Machine learning model for green building design prediction

Mustika Sari[1], Mohammed Ali Berawi[1,2], Teuku Yuri Zagloel[3], Rizka Wulan Triadji[2]
[1]Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia, Depok, Indonesia
[2]Center for Sustainable Infrastructure Development, Faculty of Engineering, Universitas Indonesia, Depok, Indonesia
[3]Department of Industrial Engineering, Faculty of Engineering, Universitas Indonesia, Depok, Indonesia

Article Info

Article history:

Received Feb 16, 2022
Revised Jun 30, 2022
Accepted Jul 29, 2022

Keywords:

Artificial neural network
Design prediction
Green building
Machine learning

ABSTRACT

Green building (GB) is a design concept that implements sustainable processes and green technologies in the building's life cycle. However, the design process of GB tends to take longer than conventional buildings due to the integration of various green requirements and performances into the building design. Advanced artificial intelligence (AI) methods such as machine learning (ML) are widely used to help designers do their jobs faster and more accurately. Therefore, this study aims to develop a GB design predictive model utilizing ML techniques that consider four GB design criteria: energy efficiency, indoor environmental quality, water efficiency, and site planning. A dataset of GB projects collected from a private construction company based in Jakarta was used to train and test the ML model. The accuracy of the models was evaluated using mean square error (MSE). The comparison of MSE values of the conducted experiments showed that the combination of the artificial neural network (ANN) method with the IF-ELSE algorithm created the most accurate ML model for GB design prediction with an MSE of 1.3.

Corresponding Author:

Mohammed Ali Berawi
Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia
Kampus Baru UI, Depok, Indonesia
Email: maberawi@eng.ui.ac.id

1. INTRODUCTION

Building and construction industry is known to imply negative impacts on the environment regarding excessive consumption of natural resources [1]. Furthermore, the building sector contributes more than 40% of greenhouse gas emissions and consumes not less than 40% of the global energy production [2]. Practitioners, professionals, and academics from the building and construction industry have attempted to find alternative approaches to practice energy conservation in the building life cycle. One of the efforts is implementing the green building (GB) concept [3]. GB concept refers to environmentally friendly and sustainable principles implemented in buildings' life cycle from the early stage of project planning, operation and maintenance to the decommissioning phase. It has been widely perceived as a strategy to minimize energy usage in the building and construction sector [4], [5]. GB concept applies principles and technologies to buildings throughout their life cycle to obtain sustainable purposes, such as minimizing the negative impacts on the environment caused by buildings and the human activities inside [6], [7].

Decisions made at the initial building design stage can significantly affect the environment [8]. However, due to various design aspects and building performances that must be set to achieve sustainability optimally, the design of GB tends to be more complex than conventional buildings [9]. Consequently, the design process of the GB can take longer due to the need for a multidisciplinary teamwork project where the team members should elaborate each GB aspect into the design [10], [11].

Technology advances that enable digitization, automation, and integration in the project life cycle have helped construction transform into a technology-driven industry by generating integrated systems and simplifying complex mechanisms that make the decision-making process more efficient [12], [13]. Furthermore, technology implementation has been proven to increase productivity in GB projects [14]. Machine learning (ML) is a technique that equips a system with the ability to learn and improve through its own experiences without being programmed [15]. It has been extensively researched and applied in the building life cycle [16], [17]. In the building design stage, this approach has been developed to optimize the building performance of GB design.

Previous studies conducted in these past few years regarding ML utilization in the building design process have given substantial contributions to the development of digital technology adoption in the building design process. It has revolutionized how the entire design process is performed [18], [19]. For example, a study [20] used the artificial neural networks (ANN) method to develop an ML model to predict reliable energy performance in office buildings that requires computation time that is 50 times faster than the standard building performance simulation tools. On the other hand, Statistical Neural Network & Gaussian Regression algorithms employed to develop an ML model to make fuel consumption predictions in a commercial building by Rahman and Smith [21] were proven to have better accuracy in doing so.

Furthermore, Geyer and Singaravel [22] developed a component-based ML model using the ANN method to predict thermal energy performance in office buildings. The computation time required to generate the prediction is drastically reduced with a small result of less than 3.9% error. It is in line with another study that compared the ANN and regression method for indoor air thermal condition prediction in residential buildings. The study results showed that even though the ML model with ANN takes time and needs much data to develop, it has a higher accuracy value of prediction results [23]. A framework to predict building performance at the design stage based on the interaction between buildings and humans developed using ANN algorithms was also proven to have an improved estimate [24].

These previous studies showed that the proposed predictive models using ML methods could significantly reduce the computation time required in the design process, increasing the productivity of architects and engineers designing GB. Despite the various development, however, there is still minimal evidence found on the usage of the ML approach in developing a prediction model for the design of the GB. This study attempts to create a design prediction model for GB using the ANN method as one of the ML techniques to address this gap. This paper is expected to provide references and give insights to building practitioners regarding the utilization of the ML approach in increasing the time efficiency of the GB design process, which can make a significant contribution toward the acceleration of technology-based development in the building and construction sector.

2. METHOD

This study was done in two stages to develop a predictive ML model for the design of GB as shown in Figure 1. The first stage is defining the GB design variables in the form of GB criteria and indicators used as parameters for the input and output of the ML model. These variables were obtained by performing a literature study of relevant research on GB published in the last five years, such as [22], [25]–[34], as well as GB assessment tools [35], [36], and regulations [37].

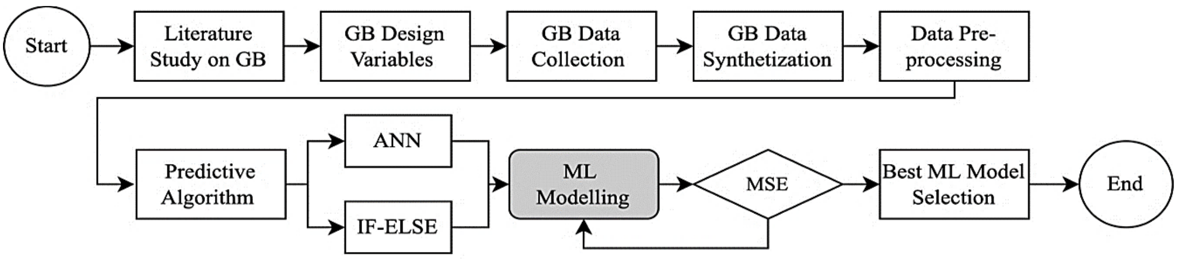


Figure 1. Research workflow

The experiments for the ML model development were performed in the second stage using the design variables and parameters obtained as the features for the ML model. Before the predetermined design variables were inputted into the experiments, a preprocessing step was performed to prepare the data.

Furthermore, the upper and lower limits for the value of each variable were determined [38], in which the values were based on the GB regulations applied both in Indonesia and other countries, as well as the archive analysis carried out on various documents discussing the design criteria of GB. The existing data on the GB projects, the ANN method, and IF-ELSE statements used to develop the ML model will be explained in the following sections.

2.1. Green building data collection

The historical data of GB design parameters used in this study were initially collected from a construction company based in Jakarta, Indonesia. However, due to the Non-Disclosure Agreement (NDA) between the contractor and owners of the GB projects, actual data cannot be fully provided. Subsequently, additional synthetic data were engineered to complete the data of the GB historical projects. Synthetic data is artificial data generated with the purpose of maintaining privacy for data sharing, which was used as training and testing data for ML model development [39], [40]. This data acquisition method has been used in ML development if the required data are not publicly accessed [41]. Synthetic data can be generated by adding actual or entirely synthesized data [42]. The synthetic data for ML model training and testing should be representative of the original dataset and based on existing standards [40].

The synthetic data was built based on the ranges of parameter values obtained from the Green Building Council Indonesia (GBCI), Jakarta Governor Regulation No. 38 of 2012 on Green Buildings, and the Jakarta Green Building User Guide issued by the Jakarta provincial government. Furthermore, Building Research Establishment Environmental Assessment Method (BREEAM) was also used as a reference for meeting data requirements [36].

2.2. Data analysis

Content analysis is the data analysis technique used to determine the variables and provide conclusions obtained from the literature study. It is a solid analytical technique for qualitative data with the systematic process used to conclude data in order of the text [43]. Due to various relevant studies' diverse views and perceptions, the analysis results were presented in tabular form. The table would be interpreted in four columns: criteria, variables, indicators, and references.

The missing data from the collected building data for ML training and testing were then completed by creating synthetic data using estimated ranges of values derived from the applied GB standards and regulations. The Microsoft Excel spreadsheet functions used in the data preprocessing step are random functions shown (1),

$$= RANDBETWEEN(lower\ limit, upper\ limit) \tag{1}$$

The random function was used to process data in the form of a minimum and maximum standards. The function then generated an integer random number from the two constraints that have been defined. Meanwhile, if the standard is a decimal number and then use the function,

$$= RAND(\ ) * (upper\ limit - lower\ limit) + lower\ limit \tag{2}$$

2.3. ML algorithms

The ML model developed was begun by importing the dataset completed in the previous step into the Python 3.7 programming language. The major specifications of the development environment are, 2.5 GHz Intel Core i5 CPU, an Intel HD Graphics 4000 with 1.5GB (1536 MB) of VRAM integrated GPU, and 4 GB RAM running. The programming code was compiled using the Google Collaboratory, a cloud service based on Jupyter Notebooks that disseminates machine learning research [44]. The packages used in this development were NumPy, pandas, Matplotlib, Scikit-learn, Tensorflow, and Keras. Furthermore, the sklearn preprocessing package that has the ability to transform raw datasets into a suitable representation was also used to perform data standardization quickly and straightforwardly.

The ML algorithms used in this experiment are the ANN and IF-ELSE algorithms. ANN is an artificial adaptive system inspired by human brain processes [45], the essential elements include node points known as processing elements (PE) and their relationships. Each node point has its input from communication between points or the environment and its output. Each of these vertices has a function that converts its general input into output. The nodes interact through the connections to generate the prediction. Since the GB design prediction include multiple inputs dan outputs; therefore, ANN as an algorithm that can provide predictions that resemble the learning processes of complex problems was selected. Furthermore, it has a high degree of flexibility in representing data regression [22].

Each relationship is characterized by the strength of the pair of nodes which gives a positive or negative value. A positive value means triggering, while a negative value means inhibiting [46]. The relationships between nodes can modify themselves, so this dynamic begins a learning process in the entire ANN, which is a key mechanism that characterizes ANNs [47]. All Pes in the ANN are interconnected with connection weights which are the basis of ANN's learning capabilities.

ANN can execute the data experimental knowledge in the training process and provide accurate predictions [38]. It consists of three main layers: the input, hidden, and output. The hyperparameters, which include the number of hidden layers, the number of nodes in the hidden layer, and the activation function, can be adjusted to the model's requirements at the time of model development to improve the quality of learning and provide an optimal model [48]. Since the architecture of the ANN network can be different for each ML model, the model selected is the one with the lowest deviation rate. The advantages of ANNs are their representational capabilities and universal function estimation capabilities, which are offered by feedforward neural networks [49]. The function of hidden neurons is to intervene in the external input and output of the network and allow the network to extract statistics at a higher level [50].

On the other hand, the IF-ELSE algorithm is usually used in making a decision among conditions or statements. It has several blocks that have different state conditions in each block. If the IF condition is true, then the true block of statements in the IF structure will be executed. However, when the IF condition is false, the false block of statements in the ELSE will be executed [51].

Among other metrics, mean squared error (MSE) is used to evaluate the performances of the ML model developed in each experiment due to its theoretical relevance in statistical modeling and sensitivity to outliers [52]. A model's MSE is the mean of the squared predictions error over all occurrences in the test set, in which prediction error shows the difference between the actual value and the predicted value [53]. MSE compresses all the training data and model predictions into a particular value measuring how well an ML model imitates reality.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2 \tag{3}$$

Where,
$n =$ number of items
$\sum =$ summation notation
$Y_i =$ actual
$\bar{Y}_i =$ prediction

3. RESULTS AND DISCUSSION

3.1. Green building criteria and indicators

The GB criteria used in the ML model development were obtained from the literature study. There are several leading design factors frequently discussed in GB guidelines and scientific publications, including indoor environmental quality, energy, water, material, waste, site planning, and innovation [33], [54]. However, to achieve the objective of this study, four particular design criteria that can be quantified were selected as the features for the ML development. The sub-criteria and indicators for the GB design criteria were also determined as shown in Table 1.

Table 1. GB indicators for predictive ML development

Criteria	Sub-criteria	Indicators	References
Energy	Building Geometry	Building Area	[22], [25]
		Building Orientation	[22], [25],
		Number of Floors	[22], [25], [26]
	Fenestration	Window wall ratio (WWR)	[22], [25], [26]
		Glazing Type	[27]
Indoor Environmental Quality	Visual Comfort	Indoor Illuminance	[27], [35]
	Thermal Comfort	Air Temperature	[27]–[29], [35]
		Relative Humidity	[27]–[29], [35]
	Acoustic Comfort	Sound Level	[35], [29], [28]
Water	Water Usage	Washbasin	[30], [35]
		Toilet Flush	[30], [35]
		Urinals	[30], [35]
	Water Recycling	Rainwater Harvesting	[31]–[35]
Site Planning	Site Planning	Landscape Area	[35]
		Cyclist Facilities	[31], [34], [35]

3.2. Parameters of green building design criteria

There are two sub-criteria in the energy efficiency criteria: building geometry and fenestration, with five indicators with determined parameters. The parameters for the indicators in building geometry that include building area and the number of floors were obtained from the historical data of the constructed GB. The GB projects varied from office, school, apartment, mall, and industrial building functions. The parameters of the building orientation are south-north and east-west.

The fulfillment of WWR standards of 20%–27% [55], in which the WWR calculation for overall thermal transfer value (OTTV) should not exceed 45 watts per square meter, as required in Governor Regulation No.38 of 2012 Article 6. Subsequently, the OTTV is 35.06–43.82 W/m^2. Furthermore, the thermal performance of the glazing type is shown by the U value, the measurement of heat loss (or heat flow) per square meter of surface area per 1-degree (Kelvin) temperature difference. The U value for the glazing type refers to the Jakarta Green Building User Guide document, stating that Indonesia's locally available U values are 4.94 W/m^2, 4.55 W/m^2, and 5.18 W/m^2.

In the indoor environmental quality criteria, the air temperature and relative humidity indicators were based on the GBCI and Governor Regulation No. 38 of 2012 Article 8 regarding the benchmark for thermal comfort that sets the air temperature plan at 25℃ and relative humidity at 60%. As for indoor illuminance, the lighting levels for different building functions were based on the GBCI, Governor regulation, and referring SNI-03-6197-2011 concerning Energy Conservation in Lighting Systems. Furthermore, the sound level was based on the GBCI regulation, which refers to SNI-03-6386-2000 concerning Specifications for Sound Levels and Reverberation Time in Buildings and Housing.

Parameters for sinks, toilet flushes, and urinals in the water efficiency criteria were based on standards by GBCI and Jakarta Green Building User Guide, providing a maximum value of water capacity of 8L/min for the sink, 4.5L/flush for the toilet flush, and 1.5L/flush for the urinal. In contrast, the minimum value is based on BREEAM UK: 3 liter/minute for the sink, 3 liter/flush for the toilet flush, and 0 liter /flush for the urinal. According to Governor Regulation No.38 of 2012 article 22, the volume of rainwater storage must be provided 5% of the ground floor area (GFA).

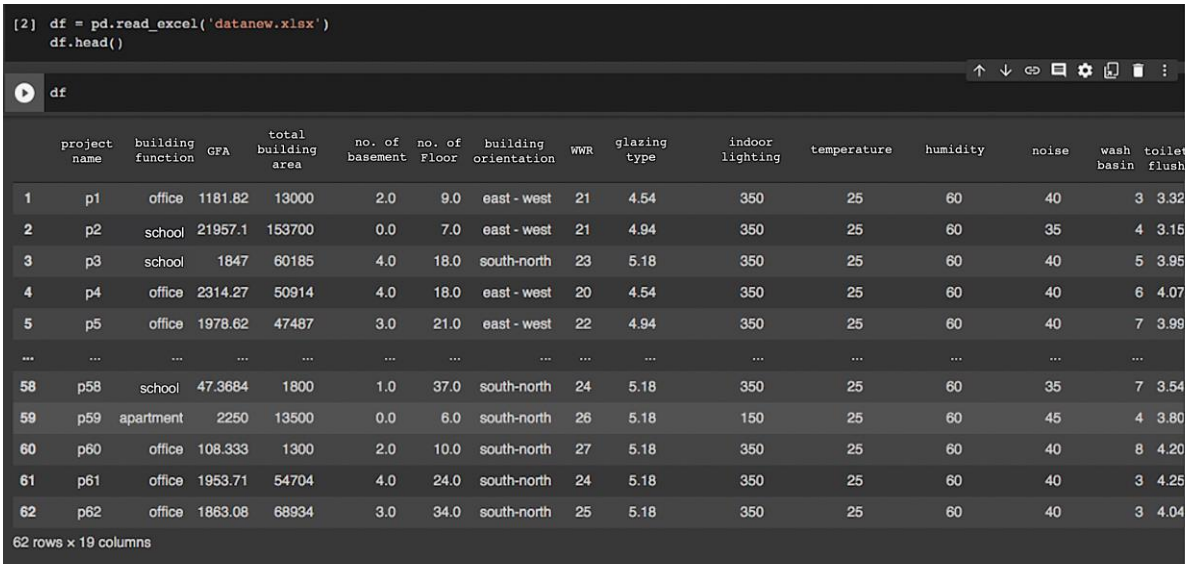
Based on Jakarta Governor Regulation No.38 of 2012 Article 21, the landscaping area in the building is 15% of the GFA for 5-story buildings, 30% of GFA for 9-story buildings, and 45% of the GFA for buildings higher than that. Moreover, referring to Article 25, bicycle parking facilities are at least one bicycle rack for every multiple of 2,500 square meters of building area. Table 2 summarizes the standards required for the GB indicators.

Table 2. GB variables and parameter standards

Variables	Indicators	Min	Max
Energy	Building Area	Historical Building Project Data	
	Building Orientation	South-North & East-West	
	Number of Floors	Historical Building Project Data	
	Window Wall Ratio (WWR)	20%	27%
	Glazing Type	4.54 W/m^{2}K	5.18 W/m^{2}K
Indoor Environmental Quality	Indoor Illuminance	Office & School:350 lux Apartment: 150 Lux Mall: 500 lux Industry: 500 Lux	
	Air Temperature	25 ℃	
	Relative Humidity	60%	
	Sound Level	Office: 40 dB School: 35 dB Apartments: 45 dB Mall: 45 dB Industry: 50 dB	
Water	Washbasin	3 L/min	8 L/min
	Toilet Flush	3 L/flush	4.5 L/flush
	Urinals	0 L/flush	1.5 L/flush
	Rainwater Harvesting	0.05 × ground floor area	
Site Planning	Landscape Area	15% of the ground floor area	45% of the ground floor area
	Cyclist Facilities	One bicycle rack/2,500 m2 building area	

3.3. Machine learning process

A total of 62 constructed GB was used as the training and testing datasets, with a ratio of 50 and 12 datasets, respectively. The built datasets were then inputted into the Google Collaboratory and processed by the ANN algorithm as shown in Figure 2. The text data were then converted into numeric data so that these data could be read and processed. The feature extraction process converted the data of the building function to numeric data, where (0) for the apartment, (1) for industry, (2) for the mall, (3) for school, and (4) for the office. While building orientation data was converted to (0) for east-west and (1) for south-north.



```
[2] df = pd.read_excel('datanew.xlsx')
    df.head()

df
```

	project name	building function	GFA	total building area	no. of basement	no. of Floor	building orientation	WWR	glazing type	indoor lighting	temperature	humidity	noise	wash basin	toilet flush
1	p1	office	1181.82	13000	2.0	9.0	east - west	21	4.54	350	25	60	40	3	3.32
2	p2	school	21957.1	153700	0.0	7.0	east - west	21	4.94	350	25	60	35	4	3.15
3	p3	school	1847	60185	4.0	18.0	south-north	23	5.18	350	25	60	40	5	3.95
4	p4	office	2314.27	50914	4.0	18.0	east - west	20	4.54	350	25	60	40	6	4.07
5	p5	office	1978.62	47487	3.0	21.0	east - west	22	4.94	350	25	60	40	7	3.99
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
58	p58	school	47.3684	1800	1.0	37.0	south-north	24	5.18	350	25	60	35	7	3.54
59	p59	apartment	2250	13500	0.0	6.0	south-north	26	5.18	150	25	60	45	4	3.80
60	p60	office	108.333	1300	2.0	10.0	south-north	27	5.18	350	25	60	40	8	4.20
61	p61	office	1953.71	54704	4.0	24.0	south-north	24	5.18	350	25	60	40	3	4.25
62	p62	office	1863.08	68934	3.0	34.0	south-north	25	5.18	350	25	60	40	3	4.04

62 rows × 19 columns

Figure 2. Data input process

Experiments were conducted four times during the model development using the ANN algorithm to find the best predictive model. All indicators became the output layer incorporated in the first attempt, including WWR, glazing type, temperature, relative humidity, indoor illuminance, noise levels, washbasins, toilet flush, urinals, rainwater harvesting landscaping areas, and cyclist facilities. The result is that the MSE is rather significant, at 5,119,586. In the second attempt, the MSE decreased to 443,484 because another hidden layer was added, making the model more accurate in predicting data and reducing error. Moreover, in the second experiment, data were divided into two smaller batch sizes of 31.

In the third attempt, the model tested only for one indicator, WWR, which showed a lower MSE of 128.9. It occurred due to the accumulated MSE value only coming from one indicator, while in the previous trial, the MSE value accumulated for all 12 indicators. Lastly, in the fourth attempt with the prediction outputs of washbasin, toilet flush, and urinal, the model has resulted in a low error rate compared to the previous trial process. For this reason, the model in the fourth experiment was considered the best model used in predicting the data because it has an MSE of 1.3. Therefore, the ANN algorithm was used for building function, ground floor area, building area, number of basements, number of floors, building orientation, toilet flush, urinals, and washbasin indicators as shown in Figure 3. The other indicators not included in the ANN algorithm were developed using the IF/ELSE algorithm, including WWR, type of glazing, indoor lighting level, temperature, relative humidity, noise level, rainwater harvesting, landscaping area, and cyclist facilities as shown in Figure 4. The trial-and-error processes of ML model development are summarized in Table 3.

Table 3. ML experiments

Layer	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Input Layer	Function, ground floor area, building area, number of basements, number of floors, building orientation			
Output Layer	All output indicators	All output indicators	WWR	Urinals, Toilet Flush, Washbasins
Hidden Layer	1	7	7	7
Epoch	50	50	50	50
Batch Size	-	31	31	31
Activation Function	Relu	Relu	Relu	Relu
MSE	5,119,586	443,484	128.9	1.3

The developed ML model can be used to generate a prediction for building design parameters with the building geometry data as the input. The building geometry is, a 9-story office building with 2-floor basements, GFA of 1,180 m^2, building area of 13,000 m^2, and east-west orientation. The results generated by the predictive model are 22% WWR, 4.94 W/m^2K for glazing type, 350 lux for indoor lighting, 25 °C for temperature, 60% for humidity, 40 dB for noise, 5.6 L/flush for washbasin, 3.9 L/flush for toilet flush, 0.9 L/flush for urinals, 650 m^3 for rainwater harvesting, and 354 m^2 for landscape area.

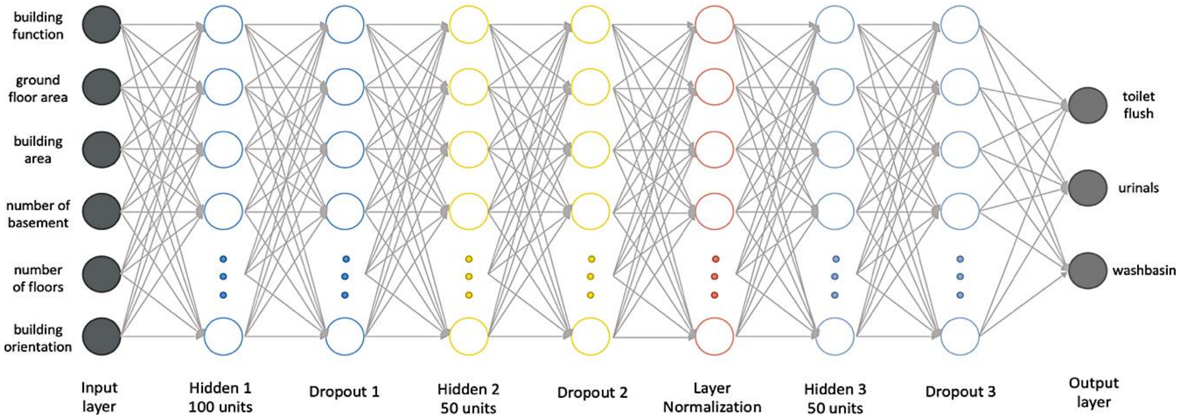


Figure 3. The Structure of the ANN algorithm

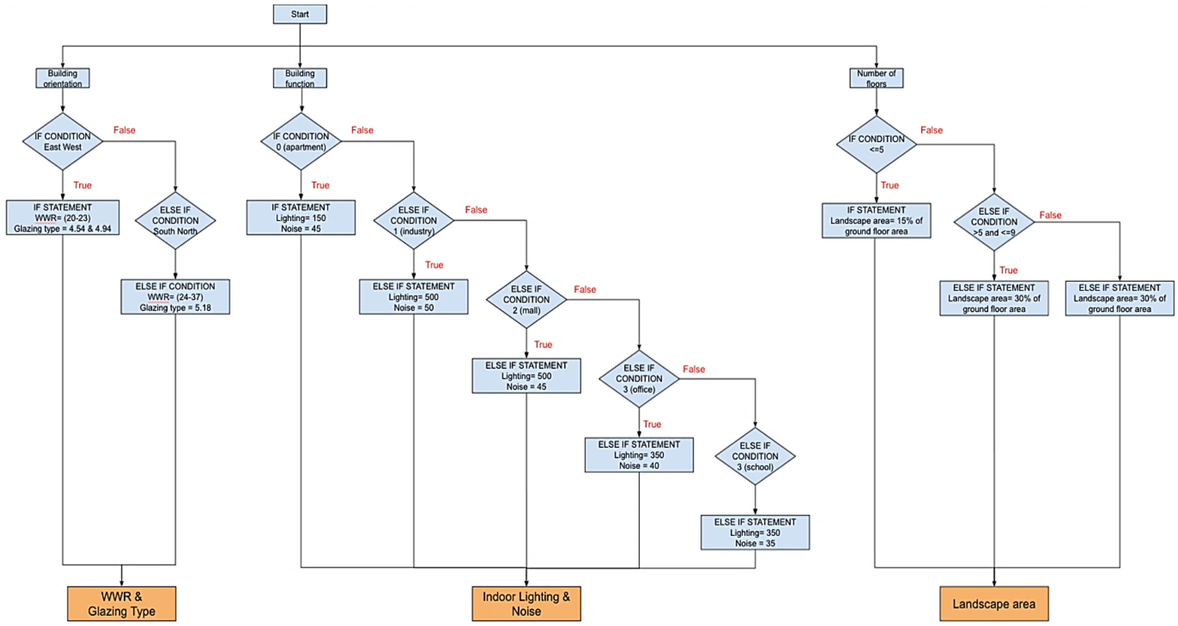


Figure 4. The structure of the IF/ELSE algorithm

4. CONCLUSION

The ML model for GB design was developed by considering the availability of existing data; the result is that the four GB design factors used in the ML model are energy efficiency, indoor environmental quality, water efficiency, and site planning. Moreover, the predictive model using IF-ELSE and ANN algorithms with an MSE of 1.3 was the most accurate. However, since this study has some limitations regarding the number of GB collected data, this study encourages future studies to develop a more robust ML model with improved accuracy performance by collecting more GB data. Furthermore, further research is needed to create the ML-based GB application design tool, easing designers' tasks during the conceptual phase of GB design.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Education, Culture, Research, and Technology, Republic of Indonesia, for the support given to this research (Agreement No. NKB-996/UN2.RST/HKP.05.00/2022).

REFERENCES

[1] M. S. Sandanayake, "Environmental Impacts of Construction in Building Industry—A Review of Knowledge Advances, Gaps and Future Directions," *Knowledge*, vol. 2, no. 1, pp. 139–156, 2022, doi: 10.3390/knowledge2010008.

[2] Z. Yas and K. Jaafer, "Factors influencing the spread of green building projects in the UAE," *Journal of Building Engineering*, vol. 27, 2020, doi: 10.1016/j.jobe.2019.100894.

[3] P. Miraj, M. A. Berawi, and S. R. Utami, "Economic feasibility of green office building: combining life cycle cost analysis and cost–benefit evaluation," *Building Research and Information*, vol. 49, no. 6, pp. 624–638, Mar. 2021, doi: 10.1080/09613218.2021.1896354.

[4] J. Zuo and Z. Y. Zhao, "Green building research-current status and future agenda: A review," *Renewable and Sustainable Energy Reviews*, vol. 30. pp. 271–281, 2014, doi: 10.1016/j.rser.2013.10.021.

[5] V. Venkataraman and J. C. P. Cheng, "Critical Success and Failure Factors for Managing Green Building Projects," *Journal of Architectural Engineering*, vol. 24, no. 4, pp. 1–10, 2018, doi: 10.1061/(ASCE)AE.1943-5568.0000327.

[6] P. K. D. Pramanik, B. Mukherjee, S. Pal, T. Pal, and S. P. Singh, *Green Smart Building*, no. June. 2019.

[7] B. Wen *et al.*, "The role and contribution of green buildings on sustainable development goals," *Building and Environment*, vol. 185, no. March, p. 107091, 2020, doi: 10.1016/j.buildenv.2020.107091.

[8] J. Basbagill, F. Flager, M. Lepech, and M. Fischer, "Application of life-cycle assessment to early stage building design for reduced embodied environmental impacts," *Building and Environment*, vol. 60, pp. 81–92, 2013, doi: 10.1016/j.buildenv.2012.11.009.

[9] F. Svalestuen, V. Knotten, O. Lædre, and J. Lohne, "Planning the building design process according to level of development," *Lean Construction Journal*, 2018.

[10] B. G. Hwang and L. P. Leong, "Comparison of schedule delay and causal factors between traditional and green construction projects," *Technological and Economic Development of Economy*, vol. 19, no. 2, pp. 310–330, 2013, doi: 10.3846/20294913.2013.798596.

[11] Y. Li, H. Song, P. Sang, P. H. Chen, and X. Liu, "Review of Critical Success Factors (CSFs) for green building projects," *Building and Environment*, vol. 158, no. May, pp. 182–191, 2019, doi: 10.1016/j.buildenv.2019.05.020.

[12] T. D. Oesterreich and F. Teuteberg, "Understanding the implications of digitisation and automation in the context of Industry 4.0: A triangulation approach and elements of a research agenda for the construction industry," *Computers in Industry*, vol. 83, pp. 121–139, Dec. 2016, doi: 10.1016/j.compind.2016.09.006.

[13] M. A. Berawi, "Managing artificial intelligence technology for added value," *International Journal of Technology*, vol. 11, no. 1, pp. 1–4, Jan. 2020, doi: 10.14716/ijtech.v11i1.3889.

[14] B.-G. Hwang, L. Zhu, and J. T. T. Ming, "Factors Affecting Productivity in Green Building Construction Projects: The Case of Singapore," *Journal of Management in Engineering*, vol. 33, no. 3, p. 4016052, 2017, doi: 10.1061/(asce)me.1943-5479.0000499.

[15] M. Kubat, *An Introduction to Machine Learning*. Springer International Publishing, 2017.

[16] N. Khean, A. Fabbri, M. H. Haeusler, S. Learning, and E. Framework, "Learning Machine Learning as an Architect , How to?," *Computing for a better tomorrow - Proceedings of the 36th eCAADe Conference, AI for Design and Built Environment*, vol. 1, pp. 95–102, 2018.

[17] T. Hong, Z. Wang, X. Luo, and W. Zhang, "State-of-the-art on research and applications of machine learning in the building life cycle," *Energy and Buildings*. 2020, doi: 10.1016/j.enbuild.2020.109831.

[18] E. Karan and S. Asadi, "Intelligent designer: A computational approach to automating design of windows in buildings," *Automation in Construction*, 2019, doi: 10.1016/j.autcon.2019.02.019.

[19] C. G. Belém, L. Santos, and A. M. Leitão, "On the Impact of Machine Learning. Architecture without Architects?" in CAAD Futures Rosetta View project Master Thesis on Structural Analysis for Algorithmic Design View project," no. March, 2019.

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

[21] A. Rahman and A. D. Smith, "Predicting fuel consumption for commercial buildings with machine learning algorithms," *Energy and Buildings*, 2017, doi: 10.1016/j.enbuild.2017.07.017.

[22] P. Geyer and S. Singaravel, "Component-based machine learning for performance prediction in building design," *Applied Energy*, 2018, doi: 10.1016/j.apenergy.2018.07.011.

[23] A. Ashtiani, P. A. Mirzaei, and F. Haghighat, "Indoor thermal condition in urban heat island: Comparison of the artificial neural network and regression methods prediction," *Energy and Buildings*, vol. 76, pp. 597–604, 2014, doi: 10.1016/j.enbuild.2014.03.018.

[24] C. Chokwitthaya, Y. Zhu, R. Dibiano, and S. Mukhopadhyay, "A machine learning algorithm to improve building performance modeling during design," *MethodsX*, vol. 7, pp. 35–49, Jan. 2020, doi: 10.1016/j.mex.2019.10.037.

[25] A. H. Al Ka'bi, "Comparison of energy simulation applications used in green building," *Annals of Telecommunications*, vol. 75, no. 7, pp. 271–290, 2020, doi: 10.1007/s12243-020-00771-6.

[26] W. Yu, B. Li, H. Jia, M. Zhang, and D. Wang, "Application of multi-objective genetic algorithm to optimize energy efficiency and thermal comfort in building design," *Energy and Buildings*, vol. 88, pp. 135–143, 2015, doi: 10.1016/j.enbuild.2014.11.063.

[27] R. Rameshwar, A. Solanki, A. Nayyar, and B. Mahapatra, "Green and Smart Buildings," no. January, 2019, pp. 146–163.

[28] C. Shen, K. Zhao, and J. Ge, "An Overview of the Green Building Performance Database," *Journal of Engineering (United Kingdom)*, vol. 2020, 2020, doi: 10.1155/2020/3780595.

[29] Y. Geng, W. Ji, Z. Wang, B. Lin, and Y. Zhu, "A review of operating performance in green buildings: Energy use, indoor environmental quality and occupant satisfaction," *Energy and Buildings*, vol. 183. Elsevier B.V., pp. 500–514, 2019, doi: 10.1016/j.enbuild.2018.11.017.

[30] I. Meireles and V. Sousa, "Assessing water, energy and emissions reduction from water conservation measures in buildings: a methodological approach," *Environmental Science and Pollution Research*, vol. 27, no. 5, pp. 4612–4629, 2020, doi:

10.1007/s11356-019-06377-3.
[31] G. S. Vyas and K. N. Jha, “Identification of green building attributes for the development of an assessment tool: a case study in India,” *Civil Engineering and Environmental Systems*, vol. 33, no. 4, pp. 313–334, 2016, doi: 10.1080/10286608.2016.1247832.
[32] P. C. Y. Liu, H. W. Lo, and J. J. H. Liou, “A combination of DEMATEL and BWM-based ANP methods for exploring the green building rating system in Taiwan,” *Sustainability (Switzerland)*, vol. 12, no. 8, p. 3216, 2020, doi: 10.3390/SU12083216.
[33] E. Yadegaridehkordi *et al.*, “Assessment of sustainability indicators for green building manufacturing using fuzzy multi-criteria decision making approach,” *Journal of Cleaner Production*, vol. 277, p. 122905, 2020, doi: 10.1016/j.jclepro.2020.122905.
[34] R. Shad, M. Khorrami, and M. Ghaemi, “Developing an Iranian green building assessment tool using decision making methods and geographical information system: Case study in Mashhad city,” *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 324–340, 2017, doi: 10.1016/j.rser.2016.09.004.
[35] Green Building Council Indonesia, *Greenship for New Building: Summary of Criteria and Benefits*. Jakarta, Indonesia: GBCI, 2014.
[36] BREEAM, “Technical Standards | BREEAM - Sustainability Assessment Method.” 2021, Accessed: Feb. 14, 2022. [Online]. Available: https://www.breeam.com/discover/technical-standards/.
[37] Governor of Jakarta, “The Governor Regulation of DKI Jakarta No. 38/2012 on Green Buildings.” 2012, [Online]. Available: https://www.iea.org/policies/2523-jakarta-regulation-no-382012-on-green-buildings.
[38] Z. Liu *et al.*, “Accuracy analyses and model comparison of machine learning adopted in building energy consumption prediction,” *Energy Exploration and Exploitation*. 2019, doi: 10.1177/0144598718822400.
[39] J. Vaidya, B. Shafiq, M. Asani, N. Adam, X. Jiang, and L. Ohno-Machado, “A Scalable Privacy-preserving Data Generation Methodology for Exploratory Analysis,” *AMIA ... Annual Symposium proceedings. AMIA Symposium*, vol. 2017, pp. 1695–1704, 2017.
[40] R. Heyburn *et al.*, “Machine learning using synthetic and real data: Similarity of evaluation metrics for different healthcare datasets and for different algorithms,” 2018, pp. 1281–1291, doi: 10.1142/9789813273238_0160.
[41] D. Rankin, M. Black, R. Bond, J. Wallace, M. Mulvenna, and G. Epelde, “Reliability of supervised machine learning using synthetic data in health care: Model to preserve privacy for data sharing,” *JMIR Medical Informatics*, vol. 8, no. 7, 2020, doi: 10.2196/18910.
[42] N. Dalsania, Z. Patel, S. Purohit, and B. Chaudhury, “An Application of Machine Learning for Plasma Current Quench Studies via Synthetic Data Generation,” *Fusion Engineering and Design*, vol. 171, no. November 2020, p. 112578, 2021, doi: 10.1016/j.fusengdes.2021.112578.
[43] H. R. Bin Hasbollah and D. Baldry, “A theoretical framework for conserving cultural values of heritage buildings in Malaysia from the perspective of facilities management,” in *Research Methodology in the Built Environment: A Selection of Case Studies*, V. Ahmed, A. Opoku, and Z. Aziz, Eds. Routledge, 2016, pp. 95–106.
[44] T. Carneiro, R. V. M. Da Nobrega, T. Nepomuceno, G. Bin Bian, V. H. C. De Albuquerque, and P. P. R. Filho, “Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications,” *IEEE Access*, vol. 6, pp. 61677–61685, 2018, doi: 10.1109/ACCESS.2018.2874767.
[45] R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, M. F. Chiang, and J. Peter Campbell, “Introduction to machine learning, neural networks, and deep learning,” *Translational Vision Science and Technology*, vol. 9, no. 2, pp. 1–12, 2020, doi: 10.1167/tvst.9.2.14.
[46] J. L. McClelland and D. E. Rumelhart, “Explorations in Parallel Distributed Processing,” *MIT Press*. 1988.
[47] S. I. Gallant, “Perceptron-Based Learning Algorithms,” *IEEE Transactions on Neural Networks*, 1990, doi: 10.1109/72.80230.
[48] A. Almalaq and J. J. Zhang, “Evolutionary Deep Learning-Based Energy Consumption Prediction for Buildings,” *IEEE Access*, vol. 7, no. c, pp. 1520–1531, 2019, doi: 10.1109/ACCESS.2018.2887023.
[49] S. Shalev-Shwartz and S. Ben-David, *Understanding machine learning: From theory to algorithms*. 2013.
[50] S. Haykin, “Neural networks: a comprehensive foundation by Simon Haykin,” *The Knowledge Engineering Review*. 1999.
[51] A. Vijayvargia, *Machine Learning with Python: Design and Develop Machine Learning and Deep Learning Technique using real world code examples*. BPB Publications, 2018.
[52] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006, doi: 10.1016/j.ijforecast.2006.03.001.
[53] J. Fürnkranz *et al.*, “Mean Squared Error,” in *Encyclopedia of Machine Learning*, Springer, Boston, MA, 2011, p. 653.
[54] A. Alwisy, S. BuHamdan, and M. Gül, “Criteria-based ranking of green building design factors according to leading rating systems,” *Energy and Buildings*, vol. 178. pp. 347–359, 2018, doi: 10.1016/j.enbuild.2018.08.043.
[55] T. R. Biyanto, “Analisis Window to Wall Ratio Terhadap Kenyamanan Termal dan Pencahayaan Pada Ruang Kerja (Window to Wall Ratio Analysis on Thermal Comfort and Lighting in the Workspace),” *Kern: Jurnal Ilmiah Teknik Sipil*, vol. 4, no. 1, pp. 45–54, 2014, [Online]. Available: http://ejournal.upnjatim.ac.id/index.php/kern/article/view/1669.

BIOGRAPHIES OF AUTHORS

Mustika Sari is a Ph.D. researcher at the Department of Civil Engineering, Universitas Indonesia. She previously obtained her master's degree in Project Management in 2013 and bachelor's degree in Architecture in 2008. Her primary research interests include building technology, infrastructure development and construction management, emphasizing innovations for economic and sustainable development. She joined the Center for Sustainable Infrastructure Development (CSID), Faculty of Engineering, Universitas Indonesia, after finishing her master's degree and has involved in many national and international research activities. She can be contacted at email: mustika.sari01@ui.ac.id

Mohammed Ali Berawi is a professor at the Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia. He has extensive research experience in value engineering and innovation in building, construction, infrastructure, and manufacturing. He has been awarded as the World's Top 2% Scientists 2021 From ASEAN Countries by Stanford University (2021), Top 500 Indonesian Best Researchers by the Ministry of Research and Technology of the Republic of Indonesia (2020), and listed by Webometrics as one of the Top Scientists in Indonesia (2015-2017). He currently serves as Executive Director of the Center for Sustainable Infrastructure Development (CSID) Universitas Indonesia and Director of the Association of Southeast Asian Nations (ASEAN) University Network for Sustainable City and Urban Development (AUN–SCUD). He can be contacted at email: maberawi@eng.ui.ac.id

Teuku Yuri Zagloel is a professor at the Department of Industrial Engineering, Faculty of Engineering, Universitas Indonesia. He started his higher education in Mechanical Engineering at the University of Indonesia and graduated in 1987. Then, he continued his study at the University of New South Wales and obtained his master's degree (M.Eng.Sc.) in 1991. In 2000, he completed his education path by finishing his doctoral degree at the Universitas Indonesia. His research interests are in the field of Quality Management and Production Systems. He can be contacted at email: yuri@ie.ui.ac.id

Rizka Wulan Triadji received a Bachelor of Engineering (B.Eng.) in Civil Engineering from Universitas Trisakti in 2016 and received a Master of Engineering (M.Eng.) in Civil Engineering from Universitas Indonesia in 2021. Her research includes machine learning, green building, and building performance. She can be contacted at email: rizkawtriadji@gmail.com