buildings in Indonesia. *E3S Web of Conferences*, *42*. https://doi.org/10.1051/e3sconf/20184201003

Wahid, F., and Kim, D. H. (2016). A prediction approach for demand analysis of energy consumption using K-nearest neighbor in residential buildings. *International Journal of Smart Home*, *10*(2), 97–108. https://doi.org/10.14257/ijsh.2016.10.2.10

Wang, Z., Wang, L., Dounis, A. I., and Yang, R. (2012). Integration of plug-in hybrid electric vehicles into energy and comfort management for smart building. *Energy and Buildings*, *47*, 260–266. https://doi.org/10.1016/j.enbuild.2011.11.048

Westermann, P., Welzel, M., and Evins, R. (2020). Using a deep temporal convolutional network as a building energy surrogate model that spans multiple climate zones. *Applied Energy*, *278*. https://doi.org/10.1016/j.apenergy.2020.115563

Xu, H., König, L., Cáliz, D., and Schmeck, H. (2018). A generic user interface for energy management in smart homes. *Energy Informatics*, *1*(1). https://doi.org/10.1186/s42162-018-0060-0

Yas, Z., and Jaafer, K. (2020). Factors influencing the spread of green building projects in the UAE. *Journal of Building Engineering*, *27*. https://doi.org/10.1016/j.jobe.2019.100894