Supplementary file

1. Alternative models for assessing effectiveness of BIM implementation in green buildings

1.1 AHP method

In this study, the criteria listed in Table 2 and five-Likert scores obtained from field survey were regarded as factor set and assessment set, respecitvely, which can be expressed as $U = \{U_1, U_2, ..., U_m\}$ and $V = \{V_1, V_2, ..., V_n\}$. To obtained more accurate results, this study employed fuzzy evaluation matrix to match the membership between factors and the corresponding grade by interviewing experts. As a consequence, we can obtain normalized fuzzy evaluation set $R = (r_{ij})_{m\times n}$, where m is the number of factors and n is the dimension of assessment types. This study adopted the combinational weight method to obtain the weight of each factor. More specifically, this study firstly determined the relative importance of different factors by establishing the judgement matrix ($P = (p_{ij})_{m\times m}$) derived from the pairwise comparison. On this basis, the eigenvector A extracted from the factor set can be calculated as:

$$A = [A_1, A_2, \cdots, A_m]^T \text{ (S1)}$$

Where $A_i = \frac{\overline{A_i}}{\sum_{i=1}^{m} \overline{A_i}}$, $\overline{A_i} = M_i^{1/m}$, and $M_i = \prod_{j}^{m} p_{ij}$.

Furthermore, the entropy of the sample data can be defined as:

$$H_i = -k \sum_{j=1}^{n} f_{ij} \ln f_{ij}, i = 1, 2, \cdots, m \text{ (S2)}$$

Where $f_{ij} = \frac{r_{ij}}{\sum_{i=1}^{n} r_{ij}}$ and $k = \frac{1}{\ln n}$.

As a result, the entropy weight is equal to:

$$w_i = \frac{(1 - H_i)}{\left(m - \sum_{i=1}^{m} H_i\right)} \text{ (S3)}$$

Then according to the combinational weight method, the weight vector for factor U_i can be expressed as:

$$t_{ij} = (a_{ij} w_{ij}) \Big/ \left(\sum_{j=1}^{m_i} a_{ij} w_{ij}\right) \text{ (S4)}$$

Finally, the assessment results can be calculated as:

$B = ATR$ (S5)

It's worth noting that AHP still needs additional external data to determine the indicator weights, which may suffer from variations in the experimental condition when comparing with other methods. Therefore, the results of AHP can be only used as a reference for verifying the validity and reliability of developed CNN method.

1.2 BP neural network model

According to the criteria index in this study, there were three first first-level indicators, 15 second-level indicators, and 53 third-level indicators. Therefore, the number of nodes in the input layer of neural network is 53. To determine the number of nodes in the hidden layer, this study used the fomular as below:

$$q = a + \sqrt{n + p} \quad \text{(S6)}$$

Where q is the number of nodes in the hidden layer, n is the number of nodes in the input layer, which is equal to 53 in this study, p is the number of nodes in the output layer, and a is a constant ranging from 1 to 10.

The aim of BP neural network model is to predict the application value of BIM implementation in green buildings. Therefore, the output layer includes five nodes, which is consistent with the Five-Point Likert scale (See Table S1).

Table S1 Outputs from BP neural network model

Grade	Point	Output vector
Poor	1	(1,0,0,0,0)
Fair	2	(0,10,0,0)
Average	3	(0,0,1,0,0)
Good	4	(0,0,0,1,0)
Excellent	5	(0,0,0,0,1)

This study adotped the Sigmoid function to predict the results at different layers. Here we have the predicted results in the hidden layer:

$$y_j = 1 \Big/ 1 + \exp(\textstyle\sum_i w_{ij} x_i + b_i) \quad \text{(S7)}$$

Similarly, the vector derived from the output layer can be expressed as:

$$z_k = 1 \Big/ 1 + \exp(\textstyle\sum_j w_{jk} y_j + b_k) \quad \text{(S8)}$$

Where x_i denotes the input vector in the input layer, w_{ij} and w_{jk} is the weights of the corresponding nodes in the ith line and jth column and jth line and kth column. b_i and b_k are the maximum values of the hidden layer and output layer.

After setting above parameters, the developed mode should be further improved by continuously adjusting the weight vector until the error between the predictions and the actual data below an acceptable level. The modification coefficient from the output layer to the hidden layer can be expressed as:

$$\Delta w_{ij} = -\eta \frac{\partial \varepsilon}{\partial w_{ij}} \text{ (S11)}$$

Similarly, the modification coefficient from the hidden layer to the input layer can be expressed as:

$$\Delta w_{ji} = -\eta \frac{\partial \varepsilon}{\partial w_{ji}} \text{ (S12)}$$

Where ε is the error and η is the learning rate of BP neural network, which indicates the convergent rate of neural network. In this study, η is equal to 0.01.

1.3 SVM model

The input data of SVM are the assessment results obtained from field survey. For each valid questionnaire, the vector S reports the scores of all the 53 factors and overall performance of BIM implementation, which can be expressed as $S = \{x_{1,} x_2, ..., x_{53}, y\}_{54}$. All the input data were evenly divided into training set and testing set. To conduct SVM analysis, it is of necessity to select appropriate parameters, such as error cost (C) and Kernel function parameter (g), which is beneficial for obtaining the results with high accuracy. The overestimation of C can cause over-learning, leading to a high training accuracy but low testing accuracy with a low generalization capability, and vice versa. Consequently, the Kernel cross validation (K-CV) method is selected to obtain the best estimation of parameter C and g, which aims to ensure the accuracy of predictions by avoiding over-learning or under-learning issues. In this study, C=1.5157 and g=0.25 are demonstrated to get the smallest error, which are selected for further analysis.

Furthermore, this study adopts a four-level vector machine classifier to identify the application value of BIM. The Radial basis function (RBF), which is a nonlinear mapping function capable of mapping data into high-dimensional feature space, is adopted as Kernel function of SVM model. The RBF can be expressed as:

$$\Phi(x) = \exp(-\|x - x'\| / 2\sigma^2)$$

Where $\|x - x'\|$ measures the distance between a specific node and the center, $2\sigma^2$ is the reciprocal of parameter g.

Subsequently, the predicted results can be calculated as:

$$f(x) = w\Phi(x) + b$$

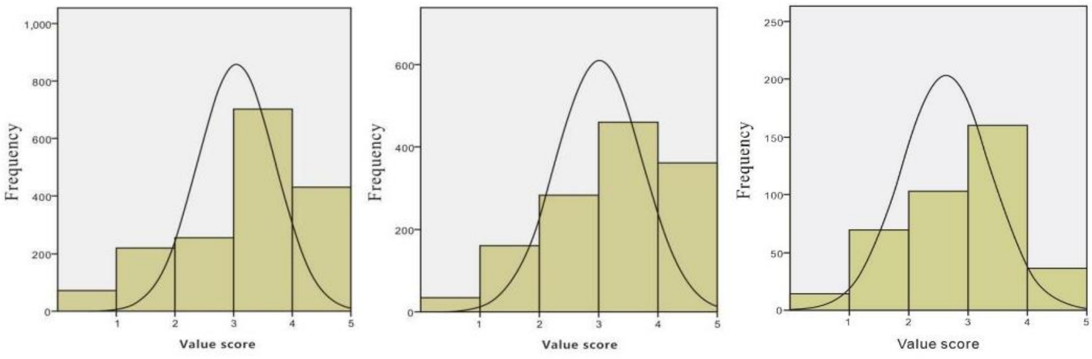
Where w is the weight between hidden layer and output layer, b is a constant. To determine the accuracy of predicted results, this study employed a loss function to measure the disparities:

$$L(f(x), y, \theta) = \begin{cases} 0 & , |y - f(x)| \le \theta \\ |y - f(x)| - \theta & , |y - f(x)| > \theta \end{cases}$$

Where $f(x)$ is the predicted results, y is the actual value, and θ indicates the threshold of differences between predicted and actual value. When the difference is lower than θ, then the loss is equal to zero.

2. The descriptive analysis of sample data

Figure S1 represents the value distribution of the total sample set, training set, and testing set. It can be found that most of value scores were distributed at the four and five points.



(a) Total sample set (b) Training set (c) Testing set

Figure S1 Value distribution of the total questionnaire, training set, and testing set

3. Model validation

3.1 Accuracy analysis

Figure S2 shows the results of accuracy analysis, which reprsents the comparative prediction results of different models. It can be observed that regarding the model accuracy, the CNN model outperformed for value prediction with the accuracy rate at 99.4%, followed by BP neural network (96.4%), SVM model (95.2%), and AHP-based assessment (91.2%).

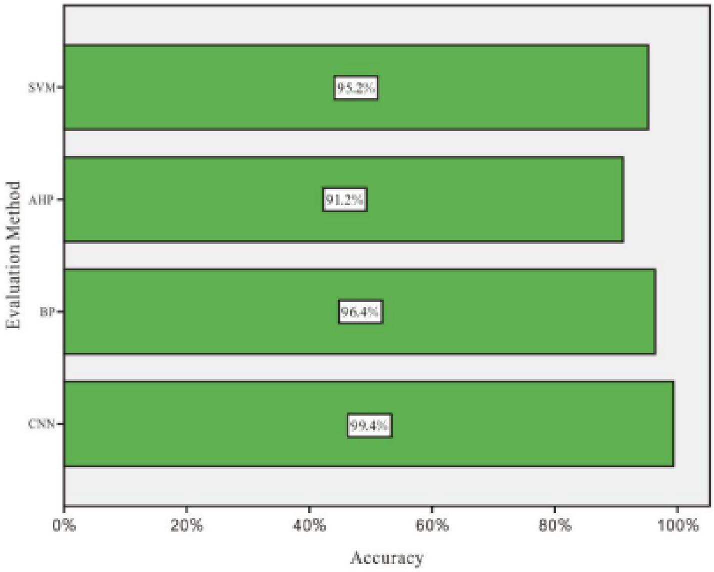


Figure S2 Comparative analysis of different models

(Note: Please note that BP and SVM are analyzed with the same training set, which is comparable with the results of CNN method. AHP is analyzed by including additional data to determine the indicator weights, thus suffering from limitations in comparability.)

3.2 Consistency analysis

The results of consistency analysis indicate that there is a relatively high consistency among different assessment results with the Kappa coefficient at 72.6%. Moreover, according to Figure S3, it is worth noting that the developed CNN model performed more confidence and higher accuracy when the target project featured with high effectiveness of BIM application.

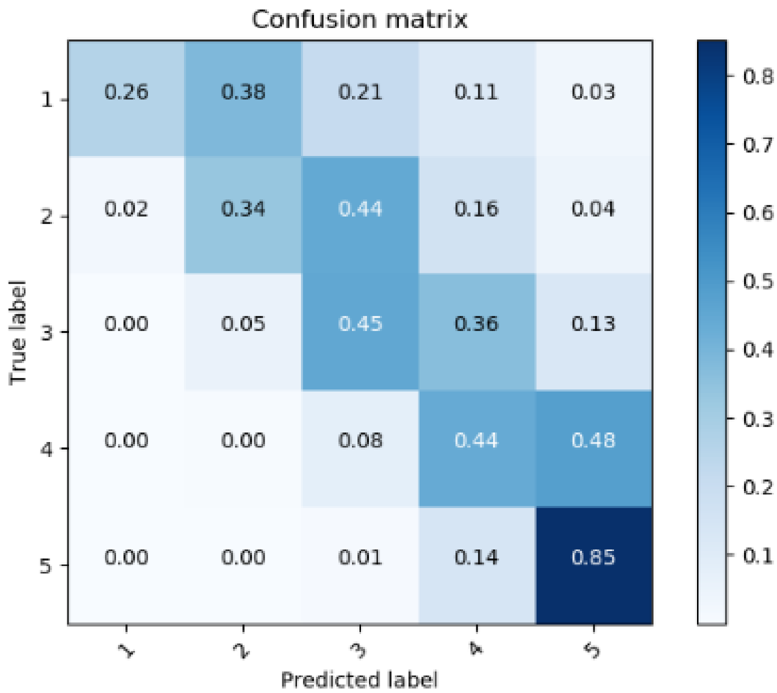


Figure S3 The confusion matrix

3.3 Reliability analysis

Table S1 summarizes the results of precision, recall, and F1 score. The results of reliability analysis revealed that the developed CNN model obtained both high precision, recall, and F_1 score for each predicted score point, indicating that the developed CNN model has no bias for the results.

Table S1 A summary of precision, recall, and F1 score of different values

	Precision	Recall	F1
Score 1	1	1	1
Score 2	0.990	1	0.995
Score 3	0.995	0.995	0.995
Score 4	1	0.976	0.988
Score 5	1	1	1