ORIGINAL PAPER

Ensemble learning models to predict the compressive strength of geopolymer concrete: a comparative study for geopolymer composition design

Qiong Tian[1] · Zhanlin Su[2] · Nicholas Fiorentini[3] · Ji Zhou[1] · Hao Luo[4] · Yijun Lu[4] · Xingquan Xu[5] ·
Chupeng Chen[4] · Jiandong Huang[4]

Abstract

Currently, single machine learning models are mostly used for predicting the compressive strength of geopolymer concrete, but the use of single models has limitations. This study proposes the use of an integrated model to predict the compressive strength of geopolymer concrete. However, there are few applications of ensemble learning model and lack of model optimization. In this study, an improved beetle antennae search (IBAS) algorithm was proposed to tune the hyperparameters of decision tree (DT). random forest (RF), and K-nearest neighbor (KNN) models to predict the compressive strength of geo-polymer concrete. The focus of this paper is to compare the reliability and efficiency of IBAS algorithm applied to three integrated learning models for the prediction of geopolymer concrete compressive strength. The test results show that the corresponding R values are 0.9043, 0.6866, 0.9024, respectively. Therefore, it can be judged that the DT-IBAS integrated model has the worst prediction effect in these three models. In addition, the minimum RMSE values obtained by RF-IBAS and KNN-IBAS models in the ten-fold verification were 5.9 and 7.1, respectively. Therefore, RF-IBAS has the best predictive performance in comparison. On the other hand, the molar concentration of NaOH is the most important factor affecting the compressive strength of geopolymer concrete. Through the importance score test, the importance score of NaOH molar concentration (4.2981) far exceeds that of other input variables. Therefore, it is necessary to focus on the molar concentration of NaOH when making geopolymer concrete.

✉ Chupeng Chen
1112216034@e.gzhu.edu

✉ Jiandong Huang
jiandong.huang@hotmail.com

Qiong Tian
hnkjxytq@huse.edu.cn

Zhanlin Su
18853861527@163.com

Nicholas Fiorentini
nicholas.fiorentini@ing.unipi.it

Ji Zhou
hnkjxyzj@huse.edu.cn

Hao Luo
hao.luo@e.gzhu.edu.cn

Yijun Lu
yijun.lu@e.gzhu.edu.cn

Xingquan Xu
xuxinquan998@126.com

[1] College of Civil and Environmental Engineering, Hunan University of Science and Engineering, Yongzhou 425006, Hunan, People's Republic of China
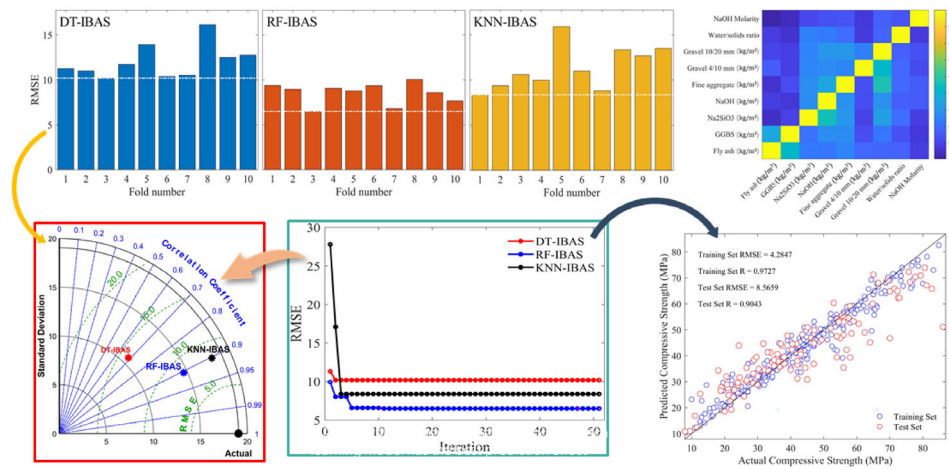
[2] Shandong Energy Group, Xinwen Mining Co., Ltd., Taian 271219, China

[3] Department of Civil and Industrial Engineering (DICI), University of Pisa, 56122 Pisa, Italy

[4] School of Civil Engineering, Guangzhou University, Guangzhou, China

[5] Guangdong Hualu Transport Technology Co., Ltd, Guangzhou, China

Graphical abstract

Keywords Ensemble learning model · Beetle antennae search · Geopolymer concrete · NaOH molarity

Abbreviations

BAS	Beetle antennae search
IBAS	Improved beetle antennae search
KNN	K-nearest neighbor
DT	Decision tree
RF	Random forest
ANN	Artificial neural network
MNLR	The multivariate nonlinear regression
MLR	The multi-linear regression
GEP	Genetic programming
GGBS	Ground granulated blast-furnace slag
RMSE	Root mean squared error
R	Correlation coefficient
SI	Scatter Index

1 Introduction

1.1 Background

Geopolymer concrete is a type of concrete that uses geopolymer material as a gelling agent (Lloyd and Rangan 2010; Verma 2023). Compared with traditional cement-based concrete, it is different in material composition and preparation process. The gelling agent of geopolymer concrete is mainly composed of geopolymer materials instead of conventional cement. Geopolymer is a gelling material produced by the reaction of alkaline active substances such as silicates and aluminates with silicic and aluminic acids (Lavanya and Jegan 2015; Sharma et al. 2023). The gelled material can condense and harden at room temperature to form a solid structure with strength and durability. Compared to traditional cement concrete, the preparation of geopolymer concrete does not require the use of a large amount of cement, thus reducing the energy consumption and carbon dioxide emissions required for cement production (Cao et al. 2018; Nguyen et al. 2023). At the same time, geopolymer concrete can use waste, industrial by-products, and other resources as raw materials, reducing the demand for natural resources, and having good environmental sustainability. A three-dimensional polymer network structure is formed in the condensation process of geopolymer concrete, which has good strength and durability. It usually shows high compressive strength, tensile strength, and permeability resistance, which can meet the requirements of engineering structures (Mehta and Siddique 2018). Geopolymer concrete has good corrosion resistance to chemical substances such as acid, alkali, and salt. This makes it have better adaptability and durability under some special environmental conditions, such as acidic soil and seawater environment (Meng et al. 2019). Because less cement is used in the preparation of geopolymer concrete, its carbon footprint is relatively low. This has positive implications for reducing the carbon footprint of the construction industry and combating climate change (Ahmad et al. 2021).

The compressive strength of geopolymer concrete can be adjusted and controlled according to the specific ratio and preparation process. The compressive strength of geopolymer concrete can usually reach the level of ordinary concrete or higher, depending on several factors, including the type of geopolymer, activity, gelling agent content, curing time, etc. (Ahmed et al. 2021; Rahman and Al-Ameri 2021). Under

Table 1 Common methods to predict the compressive strength of geopolymer concrete

Prediction methods	Detailed approaches	Features
Empirical model (Garces et al. 2022; Le and Bui 2021; Sudhir et al. 2022)	Based on known data and experimental results, the compressive strength can be estimated using empirical formulas. Based on statistical data and empirical rules, these formulas can give approximate compressive strength estimates (Sudhir et al. 2022; Özbayrak et al. 2023; Jonbi and Fulazzaky 2020)	These models are often appropriate for specific conditions and materials and may not be accurate in other situations (Sudhir et al. 2022; Rai et al. 2018; Bellum et al. 2019)
Statistical model (Rai et al. 2018; Bellum et al. 2019; Dolamary et al. 2018; Ahmed et al. 2022a)	By collecting and analyzing test data from a large number of polymer concrete samples, statistical models can be constructed to predict compressive strength. Common methods include regression analysis, artificial neural (Huang et al. 2022a) networks, support vector machine, and so on (Rai et al. 2018; Dolamary et al. 2018)	These models can take into account multiple influencing factors, such as material composition, curing conditions, etc., and predict the numerical range of compressive strength (Lavanya and Jegan 2015; Ali et al. 2022; Kishore et al. 2022)
Physical models and numerical simulations (Cao et al. 2018; Chen et al. 2022; Veerapandian et al. 2022; Zhang et al. 2021)	By using a physical model and numerical simulation method, the compressive strength of geopolymer concrete can be predicted by establishing a mechanical property model and corresponding calculation method (Meng et al. 2019; Zhang et al. 2021; Colangelo et al. 2013)	This method can consider more material properties and influencing factors, but it needs corresponding experimental data and model validation (Colangelo et al. 2013)
Machine learning algorithm (Ahmad et al. 2021; Rahmati and Toufigh 2022; Faraj et al. 2022a; Ahmed et al. 2022b)	Using machine learning algorithms, such as decision trees (Nguyen et al. 2023), random forest (Verma 2023), neural networks (Bhogayata et al. 2021), etc., the model can be trained to predict the compressive strength of geopolymer concrete by the input of known material properties and other relevant factors (Rahman and Al-Ameri 2021, 2022)	This method can learn patterns and correlations from large amounts of data, providing relatively accurate predictions (Sharma et al. 2023; Gupta and Rao 2022)

normal circumstances, the compressive strength of polymer concrete can be in the range of 20–100 MPa. This range is relatively wide, and the specific compressive strength depends on several factors, such as the characteristics of the geopolymer material used, the selection and ratio of aggregates, and the curing temperature (Ali et al. 2022; Chen et al. 2022). Predicting the compressive strength of geopolymer concrete is a complex task, which requires consideration of many factors and estimation with appropriate models or methods (Garces et al. 2022). Table 1 shows some common methods.

Among these methods, machine learning models can process and analyze large amounts of data in a highly automated manner, thereby quickly and accurately predicting the performance of geopolymer concrete, saving the time and cost of manual testing (Gupta and Rao 2022; Ahmed et al. 1868, 2022c). At the same time, it considers multiple factors that affect the performance of geopolymer concrete, such as material composition, process parameters, environmental conditions, etc., to provide more comprehensive prediction results (Rahman and Al-Ameri 2022; Ahmed et al. 2022d;

Ghafor et al. 2022). Machine learning can be flexible in dealing with nonlinear relationships, because the performance of geopolymer concrete is often affected by the complex interaction of multiple factors (Choudhary and Gianey 2017; Huang et al. 2022b). Machine learning models can capture these nonlinear relationships and provide more accurate predictions. Not only that, but machine learning models can improve their predictive performance through continuous training and optimization (Grazzi et al. 2020; Faraj et al. 2022b). As the amount of data and the complexity of the model increase, the predictive power of the model can be continuously improved (Ahmed et al. 2023a, 2022e). However, it should be noted that the performance of the machine learning model depends heavily on the training data available. If the data quality is poor or insufficient, the predictive power of the model may be limited. Also, if a machine learning model is too complex or has insufficient training data, overfitting problems can occur, in which the model performs well on training data but has poor generalization ability on new data. Gupta et al. employed the artificial neural network

(ANN), multiple linear regression, and the multivariate non-linear regression (MNLR) models to predict the compressive strength of the geopolymer concrete. The results showed that the performance of the ANN model was better than the multi linear regression (MLR) and multi non-linear regression (MNLR) models (Gupta and Rao 2022). Awoyera et al. used genetic programming (GEP) and ANN techniques to predict the strength properties of geopolymer self-compacting concrete. It is confirmed that both GEP and ANN methods exhibited good prediction of the experimental data, with minimal errors (Awoyera et al. 2020). Tanyildizi tried to predict the geopolymerization process of fly ash-based geopolymer using deep long short-term memory and machine learning (Tanyildizi 2021). Ayar Mazumder et al. developed the gene-based expression programming (GBEP) model to predict the mechanical properties of self-compacting geopolymer concrete and the developed model can predict the experimental data results very precisely and with very few errors (Chai and Draxler 2014). Rizvon et al. combined random forest models with Artificial Neural networks and the Least absolute shrinkage and selection operator (LASSO). To predict the compressive strength and physical properties of eco-friendly, cement-based materials (Rizvon and Jayakumar 2021a, b, 2023).

Compared with the traditional model, the machine learning model does have greater advantages in predicting the compressive strength of geopolymer concrete (Kakasor Ismael Jaf et al. 2023; Unis Ahmed et al. 2023). However, it should be noted that due to the complexity of the machine learning model itself and the size of the data set, the overly sensitive performance of a single machine learning model in the prediction process may lead to overfitting or underfitting problems (Ahmed et al. 2023b; Huang et al. 2021; Zhu et al. 2022). Overfitting means that the model can predict the data too well in the training stage, but has poor generalization ability to the unencountered data. Underfitting refers to the insufficient fitting effect of the model on the training data and the inability to find the hidden internal relationship in the data. Although the results obtained from the examples mentioned above meet the requirements of experimental prediction accuracy in terms of data. However, single machine learning models are often trained and predicted based on specific data sets, and if the data distribution changes, the model's prediction effect will be affected (Huang et al. 2022c). On the other hand, a single machine learning model can usually only be used to deal with a specific problem or a specific data type, and it is necessary to find a suitable machine learning model for different problems. But choosing the right model often requires domain knowledge and experience. The wrong model choice can result in performance degradation that fails to meet problem solving requirements (Huang and Xue 2022; Huang et al. 2022d).

Therefore, in order to solve the above problems, it is necessary to use the integrated model to predict the compressive strength of geopolymer concrete.

In order to solve the limitation of single machine learning model on the compressive strength of geopolymer concrete. In this study, three ensemble learning models are adopted, and an improved BAS (IBAS) algorithm is proposed to optimize the hyperparameters of the other three models to form an ensemble model. And focus on the comparison of the IBAS algorithm cited in decision tree (DT), random forest (RF), and K-nearest neighbor (KNN) models to build the ensemble learning models (Huang et al. 2022e). By comparing the predicted data, a more suitable prediction model for the compressive strength of geopolymer concrete is further selected, which is an important step for the manufacture and use of geopolymer concrete. The importance of the influencing factors in geopolymer concrete is analyzed by using the prediction model, which can provide reference for the efficient design of geopolymer concrete in the future. In future research, it is meaningful to directly establish the visualization software of the output variables. However, this paper mainly discusses the prediction effect of different integrated models on the compressive strength of geopolymer concrete, and selects the prediction model with the best prediction effect. In future studies, if the results of this study can be adopted, it is intended to build visualization software that can output different results by changing the input variables.

1.2 Significance of the study

The purpose of this study is to explore the prediction accuracy of geopolymer concrete compressive strength and find out the influence degree of influencing factors, so as to provide new ideas and methods for solving this problem. This study is of great significance for promoting the development of geopolymer concrete. Compared with previous studies, the innovation of this study lies in the adoption of ensemble learning model in prediction, which makes up for the deficiency of prediction generalization ability of single machine learning model. The integrated machine learning model proposed in this study provides a reference for the future design of geopolymer concrete by predicting the compressive strength of geopolymer concrete.

2 Methodology

2.1 Determination of the input variables

The compressive strength of geopolymer concrete is affected by many factors, including cementing materials (such as fly ash, metallurgical slag, etc.), aggregates, and activators. The

ratio of raw materials directly affects the chemical composition and physical properties of concrete, thus affecting the compressive strength. The appropriate ratio of raw materials can improve the compressive strength of geopolymer concrete. The activator of geopolymer concrete is used to activate the silicates in the gelling material and promote the gelling reaction. The type and dosage of activator have a significant influence on the compressive strength of concrete. Different activators have different activation mechanisms and reaction effects. Appropriate selection and adjustment of the type and dosage of activators can improve compressive strength. The properties of cementitious materials play an important role in the compressive strength of geopolymer concrete. For example, the silicate content of fly ash and the activity index of metallurgical slag affect the reactivity and strength development of gelling materials. The compressive strength of geopolymer concrete can be improved by selecting suitable cementitious material and treating and activating it. Aggregate is a particle filler in geopolymer concrete, which has a certain influence on the mechanical properties of concrete. The physical and mechanical properties of aggregates (such as particle size distribution, shape, surface properties, etc.) affect the interaction between particles and the internal structure of the concrete, thus affecting the compressive strength. Based on the above analysis, this study selected the contents of fly ash (Lloyd and Rangan 2010) and ground granulated blast-furnace slag (GGBS) (Mehta and Siddique 2018) as the influencing factors of the cementing materials, Na_2SiO_3 (Mohseni 2018), NaOH (Mohseni 2018), and NaOH molarity (Ahmed et al. 2021) as the activator design parameters, the contents of the fine aggregate, gravel (4/10 mm), gravel (10/20 mm) (Ahmed et al. 2021) and water/solids ratio as the concrete design parameters to predict the compressive strength of geopolymer concrete. Table 2 summarizes the determination of the input variables.

2.2 Experimental procedure

Raw materials used in the experiment mainly include fly ash, ground granulated blastfurnace slag (GGBS), fine aggregate, alkali activator (Na_2SiO_3, NaOH), stone, sand and water, etc. To systematically study the influence of various influencing factors on the mechanical properties of geopolymer concrete and collect data for the training of mechanical learning models. A variety of experimental blocks in different situations are prepared with different mix ratios to improve the prediction accuracy of the machine learning model.

The main preparation process of geopolymer concrete is: firstly, raw materials such as fly ash, GGBS, fine aggregate, alkali exciter (Na_2SiO_3, NaOH), stone, sand and water are introduced into the mixer for mixing in turn, which is stirred slowly for 180 s at first, then stopped stirring for 20 s, and then stirred quickly for 120 s. Finally, the stirred slurry is

Table 2 Determination of the input variables

Design considerations	Input variables
Cementing materials	Contents of fly ash (Bellum et al. 2019; Huang et al. 2022f) and ground granulated blast-furnace slag (GGBS) (Mehta and Siddique 2018; Huang et al. 2022f)
Activator design parameters	Contents of Na_2SiO_3 (Mohseni 2018) and NaOH (Mohseni 2018); Molarity of NaOH (Mohseni 2018)
Concrete design parameters	Contents of the fine aggregate, gravel (4/10 mm), gravel (10/20 mm) and water/solids ratio (Lavanya and Jegan 2015; Rahman and Al-Ameri 2021; Sudhir et al. 2022)

poured into the sample mold. It was placed on the shaking table for 60 s. After the slurry was formed, it was demoulded and placed in a standard curing box for 7 days. The prismatic specimen of 150 mm × 150 mm × 300 mm was used for the axial compressive strength test.

2.3 Data collection and analysis

Based on the determination of the input variables from Table 2, this research collected the dataset from the previous studies including the design parameters and the compressive strength of the geopolymer concrete (Zou et al. 2022). Table 3 gives the detail of the input variables.

The proportion of training and testing data accounts for 70% (147) and 30% (63) of the total data set, respectively. The curing time of geopolymer concrete test block in this experiment adopts the standard of 7 days. Therefore, the curing time is not used as a reference factor in the model prediction. To test the independence of these input variables, the correlation between these input variables was tested by using the method of determining the Pearson correlation coefficient. Figure 1 gives the results.

Correlation coefficients can help identify collinearity (linear correlation) between variables. In machine learning models, multicollinearity can lead to instability and decreased explanatory power of the model. As can be seen from the figure, most variables have correlation coefficients below 0.4, which indicates that the linear correlation between them is weak and, therefore, does not lead to instability in the machine learning prediction model. It should be noted that the correlation coefficient only measures the linear correlation between variables and ignores the nonlinear relationship.

Table 3 Detail of the input variables

	Fly ash (kg/m³)	GGBS (kg/m³)	Na$_2$SiO$_3$ (kg/m³)	NaOH (kg/m³)	Fine aggregate (kg/m³)	Gravel 4/10 mm (kg/m³)	Gravel 10/20 mm (kg/m³)	Water/solids ratio	NaOH molarity	Compressive strength (MPa)
Average value	257.65	127.93	106.97	54.15	698.64	420.65	718.58	0.32	10.62	48.72
Standard deviation	90.38	83.76	38.13	17.08	43.33	125.86	119.64	0.10	2.61	17.58
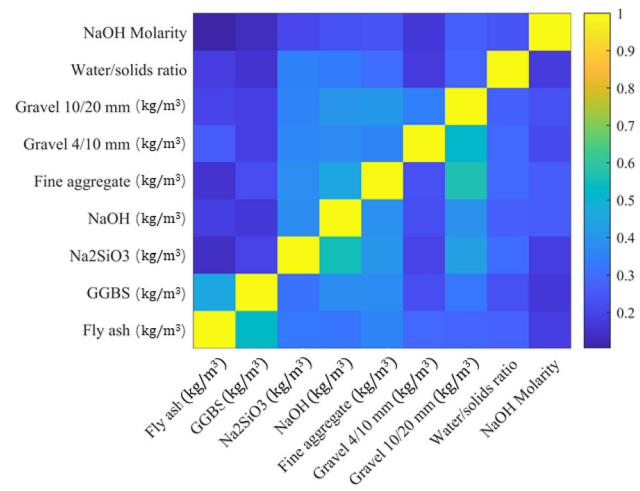
Fig. 1 Correlation coefficient of the input parameters

In some cases, there may be complex nonlinear relationships between variables, and the correlation coefficient may not accurately capture this relationship. However, the results obtained so far showed that the databases selected in this study are independent of each other in terms of input variables, and these input variables are acceptable for machine learning in this study.

2.4 Ensemble learning models used in the study

Decision tree (DT) is a commonly used machine learning algorithm that can be used for classification and regression tasks (Song and Ying 2015; Myles et al. 2004). Decision trees classify or predict data by constructing a tree-structured decision process. The tree structure of the decision tree consists of nodes and edges. Each internal node represents a feature or attribute, while each leaf node represents a category or predicted value (Myles et al. 2004). By dividing the feature layer by layer, the decision tree can assign the sample to different categories or predicted values according to the value of the feature; the RF model is an ensemble learning method based on the decision tree (Cutler et al. 2012; Belgiu and Drăguţ 2016). Random forests (RF) perform classification and regression tasks by building multiple decision trees at the same time and integrating their predictions (Wang et al. 2021). The model takes samples randomly from the original training data to form several different training subsets. A decision tree model is constructed using a decision tree algorithm. Each decision tree is independently generated from a different subset of training (Cutler et al. 2012; Belgiu and Drăguţ 2016). For regression predictions, the random forest averages the predictions of each decision tree; The K-nearest neighbor (KNN) model is a commonly used machine learning algorithm for classification and regression problems. The core idea of the KNN model is to compare the

classification or prediction results of a new sample with its nearest neighbor training sample based on similarity measures. For a new sample, the KNN model calculates the distance between it and each sample in the training set (Guo et al. 2003). Common distance measures include Euclidean distance, Manhattan distance, etc. According to the results of distance measurement, K training samples closest to the new sample are selected as the nearest neighbors. For regression prediction, the KNN model averages the predicted values of the nearest neighbors as the prediction results of the new sample (Guo et al. 2003; Huang et al. 2022f, Huang et al. 2022g).

2.5 Improved Beetle antennae search (IBAS) algorithm

The Beetle antennae search (BAS) (Wang and Chen 1807; Huang et al. 2022g) is a heuristic optimization algorithm based on the search behavior of the beetle antennae in nature. The algorithm simulates the feedback signal generated when the beetle antennae touches the environment and optimizes the solution by adjusting the search direction and distance. The BAS algorithm will initialize a certain number of solution vectors as the initial population according to the characteristic dimension of the problem, and then calculate the fitness value of each solution vector, that is, the performance of the objective function on the solution vector (Wang and Chen 1807). For each solution vector, according to its fitness value and neighborhood information, the solution vector is updated by adjusting the search direction and distance, and the global optimal solution is recorded. Determines whether to end the search based on a set termination condition, such as reaching a maximum number of iterations or meeting a specific stop criterion (Huang et al. 2020, 2021).

For traditional BAS, the beetle's step size remains the same or decreases with each iteration. Adopting this step adjustment strategy can cause some problems. If the given step size is too small, the BAS algorithm may converge slowly or fall into a local optimal state (Zhang et al. 2019). However, if the given step size is very large, global optimality may be skipped and the result may oscillate. Therefore, Levy flight and self-inertia weights are used in this study to adjust the step size of BAS (Huang et al. 2022g). To improve the search efficiency, this paper named the improved BAS the improved beetle antennae search (IBAS) algorithm. It can quickly adjust the step size according to the current fitness value, and reduce the oscillation by using the adaptive weight; also, using Levy flying, randomly expand the step size when the BAS falls into a local optimal state (Huang et al. 2021).

In the implementation of this study, since the BAS algorithm is in the local optimal state, the following formula is

triggered to increase the beetle step size:

$$\delta^{(i)} = \alpha |\text{Levy}| \otimes \delta^{(i-1)} \tag{1}$$

where α is the randomization parameter; $\otimes$ means term-by-term multiplication; Levy is a Levy distribution with infinite variance, where infinity is expressed as Levy $\sim u = t^{-\lambda}$, $(1 < \lambda \leq 3)$. Trigger Levy flight as,

$$\left| f^{(i)} - f^{(i-1)} \right| < \mu(f_w - f_b), \tag{2}$$

where μ represents the coefficient; f_w and f_b are the worst fitness value and the best fitness value, respectively. In the implementation of this study, the adaptive inertia weight adopts a monotone reduction equation, which is described as follows:

$$\delta^{i+1} = \eta^i \times \delta^i, \tag{3}$$

where δ^i is the step size of the current position; η^i represents the adaptive inertia weight, given by:

$$\eta^i = (1-\alpha)0.95 + \alpha \frac{f_w^i - f^i}{f_w^i - f_b^i}, \tag{4}$$

where f^i represents the fitting equation of the current position; f_b^i represents the best fitting value; f_w^i represents the worst fitting value; α represents the hyperparameter to trade-off between the two items (Huang et al. 2020).

2.6 Evaluation of predictive performance

In this study, n-fold cross-validation will be used to evaluate the model performance (Malhotra and Meena 2021). It is a commonly used model evaluation method to evaluate the performance and generalization ability of machine learning models. It divides the original data set into n subsets of equal size, where n-1 subsets are used as the training set, and the remaining subset is used as the validation set. This process is repeated n times, each time using a different subset as the validation set, and the average of the n evaluations is used as the performance indicator of the model. To improve the reliability of model comparison, 10-fold cross-validation was selected in this study. For quantitative comparison indicators, this study determined the standard deviation (Lee et al. 2015), root mean squared error (RMSE) (Chai and Draxler 2014), and correlation coefficient (R) (Benesty et al. 2008) for the model comparison to predict the compressive strength of the geopolymer concrete.
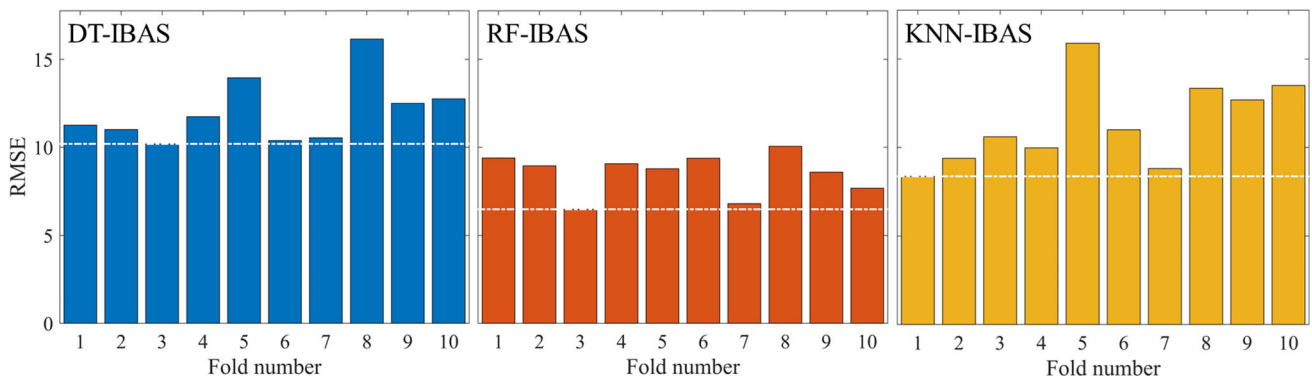
Fig. 2 10-fold cross-validation results

3 Analysis of results

3.1 Results of the tenfold cross-validation

Tenfold cross-validation evaluates a model's ability to generalize, that is, how it performs on previously unseen data. By evaluating different validation sets, the generalization performance of the model can be better understood, and overfitting or underfitting problems can be avoided (Huang et al. 2022f, h). Figure 2 shows the tenfold cross-validation results. It can be seen RF-IBAS model can earn the minimum RMSE value for the tenfold, indicating that the model has a lower probability of overfitting or underfitting. At the same time, it can be seen that compared with the other two models, the generalization ability of RF-IBAS is also strong, which can be proved by the relatively stable RMSE value obtained by the tenfold. KNN-IBAS can also obtain an ideal RMSE value, but it seems to have less generalization than DT-IBAS (see the relative stability of RMSE values of DT-IBAS and KNN-IBAS models).

3.2 RMSE values for increasing iteration times

Figure 3 gives the RMSE values for DT-IBAS, RF-IBAS, and KNN-IBAS models. As can be seen from the figure, for the three ensemble learning models, under the hyperparameter adjustment of the IBAS algorithm, RMSE usually decreases rapidly with the increase of iterations in the early stage of model training. This is because the model is learning and gradually fitting the patterns and relationships in the training data. As the number of iterations increases further, the model may reach a stage of plateau or slow decline. At this point, the performance of the model may already be close to the local optimal, and further iterations may result in only minor performance improvements. For the DT-IBAS model, it can be observed that the RMSE value decreased slightly with the increase in the number of iterations and remained unchanged. This showed that the IBAS algorithm
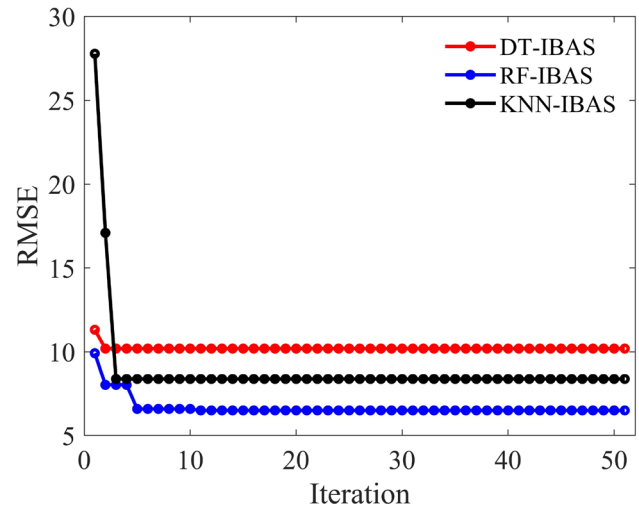

Fig. 3 RMSE values for DT-IBAS, RF-IBAS, and KNN-IBAS models

cannot effectively adjust the parameters during training so that the RMSE remains unchanged. Another possible reason is that DT-IBAS models are not sufficiently expressive to capture complex patterns and relationships in the data. In this case, even if the number of iterations is increased, the model does not improve the prediction performance. For RF-IBAS and KNN-IBAS models, the iteration will be adjusted in the direction of the gradient of the loss function. This iterative process gradually reduces the value of the loss function, thereby reducing the RMSE.

## 3.3 Predictive results of the DT-IBAS, RF-IBAS, and KNN-IBAS models

Table 4 shows the hyperparameters of the developed DT-IBAS, RF-IBAS, and KNN-IBAS models, including the initial values and the IBAS suggested values for the hyperparameters.

Figure 4 gives the results of the predicted compressive strength and actual compressive strength for the DT-IBAS,

Table 4 Obtained hyperparameters of the DT-IBAS, RF-IBAS, and KNN-IBAS models

| Developed models	Hyperparameters			
DT-IBAS	Number of samples contained by a leaf node		Number of samples contained in the intermediate node	
	Initial value	Suggested value by IBAS	Initial value	Suggested value by IBAS
	25	45	50	90
RF-IBAS	Number of trees in the forest		Number of leaf node samples per tree	
	Initial value	Suggested value by IBAS	Initial value	Suggested value by IBAS
	40	19	40	1
KNN-IBAS	Number of nearest neighbor samples considered for each sample			
	Initial value		Suggested value by IBAS	
	30		1	
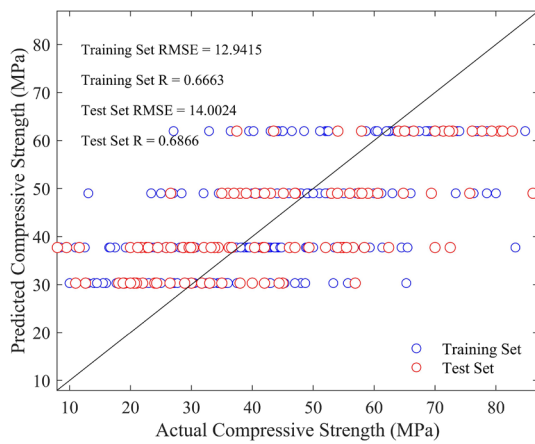
a) DT-IBAS
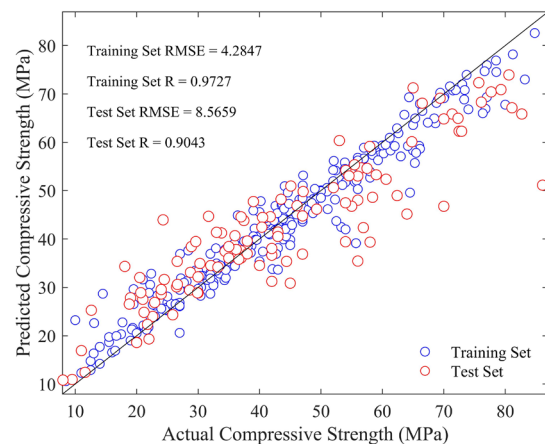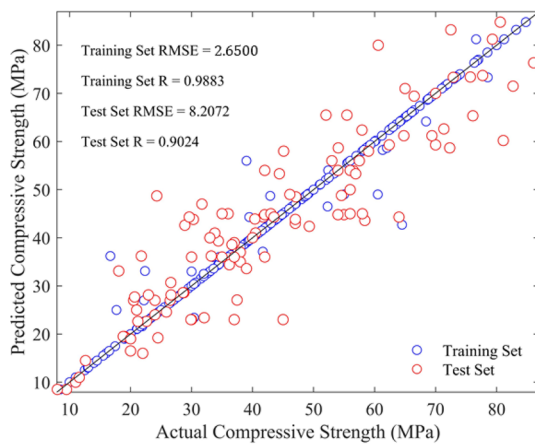
b) KNNIBAS

c) RF-IBAS

Fig. 4 Predicted and actual compressive strength of the geopolymer concrete for the developed models

RF-IBAS, and KNN-IBAS models. It can be seen from the figure that the good prediction effect of KNN-IBAS and RF-IBAS, which can be seen from the predicted compressive strength and the actual compressive strength close to the "1:1" curve. Specifically, KNN-IBAS achieves RMSE values of 2.65 and 8.2072 for the training set and the test set, respectively; RF-IBAS achieves RMSE values of 4.2847 and 8.5659 for the training set and the test set, respectively. Such low RMSE values indicate that the developed model has a small prediction error and can better predict the compressive strength of geopolymer concrete. At the same time, comparing the distribution of the predicted and actual compressive strength of the two models, it can be found that the test set results of KNN-IBAS are more symmetric in the "1:1" curve. However, the prediction results of RF-IBAS are larger in smaller compressive strength regions and smaller in larger compressive strength regions. This suggests that the RF-IBAS model may overestimate the predicted compressive strength of the geopolymer concrete with low compressive strength and underestimate the predicted compressive strength of geopolymer concrete with high compressive strength. For the DT-IBAS model, it can be clearly seen that the regression fitting effect is poor due to the low learning rate, and the loss function target is almost not reached. Such a prediction model is difficult to be adopted in the strength modeling of geopolymer concrete.

3.4 Models' comparison

Figure 5 shows the model comparison of the three developed using the form of a Taylor diagram, indicating the values of the RMSE, standard deviation, and correlation coefficient.

The Taylor diagram shows the difference between the predictions of multiple models and the observed data in the form of a circular coordinate system. Each point in the plot represents a model, where the position of each point represents the model's standard deviation (the difference from the observed data) and its correlation coefficient (the correlation with the observed data). In the Taylor diagram, the better model will be located closer to the observed data, with a small standard deviation and a high correlation coefficient. This allows us to intuitively compare the performance between different models and understand how they differ in terms of predictive power and correlation (Table 5). It can be seen from the figure that similar to the previous results, RF-IBAS and KNN-IBAS achieved better results in terms of prediction accuracy and reliability. Specifically, RF-IBAS and KNN-IBAS have similar performance in terms of the correlation coefficient. KNN-IBAS has a slight advantage in RMSE compared with RF-IBAS, while RF-IBAS has a better performance in standard deviation than KNN-IBAS; DT-IBAS has a small standard deviation, but its performance in correlation coefficient and RMSE value is unsatisfactory.
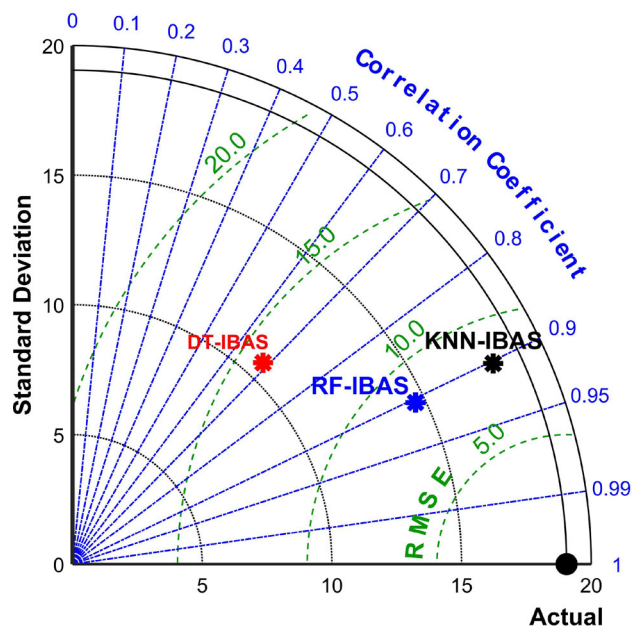
Fig. 5 Comparison from the perspectives of the correlation coefficient, standard deviation, and RMSE values

The above table shows the evaluation index of the three models on the compressive strength prediction results of geopolymer concrete after data training. It can be clearly seen from the table above that the KNN-IBAS model has the lowest RMSE, Scatter Index(SI) and OBJ values and the highest R values. Therefore, it can be shown that KNN-IBAS model has the best prediction ability in the prediction of geopolymer concrete compressive strength.

3.5 Importance analysis and sensitivity analysis of input variables

Figure 6 gives the importance score of these influential variables determining the compressive strength of geopolymer concrete. Importance Score is an indicator used to evaluate the importance of parameters in the calculation, and the value obtained by it can measure the contribution of each feature to the prediction result of the model. Through the importance score of each feature, we can know which features have the greatest impact on the predictive performance of the model, so that we can better understand and interpret the prediction results of the model. It can be clearly seen from the figure that NaOH molarity is the most important factor affecting the compressive strength of geopolymer concrete and its influence far outweighs other possible factors. This is because a higher concentration of sodium hydroxide solution can be used as an activator to promote polymer reactions in geopolymer concrete. These polymer reactions lead to the formation of polymer networks in the concrete and increase the compressive strength of the concrete. However, too high

Table 5 Forecast the result of the DT-IBAS, RF-IBAS, and KNN-IBAS models

	Training set RMSE	Test set RMSE	Training set R	Test set R	SI	OBJ
DT-IBAS	12.9415	14.0024	0.6663	0.6866	0.32189	12.6021
KNN-IBAS	2.65	8.2072	0.9883	0.9024	0.18867	7.63269
RF-IBAS	4.2847	8.5659	0.9727	0.9043	0.19691	7.88062
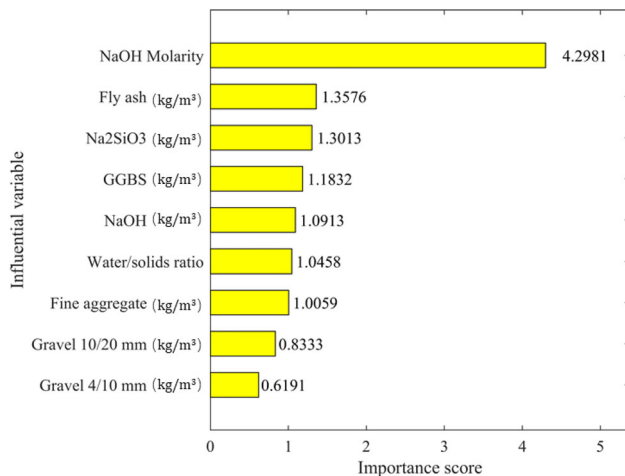
Fig. 6 Importance analysis of input variables for the compressive strength

NaOH molarity may cause the reaction to be too fast or violent, which adversely affects the performance of the concrete. Therefore, such polymer reaction is significant for the formation of the compressive strength of geopolymer concrete. Similar to NaOH, $NaSi_2O_3$ can also be used as an activator to promote polymer reactions in geopolymer concrete. Its main role in concrete is to form hydrated silicate gels by reacting with calcium ions in water, which helps to strengthen the structure of concrete and improve compressive strength. This is the reason why its content has achieved an importance score of 1.3013. However, it should be noted that the importance of NaOH molarity is far greater than that of the NaOH content in forming the compressive strength of geopolymer concrete, indicated by the importance score of 1.0913 for the NaOH content. For those concrete design parameters (including water/solids ratio, fine aggregate, gravel 10/20 mm, and gravel 4/10 mm), the importance is less than that of the polymer cement design parameters, This should be paid enough attention in the design of geopolymer concrete in the future.

4 Conclusion

Geopolymer concrete has been developed to contribute to sustainable construction, but predicting the compressive strength of geopolymer concrete is a complex task, which requires consideration of many factors and estimation with appropriate models or methods. This study employed three ensemble learning models and proposed an improved BAS algorithm to predict the compressive strength of geopolymer concrete. Also, this study focused on comparing the reliability and efficiency effects of this improved BAS algorithm when applied to the three ensemble learning models. Through this comparative study, it can build an important step for the establishment of the geopolymer concrete compressive strength prediction model. Using the established model, the importance of these factors affecting the compressive strength of geopolymer concrete is also analyzed. The following are the conclusions that can be obtained.

1. The developed RF-IBAS model can earn the minimum RMSE value for the tenfold and the model has a lower probability of overfitting or underfitting. Compared with the other two models, the generalization ability of RF-IBAS is strong, which can be proved by the relatively stable RMSE value obtained by the tenfold. KNN-IBAS can obtain an ideal RMSE value but show less generalization than DT-IBAS. Therefore, RF-IBAS model has the best prediction effect on the compressive strength of geopolymer concrete.

2. Under the hyperparameter tuning of IBAS algorithm, the RMSE of the three-medium integrated model proposed in this study usually decreases rapidly with the increase of iterations in the early stage of model training. This suggests that IBAS algorithm has effective hyperparameter tuning capability for these three integrated models. As the number of iterations increases further, the model may reach a stage of plateau or slow decline, indicating the performance of the model may already be close to local optimal, and further iterations may result in only minor performance improvements.

3. KNN-IBAS and RF-IBAS showed good prediction effects, indicated by the fact that the predicted compressive strength and the actual compressive strength are close to the "1:1" curve. Specifically, KNN-IBAS achieves RMSE values of 2.6500 and 8.2072 for the training set and the test set, respectively; RF-IBAS achieves RMSE values of 4.2847 and 8.5659 for the training set and the test set, respectively. Such low RMSE values indicate that the developed model has a small prediction error and can better predict the compressive strength of geopolymer concrete.

4. NaOH molarity is the most important factor affecting the compressive strength of geopolymer concrete and its influence far outweighs other possible factors. Similar to NaOH, $NaSi_2O_3$ can also be used as an activator to promote polymer reactions in geopolymer concrete. The importance of NaOH molarity is far greater than that of the NaOH content in forming the compressive strength of geopolymer concrete. Those concrete design parameters are less important than the polymer cement design parameters to form the compressive strength of the geopolymer concrete.

Although the model presented in this study has certain advantages and prediction ability in the prediction of geopolymer concrete compressive capacity, it also has some limitations. First, these three models are trained based on specific data sets and experimental conditions, and therefore, may not be applicable to all types of cement-based concrete composites. Second, it may not be possible to accurately predict the properties of concrete composites in different strength ranges. This is because concrete composites with different strength ranges may have different compositions and structures, thus affecting the predictive power of the model. Nevertheless, because the model is an integrated model, it has certain universality and extensibility. By properly adjusting the parameters and input characteristics of the model, the application range of the model can be extended to adapt to different types and strength ranges of concrete composites. It is hoped that the launch of the model will promote the development of geopolymer concrete, and further improve the use of geopolymer concrete in the construction field, so as to implement the policy of environmental protection.

In future research, it is meaningful to directly establish the visualization software of the output variables. However, this paper mainly discusses the prediction effect of different integrated models on the compressive strength of geopolymer concrete, and selects the prediction model with the best prediction effect. In future studies, if the results of this study can be adopted, it is intended to build visualization software that can output different results by changing the input variables.

Author contributions QT, ZS, NF and JZ wrote the main manuscript text, XX, CC and JH prepared figures. All authors reviewed the manuscript.

Data availability statement The data are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest in this work.

References

Ahmad A, Ahmad W, Chaiyasarn K, Ostrowski KA, Aslam F, Zajdel P, Joyklad P (2021) Prediction of geopolymer concrete compressive strength using novel machine learning algorithms. Polymers 13:3389

Ahmed HU, Abdalla AA, Mohammed AS, Mohammed AA, Mosavi A (1868) Statistical methods for modeling the compressive strength of geopolymer mortar. Materials 2022:15

Ahmed HU, Mohammed AA, Rafiq S, Mohammed AS, Mosavi A, Sor NH, Qaidi S (2021) Compressive strength of sustainable geopolymer concrete composites: a state-of-the-art review. Sustainability 13:13502

Ahmed HU, Abdalla AA, Mohammed AS, Mohammed AA (2022a) Mathematical modeling techniques to predict the compressive strength of high-strength concrete incorporated metakaolin with multiple mix proportions. Clean Mater 5:100132. https://doi.org/10.1016/j.clema.2022.100132

Ahmed HU, Mohammed AS, Mohammed AA (2022b) Proposing several model techniques including ANN and M5P-tree to predict the compressive strength of geopolymer concretes incorporated with nano-silica. Environ Sci Pollut Res 29:71232–71256. https://doi.org/10.1007/s11356-022-20863-1

Ahmed HU, Mohammed AS, Faraj RH, Qaidi SMA, Mohammed AA (2022c) Compressive strength of geopolymer concrete modified with nano-silica: experimental and modeling investigations. Case Stud Constr Mater 16:e01036. https://doi.org/10.1016/j.cscm.2022.e01036

Ahmed HU, Mohammed AA, Mohammed A (2022d) Soft computing models to predict the compressive strength of GGBS/FA- geopolymer concrete. PLoS ONE 17:e0265846. https://doi.org/10.1371/journal.pone.0265846

Ahmed HU, Mohammed AS, Mohammed AA (2022e) Multivariable models including artificial neural network and M5P-tree to forecast the stress at the failure of alkali-activated concrete at ambient curing condition and various mixture proportions. Neural Comput Appl 34:17853–17876. https://doi.org/10.1007/s00521-022-07427-7

Ahmed HU, Mostafa RR, Mohammed A, Sihag P, Qadir A (2023a) Support vector regression (SVR) and grey wolf optimization (GWO) to predict the compressive strength of GGBFS-based geopolymer concrete. Neural Comput Appl 35:2909–2926. https://doi.org/10.1007/s00521-022-07724-1

Ahmed HU, Mohammed AS, Faraj RH, Abdalla AA, Qaidi SMA, Sor NH, Mohammed AA (2023b) Innovative modeling techniques including MEP, ANN and FQ to forecast the compressive strength of geopolymer concrete modified with nanoparticles. Neural Comput Appl 35:12453–12479. https://doi.org/10.1007/s00521-023-08378-3

Ali AA, Al-Attar TS, Abbas WA (2022) A statistical model to predict the strength development of geopolymer concrete based on SiO_2/Al_2O_3 ratio variation. Civ Eng J 8:454–471

Awoyera PO, Kirgiz MS, Viloria A, Ovallos-Gazabon D (2020) Estimating strength properties of geopolymer self-compacting concrete using machine learning techniques. J Mark Res 9:9016–9028

Belgiu M, Drăguţ L (2016) Random forest in remote sensing: a review of applications and future directions. ISPRS J Photogramm Remote Sens 114:24–31

Bellum RR, Muniraj K, Madduru SRC (2019) Empirical relationships on mechanical properties of class-F fly ash and GGBS based geopolymer concrete. Ann Chim-Sci Matér 43:189–197

Benesty J, Chen J, Huang Y (2008) On the importance of the Pearson correlation coefficient in noise reduction. IEEE Trans Audio Speech Lang Process 16:757–765

Bhogayata A, Kakadiya S, Makwana R (2021) Neural network for mixture design optimization of geopolymer concrete. ACI Mater J 118

Cao VD, Pilehvar S, Salas-Bringas C, Szczotok AM, Bui TQ, Carmona M, Rodriguez JF, Kjøniksen A-L (2018) Thermal performance and numerical simulation of geopolymer concrete containing different types of thermoregulating materials for passive building applications. Energy Build 173:678–688

Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)?–arguments against avoiding RMSE in the literature. Geosci Model Dev 7:1247–1250

Chen C, Zhang X, Hao H, Cui J (2022) Discussion on the suitability of dynamic constitutive models for prediction of geopolymer concrete structural responses under blast and impact loading. Int J Impact Eng 160:104064

Choudhary R, Gianey HK (2017) Comprehensive review on supervised machine learning algorithms. In: Proceedings of the 2017 international conference on machine learning and data science (MLDS), pp 37–43

Colangelo F, De Luca G, Ferone C, Mauro A (2013) Experimental and numerical analysis of thermal and hygrometric characteristics of building structures employing recycled plastic aggregates and geopolymer concrete. Energies 6:6077–6101

Cutler A, Cutler DR, Stevens JR (2012) Random forests. Ensemble machine learning: methods and applications, pp 157–175

Dolamary PY, Dilshad J, Arbili MM, Karpuzcu M (2018) Validation of feret regression model for fly ash based geopolymer concrete. Polytech J 8:173–189

Faraj RH, Mohammed AA, Omer KM, Ahmed HU (2022a) Soft computing techniques to predict the compressive strength of green self-compacting concrete incorporating recycled plastic aggregates and industrial waste ashes. Clean Technol Environ Policy 24:2253–2281. https://doi.org/10.1007/s10098-022-02318-w

Faraj RH, Ahmed HU, Rafiq S, Sor NH, Ibrahim DF, Qaidi SMA (2022b) Performance of self-compacting mortars modified with nanoparticles: a systematic review and modeling. Clean Mater 4:100086. https://doi.org/10.1016/j.clema.2022.100086

Garces JIT, Beltran AB, Tan RR, Ongpeng JMC, Promentilla MAB (2022) Carbon footprint of self-healing geopolymer concrete with variable mix model. Clean Chem Eng 2:100027

Ghafor K, Ahmed HU, Faraj RH, Mohammed AS, Kurda R, Qadir WS, Mahmood W, Abdalla AA (2022) Computing models to predict the compressive strength of engineered cementitious composites (ECC) at various mix proportions. Sustainability 14:12876

Grazzi R, Franceschi L, Pontil M, Salzo S (2020) On the iteration complexity of hypergradient computation. In: Proceedings of the international conference on machine learning, pp 3748–3758

Guo G, Wang H, Bell D, Bi Y, Greer K (2003) KNN model-based approach in classification. In: Proceedings of the on the move to meaningful internet systems 2003: CoopIS, DOA, and ODBASE: OTM confederated international conferences, CoopIS, DOA, and ODBASE 2003, Catania, November 3–7, 2003. Proceedings, pp 986–996

Gupta T, Rao MC (2022) Prediction of compressive strength of geopolymer concrete using machine learning techniques. Struct Concr 23:3073–3090

Huang J, Xue J (2022) Optimization of svr functions for flyrock evaluation in mine blasting operations. Environ Earth Sci 81:434

Huang J, Duan T, Zhang Y, Liu J, Zhang J, Lei Y (2020) Predicting the permeability of pervious concrete based on the beetle antennae search algorithm and random forest model. Adv Civ Eng 2020:8863181. https://doi.org/10.1155/2020/8863181

Huang J, Kumar GS, Ren J, Zhang J, Sun Y (2021) Accurately predicting dynamic modulus of asphalt mixtures in low-temperature regions using hybrid artificial intelligence model. Constr Build Mater 297:123655. https://doi.org/10.1016/j.conbuildmat.2021.123655

Huang J, Zhou M, Zhang J, Ren J, Vatin NI, Sabri MMS (2022a) Development of a new stacking model to evaluate the strength parameters of concrete samples in laboratory. Iran J Sci Technol Trans Civ Eng 46:4355–4370. https://doi.org/10.1007/s40996-022-00912-y

Huang J, Zhang J, Gao Y (2022b) Evaluating the clogging behavior of pervious concrete (PC) using the machine learning techniques. In: CMES-computer modeling in engineering & sciences, p 130

Huang J, Zhang J, Li X, Qiao Y, Zhang R, Kumar GS (2022c) Investigating the effects of ensemble and weight optimization approaches on neural networks' performance to estimate the dynamic modulus of asphalt concrete. Road Mater Pavement Design 1–21

Huang J, Zhou M, Yuan H, Sabri MMS, Li X (2022d) Prediction of the compressive strength for cement-based materials with metakaolin based on the hybrid machine learning method. Materials 15:3500

Huang J, Sabri MM, Ulrikh DV, Ahmad M, Alsaffar KA (2022e) Predicting the compressive strength of the cement-fly ash–slag ternary concrete using the firefly algorithm (FA) and random forest (RF) hybrid machine-learning method. Materials 15:4193. https://doi.org/10.3390/ma15124193

Huang J, Zhou M, Yuan H, Sabri MM, Li X (2022f) Towards sustainable construction materials: a comparative study of prediction models for green concrete with metakaolin. Buildings 12:772. https://doi.org/10.3390/buildings12060772

Huang J, Zhou M, Sabri MMS, Yuan H (2022g) A novel neural computing model applied to estimate the dynamic modulus (dm) of asphalt mixtures by the improved beetle antennae search. Sustainability 14:5938

Huang J, Zhou M, Zhang J, Ren J, Vatin NI, Sabri MMS (2022h) The use of GA and PSO in evaluating the shear strength of steel fiber reinforced concrete beams. KSCE J Civ Eng 26:3918–3931

Jonbi J, Fulazzaky MA (2020) Modeling the water absorption and compressive strength of geopolymer paving block: an empirical approach. Measurement 158:107695

Kakasor Ismael Jaf D, Abdulrahman AS, Abdulrahman PI, Salih Mohammed A, Kurda R, Ahmed HU, Faraj RH (2023) Effitioned soft computing models to evaluate the impact of silicon dioxide (SiO_2) to calcium oxide (CaO) ratio in fly ash on the compressive strength of concrete. J Build Eng 74:106820. https://doi.org/10.1016/j.jobe.2023.106820

Kishore Y, Nadimpalli SGD, Potnuru AK, Vemuri J, Khan MA (2022) Statistical analysis of sustainable geopolymer concrete. Mater Today Proc 61:212–223

Lavanya G, Jegan J (2015) Evaluation of relationship between split tensile strength and compressive strength for geopolymer concrete of varying grades and molarity. Int J Appl Eng Res 10:35523–35527

Le H-B, Bui Q-B (2022) Predicting the compressive strength of geopolymer concrete: an empirical model for both recycled and natural aggregates. In: Proceedings of the CIGOS 2021, emerging technologies and applications for green infrastructure: proceedings of the 6th international conference on geotechnics, civil engineering and structures, pp 793–802

Lee DK, In J, Lee S (2015) Standard deviation and standard error of the mean. Korean J Anesthesiol 68:220–223

Lloyd N, Rangan V (2010) Geopolymer concrete with fly ash. In: Proceedings of the second international conference on sustainable construction materials and technologies, pp 1493–1504

Malhotra R, Meena S (2021) Empirical validation of cross-version and 10-fold cross-validation for defect prediction. In: Proceedings of the 2021 second international conference on electronics and sustainable communication systems (ICESC), pp 431–438

Mehta A, Siddique R (2018) Sustainable geopolymer concrete using ground granulated blast furnace slag and rice husk ash: Strength and permeability properties. J Clean Prod 205:49–57

Meng Q, Wu C, Su Y, Li J, Liu J, Pang J (2019) Experimental and numerical investigation of blast resistant capacity of high performance geopolymer concrete panels. Compos B Eng 171:9–19

Mohseni E (2018) Assessment of Na_2SiO_3 to NaOH ratio impact on the performance of polypropylene fiber-reinforced geopolymer composites. Constr Build Mater 186:904–911

Myles AJ, Feudale RN, Liu Y, Woody NA, Brown SD (2004) An introduction to decision tree modeling. J Chemom 18:275–285

Nguyen MH, Mai H-VT, Trinh SH, Ly H-B (2023) A comparative assessment of tree-based predictive models to estimate geopolymer concrete compressive strength. Neural Comput Appl 35:6569–6588

Özbayrak A, Kucukgoncu H, Atas O, Aslanbay HH, Aslanbay YG, Altun F (2023) Determination of stress-strain relationship based on alkali activator ratios in geopolymer concretes and development of empirical formulations. In: Proceedings of the structures, pp 2048–2061

Rahman SK, Al-Ameri R (2021) Experimental investigation and artificial neural network based prediction of bond strength in self-compacting geopolymer concrete reinforced with basalt FRP bars. Appl Sci 11:4889

Rahman SK, Al-Ameri R (2022) Experimental and artificial neural network-based study on the sorptivity characteristics of geopolymer concrete with recycled cementitious materials and basalt fibres. Recycling 7:55

Rahmati M, Toufigh V (2022) Evaluation of geopolymer concrete at high temperatures: an experimental study using machine learning. J Clean Prod 372:133608

Rai B, Roy L, Rajjak M (2018) A statistical investigation of different parameters influencing compressive strength of fly ash induced geopolymer concrete. Struct Concr 19:1268–1279

Rizvon SS, Jayakumar K (2021a) Machine learning techniques for recycled aggregate concrete strength prediction and its characteristics between the hardened features of concrete. Arab J Geosci 14:2390. https://doi.org/10.1007/s12517-021-08674-z

Rizvon SS, Jayakumar K (2021b) Strength prediction models for recycled aggregate concrete using random forests, ANN and LASSO. J Build Pathol Rehabil 7:5. https://doi.org/10.1007/s41024-021-00145-y

Rizvon SS, Jayakumar K (2023) Strength-maturity correlation models for recycled aggregate concrete using Plowman's coefficient. Arab J Geosci 16:147. https://doi.org/10.1007/s12517-023-11211-9

Sharma U, Gupta N, Verma M (2023) Prediction of the compressive strength of Flyash and GGBS incorporated geopolymer concrete using artificial neural network. Asian J Civ Eng 1–14

Song Y-Y, Ying L (2015) Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry 27:130

Sudhir M, Chen S, Rai S, Jain D (2022) An empirical model for geopolymer reactions involving fly ash and GGBS. Adv Mater Sci Eng 2022:1–13

Tanyildizi H (2021) Predicting the geopolymerization process of fly ash-based geopolymer using deep long short-term memory and machine learning. Cement Concr Compos 123:104177

Unis Ahmed H, Mohammed AS, Mohammed AA (2023) Fresh and mechanical performances of recycled plastic aggregate geopolymer concrete modified with nano-silica: experimental and computational investigation. Constr Build Mater 394:132266. https://doi.org/10.1016/j.conbuildmat.2023.132266

Veerapandian V, Pandulu G, Jayaseelan R, Sathish Kumar V, Murali G, Vatin NI (2022) Numerical modelling of geopolymer concrete in-filled fibre-reinforced polymer composite columns subjected to axial compression loading. Materials 15:3390

Verma M (2023) Prediction of compressive strength of geopolymer concrete using random forest machine and deep learning. Asian J Civ Eng 1–10

Wang J, Chen H (2018) BSAS: beetle swarm antennae search algorithm for optimization problems. arXiv:1807.10470

Wang Q-A, Zhang J, Huang J (2021) Simulation of the compressive strength of cemented tailing backfill through the use of firefly algorithm and random forest model. Shock Vib 2021:1–8

Zhang J, Ma G, Huang Y, Aslani F, Nener B (2019) Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression. Constr Build Mater 210:713–719

Zhang P, Gao Z, Wang J, Wang K (2021) Numerical modeling of rebar-matrix bond behaviors of nano-SiO_2 and PVA fiber reinforced geopolymer composites. Ceram Int 47:11727–11737

Zhu F, Wu X, Zhou M, Sabri MMS, Huang J (2022) Intelligent design of building materials: development of an ai-based method for cement-slag concrete design. Materials 15:3833

Zou Y, Zheng C, Alzahrani AM, Ahmad W, Ahmad A, Mohamed AM, Khallaf R, Elattar S (2022) Evaluation of artificial intelligence methods to estimate the compressive strength of geopolymers. Gels 8:271