Analyzing the mechanical performance of fly ash-based geopolymer concrete with different machine learning techniques

Yiming Peng[1], Cise Unluer[1, *]

[1] School of Engineering, University of Glasgow, Glasgow G12 8LT, United Kingdom

* Corresponding author. E-mail address: Cise.Unluer@glasgow.ac.uk

Abstract: Fly ash (FA)-based geopolymer concrete is considered as an alternative system with potentially lower environmental impact than Portland cement mixes. However, the prediction accuracy of compressive strength still needs to be improved. This study demonstrated the feasibility of predicting the 28-day strength of geopolymer concrete through mix proportions and pre-curing conditions by using three machine learning algorithms (backpropagation neural network (BPNN), support vector machine (SVM) and extreme learning machine (ELM)) and provided a comparison of their differences, highlighting variations in prediction accuracy. As a part of the evaluation of model performance and error analysis, the prediction accuracy differences of these three models in training, validation and testing sets were discussed, and the influence weight of each input parameter on results was analyzed by permutation feature importance (PFI) index. Results showed that all models revealed good prediction performance for the overall database. BPNN model had the largest number of instances where the error percentage was within ±20%. SVM model showed the highest generalization capability and most stable prediction accuracy among all three. Out of different variables investigated, SiO_2 content in FA had the highest influence on strength, followed by Al_2O_3 content and activator content/concentration. These outcomes can enable reductions in experimental time, labor, materials and costs; and facilitate the adoption of alternative binders in the concrete industry.

Keywords: Geopolymers; compressive strength; machine learning; support vector machine; backpropagation neural network; extreme learning machine

1. Introduction

Portland cement (PC), the most frequently used cementitious material for concrete worldwide, is associated with a large energy consumption and heavy carbon dioxide (CO_2) emissions due to the procedures involved in its production. About 4 billion tons of CO_2 are emitted by PC production every year, accounting for 5-7% of total anthropogenic CO_2 emissions [1-4]. Owing to the increasing concerns for environmental protection and climate change, many initiatives have been taken to reduce impacts associated with cement production and use. These include the incorporation of supplementary cementitious materials (SCMs) such as fly ash (FA), limestone powder (LP) and ground granulated blast furnace slag (GGBS) into cement mixes and the recycling of industrial solid wastes to reduce the reliance on PC in new and existing structures [5-9]. However, the proportions of these SCMs used as a replacement for PC are often restricted. Taking FA as an example, although FA demonstrates pozzolanic effects during the different stages of cement hydration, it does not have a major role in strength development in the early stages [10-13]. The addition of FA can reduce the early hydration rate and prolong the setting time of composite pastes [14-16], which can limit its use in large contents.

Obtaining more environmentally friendly cementitious binders by means of alkali activation is one of the most frequently investigated routes for the complete replacement of PC [17]. Alkali-activated materials (AAMs) do not involve the high-temperature and energy intensive calcination process adopted in the production of PC clinker. AAMs such as geopolymers are polymeric aluminosilicate cementitious materials with three-dimensional spatial network structures that involve industrial wastes (e.g. FA) as the main binder under the action of an alkaline activator (e.g. NaOH, Na_2SiO_3) [18-21]. Geopolymers can demonstrate favorable mechanical properties and durability due to their unique chemical structure. From an environmental standpoint, the reuse of industrial wastes as the main binder component makes these systems attractive in comparison to PC-based mixes [22].

FA-based geopolymer concrete not only has comparable mechanical properties to PC-based concrete [23, 24], but also shows improved durability such as high temperature resistance [25], sulfate resistance [26] and chloride penetration resistance [27]. Different compressive strength prediction models for FA-based geopolymer concrete are presented in the literature. Zhang et al. [28] analyzed the residual compressive strength test points of low calcium FA-based geopolymer concrete after exposure to high temperatures. The corresponding coefficients in Gaussian-based mathematical model were calculated in MATLAB to get two prediction equations. The results showed that the experimental strengths were usually larger than the predicted counterparts. Cong et al. [29] established a constitutive model to predict the engineering properties of alkali-activated GGBS/FA concrete according to its stress-strain relationship and elastic modulus under different compression and tension loads. However, the dynamic increase factor obtained from the constitutive model presented in this study was not

consistent with the experimental results. Le et al. [30] used the modified Feret and De Larrard models to predict the compressive strength of FA-based geopolymer recycled aggregate concrete. Compared with Feret's model, which requires a modification that replace the parameter for cement (C) with the parameter for binder (B) to ensure reasonable prediction accuracy, De Larrard's model only needs to determine the parameters of natural and recycled aggregates in advance to get satisfactory prediction results. One of the most significant disadvantages of using constitutive or other empirical models to predict the mechanical properties of concrete is that the expression of function is solidified in the calculation process. Although the regression function itself has many undetermined parameters, the solidification of the function expression still hinders the possibility of further optimizing the prediction accuracy.

The application of artificial intelligence in concrete mix design and performance prediction has gained increased attention in the last few years [31, 32]. The most distinctive advantage of using machine learning as opposed to empirical models in the prediction of the mechanical properties of concrete mixes is its ability to fully consider the nonlinear relationship between independent variables and dependent variables, as well as the influence of various factors on the results. Nonlinearity is the most remarkable characteristic of machine learning (i.e. especially for artificial neural network (ANN)) [33]. Most machine learning algorithms usually divide the database into training set and testing set [34, 35]. If the results of the testing set are not ideal, the algorithm will automatically iterate until the prediction accuracy meets the requirements. In contrast, the empirical model can only iterate once, resulting in an output of fitting parameters, which can explain why simple regression results cannot achieve the desired effect. Some of the most frequently used machine learning models in the analysis of concrete properties include ANN [36-39] and support vector machine (SVM) [40, 41]. These algorithms provide a good support for the optimization of concrete mix design and performance.

Accordingly, the purpose of this study is to demonstrate the feasibility of predicting the strength of geopolymer concrete by using a recently emerging popular algorithm (i.e. extreme learning machine (ELM)) and provide a comparison with two traditional approaches (i.e. backpropagation neural network (BPNN) and support vector machine (SVM)) to highlight their differences and variations in prediction accuracy. Although there are several studies on concrete performance prediction based on machine learning, few studies focus on the prediction of the strength of geopolymer concrete by using different algorithms. Aiming to fill this gap, this paper presents a detailed analysis involving the use of these three algorithms to predict the 28-day compressive strength of geopolymer concrete according to the chemical composition of the main binder component, mix proportions and pre-curing conditions. Moreover, this paper analyzes the influence of each parameter on the mechanical performance and characterizes the influence weight of each feature parameter by using the permutation feature importance (PFI) index and sensitivity analysis, thereby providing a theoretical basis for optimizing the mechanical properties of geopolymer concrete.

2. Machine learning approaches and performance evaluation

2.1 Machine learning models

BPNN and SVM are two commonly used algorithms in neural network prediction. Among them, BPNN belongs to multilayer feedforward neural network and uses the difference between the actual outputs and expected outputs to correct the network parameters of each layer from back to front, thereby iteratively optimizing biases and weights, and maximizing the prediction accuracy (Fig. 1(a)). Alternatively, SVM is considered as a generalized linear classifier for binary data classification based on supervised learning. As shown in Fig. 1(b), the goal of this algorithm is to maximize the distance from the hyperplane to the nearest support vectors of each category. It is worth noting that SVM can be used for both classification and parameter prediction. The method used in this study is support vector regression (SVR).

ELM is an advanced machine learning algorithm proposed by Huang et al. [42]. For single hidden layer neural networks, ELM can initialize the input weights and biases randomly and obtain the corresponding output parameters (Fig. 1(c)). According to its operation principle, it is considered as an improvement of the backpropagation algorithm in BPNN. The weights and biases from the input layer to the hidden layer of ELM are set randomly, and there is no need to make constant reverse adjustments to the weights and biases like BPNN. This results in a significant reduction in the amount of computation. Moreover, the connection weights β_{jk} between the hidden layer and the output layer of ELM are determined by solving the equations at one time. Similar to BPNN, ELM also only needs one hidden layer to meet the requirement of prediction accuracy.

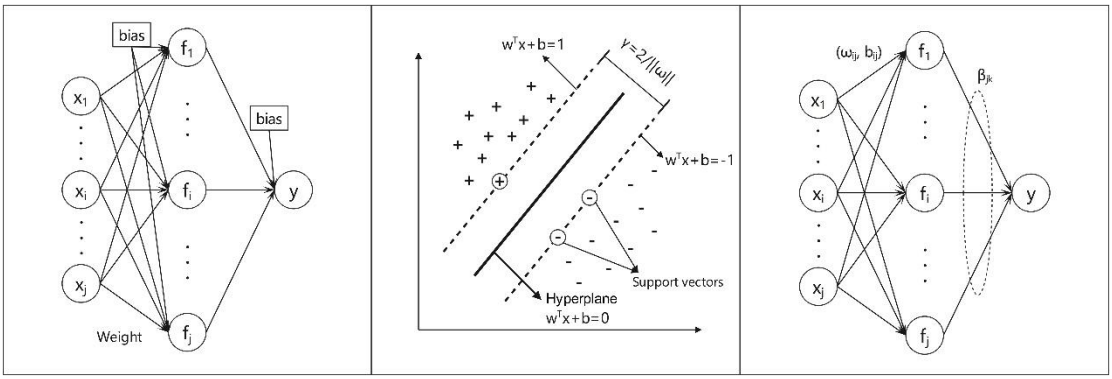


Fig. 1 Graphical representation of different machine learning models (modified from [31] and [8])

2.2 Training process and evaluation of machine learning models

Fig. 2 shows the unified calculation process of three machine learning algorithms used in this study. Firstly, the collected data was arranged in random order. In order to eliminate the possible impact of various characteristic indicators on the results due to their different dimensions, all input parameters were preprocessed by the same normalization algorithm to unify their dimensions, improve the speed of model training and prevent the decline of prediction accuracy or the failure of algorithm convergence (e.g. for BPNN). The mathematical equation used for data normalization is shown in Eq. (1), where y_{max} and y_{min} are 1 and -1 by default, and x_{max} and x_{min} are the maximum and minimum values of input parameters, respectively [43]. After normalization, all input parameters were in the range of -1 to 1. Finally, the prediction data were denormalized to get the final output data. All the algorithms were implemented in MATLAB 2016a platform.

Fig. 2 Construction and data prediction process of machine learning models

$$y = \frac{(y_{max} - y_{min}) \times (x - x_{min})}{(x_{max} - x_{min})} + y_{min} \quad (1)$$

The hyperparameters of BPNN, SVM and ELM used in this study are presented in Table 1. The activation functions of BPNN and ELM adopted the default sigmoid function, which is shown in Eq. (2).

$$s(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

The number of nodes in the hidden layer was determined by the calculation accuracy of validation data. In the training process of BPNN and ELM, results showed that when the hidden layer was set to 18 and 20, respectively, the goodness of fit reached the maximum and the calculation error was the lowest. The kernel function of SVM was RBF function. In order to enhance the generalization ability, the K-fold cross validation method is mainly utilized in SVM due to the presence of several hyperparameters. As BPNN and ELM involve a limited number of hyperparameters, they mainly optimize their hyperparameters according to the prediction outcomes of the training set. All algorithms divided the data into training set, validation set and testing set in line with

the proportion of 70%, 15% and 15%, respectively.

Table 1 Hyperparameters of machine learning models

Machine learning algorithms	Hyperparameters	Values/categories
BPNN	Activation function	Sigmoid
	Hidden layer numbers	1
	Number of hidden layer nodes	18
SVM	SVM type	e-SVR
	Kernel function	RBF
	Tolerance	0.001
	epsilon	0.1
	Shrinking	True
ELM	Hidden layer nodes	20
	Activation function	Sigmoid
	Mode	Regression

In order to evaluate the deviation between the predicted value and measured value for a single sample, error and error percentage were used to characterize their similarity, as shown in Eq. (3) and (4), where y_i' is the predicted value and y_i is the actual value.

$$Error = y'_i - y_i \tag{3}$$

$$Error\ percentage = \frac{y'_i - y_i}{y_i} \tag{4}$$

As the above evaluation criteria were only for a single sample, in order to further quantify the comprehensive prediction results, various statistical methods were used. This led to the evaluation of the performance of given models, including coefficient of determination (R^2), mean square error (*MSE)*, root mean square error (*RMSE*) and mean absolute error (*MAE*). These parameters can not only quantify the prediction accuracy of a single algorithm, but also compare the performance of different models, identifying the most suitable model with the highest applicability for the specific database. The expressions of these evaluation indexes are shown in Eq. (5)-(8).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y'_i - y_i\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{5}$$

$$MSE = \frac{\sum_{i=1}^{n}(y'_i - y_i)^2}{n} \tag{6}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y'_i - y_i)^2}{n}} \tag{7}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y'_i - y_i| \quad (8)$$

3. Data collection

The data used in this study were gathered from 8 different references [44-51], resulting in a total of 110 mix proportions. Table 2 shows a representative mix proportion, which comprehensively considers the influence of FA, coarse and fine aggregates, alkaline activator, water and chemical admixture. Alkaline activators were mainly composed of sodium hydroxide (NaOH) and sodium silicate (Na_2SiO_3). In addition to their dosage, the molar concentrations of these activators used in the solutions were considered. Polycarboxylate superplasticizer (PCE) is commonly used as a chemical admixture in these formulations. Among the relevant studies used for the main source of data, some mixed water and the alkaline activator together before mixing with powder, while others calculated the mass of water separately. In order to obtain a uniform content of water across different mixes, the total mass of water, shown in Table 2, was obtained by adding up the mass of water in NaOH and Na_2SiO_3 solutions and added water.

Table 2 A sample mix proportion of FA-based geopolymer concrete representing those used this study

FA (kg/m^3)	Coarse aggregate (kg/m^3)	Fine aggregate (kg/m^3)	NaOH solution (kg/m^3)	NaOH (M)	Na_2SiO_3 solution (kg/m^3)	Na_2SiO_3/NaOH	AA/FA	Water (kg/m^3)	PCE (kg/m^3)
400	875	875	40	16	100	2.5	0.35	112.4	10.5

Another factor that was considered in the analysis was the variation in the chemical compositions of FA used across different studies (Table 3). In line with the findings of previous studies [22, 52, 53], where the key roles of Si and Al contents in FA on the strength development of FA-based mixes were reported, the analysis presented in this study also incorporated the SiO_2 and Al_2O_3 contents in FA as input parameters. Other than the binder component, aggregate characteristics such as composition and physical properties (e.g. particle size), as shown in Table 4, were also involved in the analysis. Owing to its difficulty of being quantified as a parameter, the influence of aggregate size distribution on strength was not considered in the presented analysis. In addition, the high temperature pre-curing regimes usually employed to accelerate the early hydration rate of FA-based mixes were also considered. The combination of these approaches led to the identification of 14 factors as input variables. However, it should be noted that all the input variables must be independent. While the analysis involves 3

variables that are related to the incorporation of NaOH, they all have an independent impact on the final results. Accordingly, mix proportions can change solely on the content of NaOH solution, while keeping the molar concentration of NaOH constant, indicating that they can be regarded as mutually independent variables [44]. The relatively large range of each parameter, as listed in Table 5, could indicate the applicability of the prediction model.

Table 3 Chemical composition of FA obtained from different sources

Reference		SiO_2	Al_2O_3	CaO	SO_3	Fe_2O_3	MgO	LOI
[44]		71.50	9.20	6.72	2.40	2.37	0.60	3.67
[45]		61.89	28.05	0.87	1.32	4.11	0.38	0.49
[46]		49	31	5	-	3	3	-
[47]		62.30	28.10	0.5	0.40	2.10	1.00	2.50
[48]	Type I	47.87	28.0	3.81	0.27	14.09	0.93	0.43
	Type II	49.37	31.25	4.80	0.24	4.47	1.28	0.51
	Type III	53.82	29.95	1.03	0.34	9.24	0.58	0.63
[49]		53.71	27.20	1.90	0.30	11.17	-	0.68
[50]		45.23	19.95	15.51	-	13.15	-	-
[51]		31.32	13.96	25.79	3.29	15.64	2.94	1.30

LOI = Loss on ignition

Table 4 Particle size parameters of coarse and fine aggregates obtained from different sources

Reference	Coarse aggregates		Fine aggregates		
	Raw materials	Size range (mm)	Raw materials	Size range (mm)	Fineness modulus
[44]	-	9.5-12.5	River sand	-	2.35
[45]	-	6.3-20	Graded river sand	< 4.75	-
[46]	Crushed bluestone gravel	7	Graded sand	< 0.4	-

[47]	Pink limestone	1.18-9.5	Natural siliceous sand	0.075-4.75	2.45
[48]	Crushed basalt aggregate	7-10	Uncrushed river sand	-	3.0
[49]	Crushed granite	7-20	Sand	-	1.97
[50]	Limestone	20	River sand	-	2.90
[51]	Crushed limestone	7-16	River sand	-	2.20

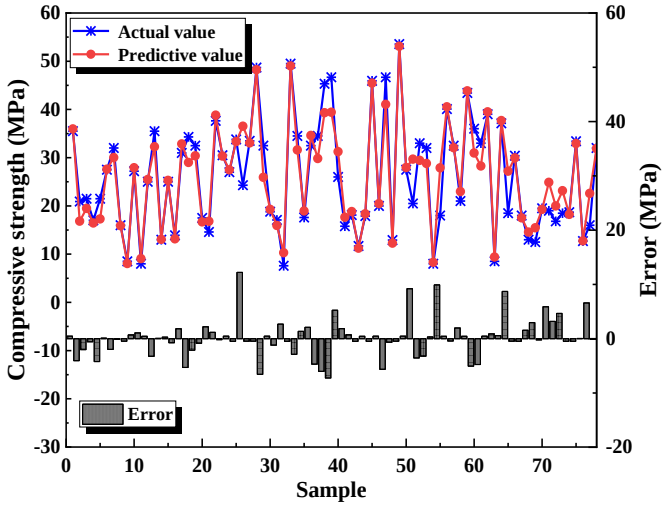
Table 5 Mean, standard deviation and range of input/output parameters

Notation	Input parameter	Mean/count	Standard deviation	Range/categories
X_1	FA (kg/m^3)	396.63	44.29	250-523
X_2	SiO_2 (%)	56.34	15.93	31.32-71.5
X_3	Al_2O_3 (%)	16.55	8.01	9.2-31.25
X_4	Coarse aggregate (kg/m^3)	1196.76	155.86	840-1567
X_5	Fine aggregate (kg/m^3)	582.25	127.59	320-910
X_6	NaOH (kg/m^3)	71.96	28.35	11.78-118
X_7	NaOH (M)	13.32	3.01	8-20
X_8	Na_2SiO_3 (kg/m^3)	116.72	34.12	29.51-292
X_9	Na_2SiO_3/NaOH	1.84	0.71	1-4.5
X_{10}	AA/FA	0.48	0.13	0.09-0.86
X_{11}	Water (kg/m^3)	119.76	23.29	52.5-197.23
X_{12}	PCE (kg/m^3)	2.36	3.41	0-10.5
X_{13}	Temperature (℃)	0/50/60/80	-	4
X_{14}	Duration (h)	0/24/48	-	3

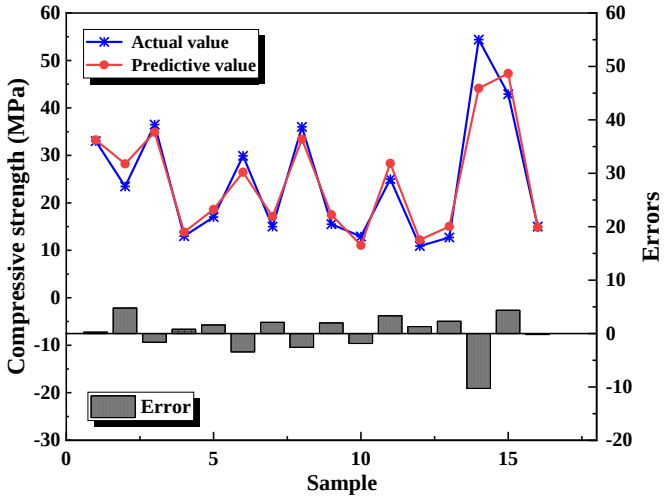
4. Results and Discussion

4.1 Performances of machine learning models

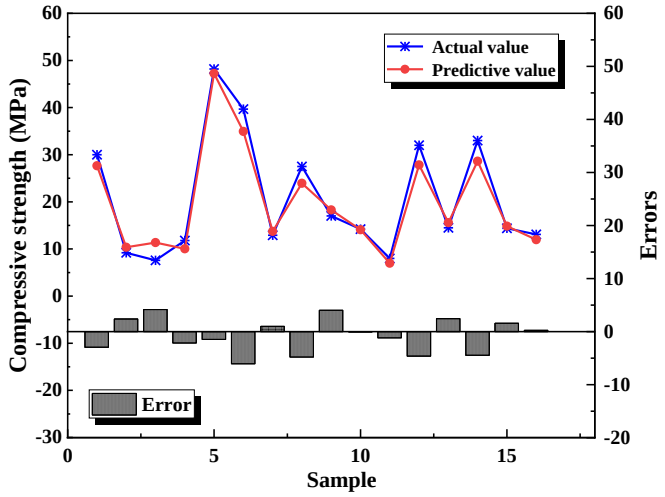
Figs. 3-5 show the machine learning prediction results of the 28-day compressive strength of FA-based geopolymer concrete mixes. The red points in each figure represent the predicted values, while the blue points are measured values. The gray histogram below these values corresponds to their errors. For all the three algorithms, the predicted and measured values of the training set were relatively close. However, several errors calculated by SVM and BPNN were more than 10 MPa. The analysis of the testing set revealed that some predictive points of SVM and ELM deviated greatly from the actual compressive strength values, while BPNN maintained relatively stable prediction results.



(a) Training set

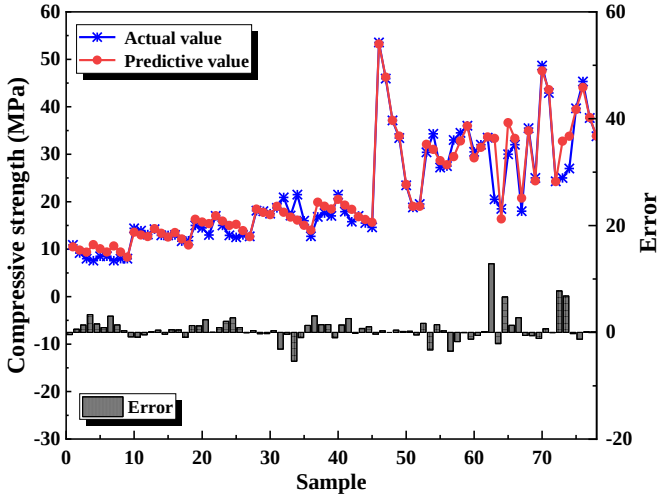


(b) Validation set

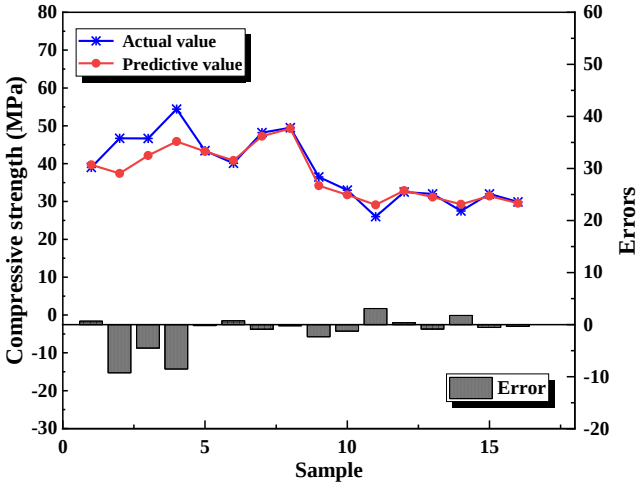


(c) Testing set

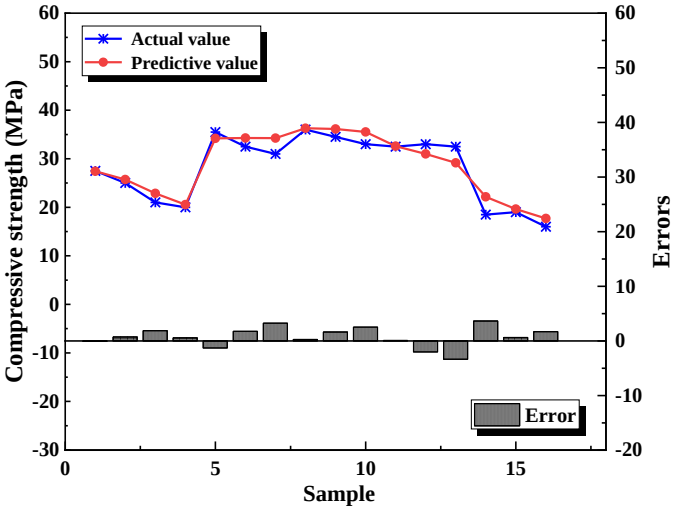
Fig. 3 Target and predicted compressive strength calculated by SVM



(a) Training set

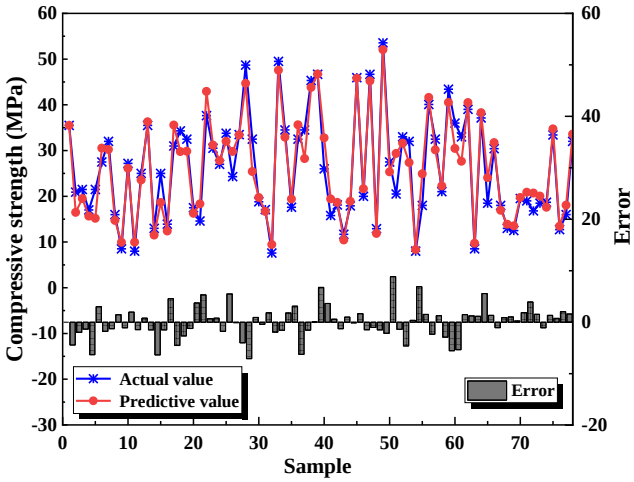


(b) Validation set

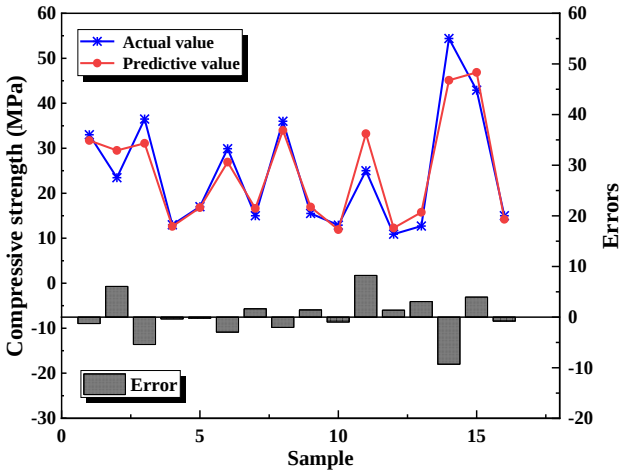


(c) Testing set

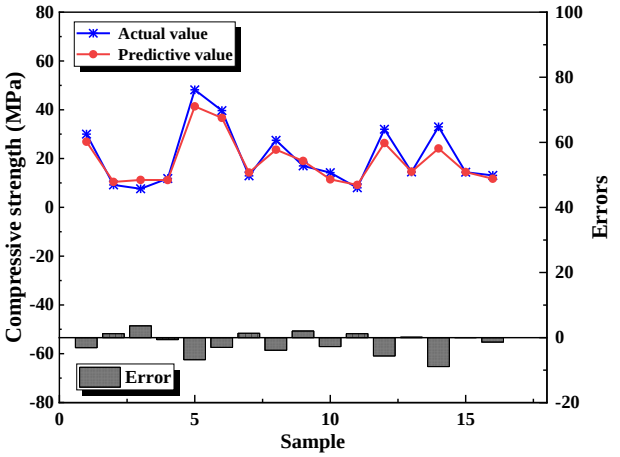
Fig. 4 Target and predicted compressive strength calculated by BPNN



(a) Training set



(b) Validation set



(c) Testing set

Fig. 5 Target and predicted compressive strength calculated by ELM

4.2 Comparison of different models

Table 6 lists the evaluation parameters obtained by the three models, which revealed similar changing trends. In general, the accuracy of BPNN, characterized by the R^2 of the total database, was the highest among all three algorithms (see the "All" column in Table 6). However, when the prediction performance of different sets was compared, the R^2 of BPNN validation set was 0.8221, which was far lower than that of the training and testing sets. Meanwhile, the R^2 of ELM testing set was only 0.7032, showing that although ELM could predict the function relationship of a given database, it does not have good generalization ability. Other evaluation parameters (*MSE*, *RMSE* and *MAE*) also verified this over fitting phenomenon. When compared to BPNN and ELM, SVM revealed a more stable prediction accuracy in all three sets, as reflected in Fig. 6.

Table 6 Precision of three machine learning algorithms for the 28-day compressive strength of FA-based geopolymer concrete

	SVM				BPNN				ELM			
	Training set	Validation set	Testing set	All	Training set	Validation set	Testing set	All	Training set	Validation set	Testing set	All
R^2	0.9038	0.9200	0.9559	0.9148	0.9472	0.8221	0.9156	0.9323	0.9224	0.8936	0.9042	0.9146
MSE	12.16	12.60	6.41	11.39	6.26	13.54	3.80	6.83	9.81	16.75	13.91	11.41
RMSE	3.49	3.55	2.53	3.37	2.5	3.68	1.95	2.61	3.13	4.09	3.73	3.38
MAE	2.28	2.68	2.04	2.30	1.48	2.92	1.59	1.61	2.41	3.07	2.85	2.57

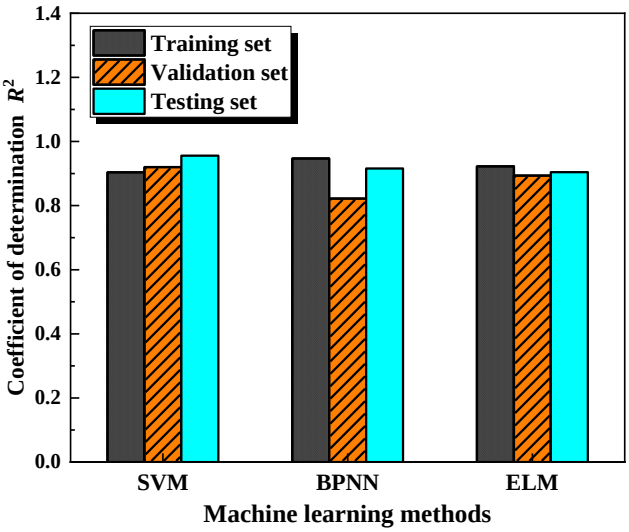
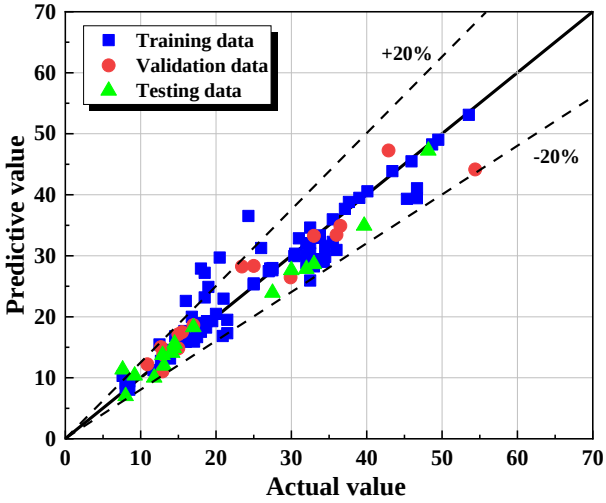


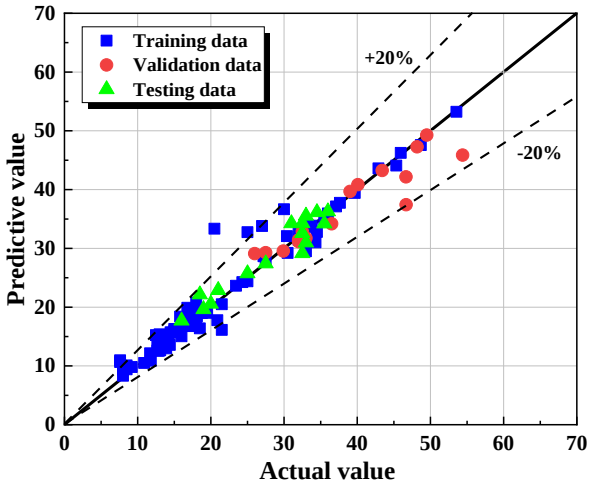
Fig. 6 Coefficient of determination (R^2) of three machine learning algorithms

Fig. 7 shows the results of predicted and experimental compressive strength obtained by SVM, BPNN and ELM. Most of the data points produced by the three algorithms were very close to the center line and were basically located in the area within an error range of 20%. Across all models, a small part of the errors between the actual and predicted values were more than 20%. The higher number of testing data points outside the error boundary of ±20% shown in Fig. 7(c) was an indication of the low performance of ELM in the testing set in comparison to other models.

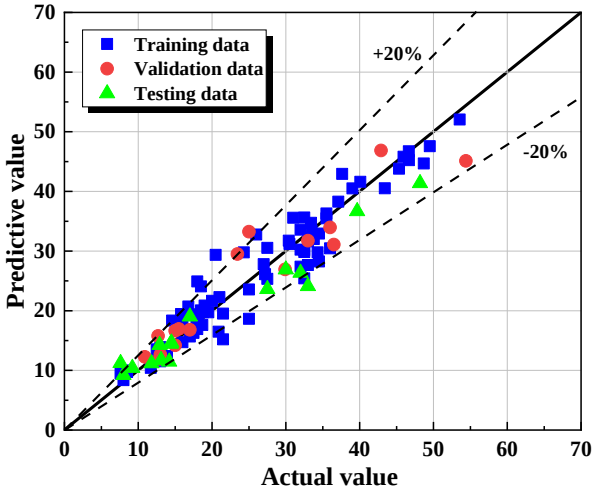
Fig. 8 shows the error percentages based on different machine learning algorithms and their number of instances in the corresponding range. Considering that relative error could reflect the credibility of measurement better, error percentage instead of error was calculated to fit normal distribution. When the actual compressive strength of a certain mix proportion was relatively small, a sample with the same absolute error showed a greater deviation than a sample with a large compressive strength. The use of error percentage could avoid this situation. According to the fitting results and percentage distributions, the average percentages of three algorithms were all close to zero. By contrast, there were more samples within the error range of -20% and 20% for BPNN, resulting in the steepest fitting curve. Considering that the most ideal error percentage distribution scenario involved minimizing of all errors by bringing them as close to 0% as possible, BPNN demonstrated the most concentrated distribution among the three algorithms. Furthermore, it can be seen from Figs. 7 and 8 that although the number of samples in the -20% to 20% range was higher in BPNN than SVM, the prediction result obtained by SVM was closer to the actual result (i.e. more sample points were concentrated near the black line shown in Fig. 7). These outcomes support the earlier observations on the higher prediction accuracy of SVM.



(a) SVM

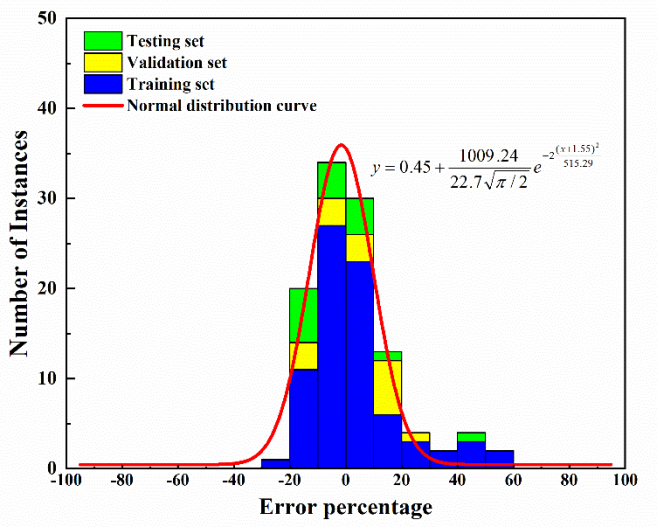


(b) BPNN

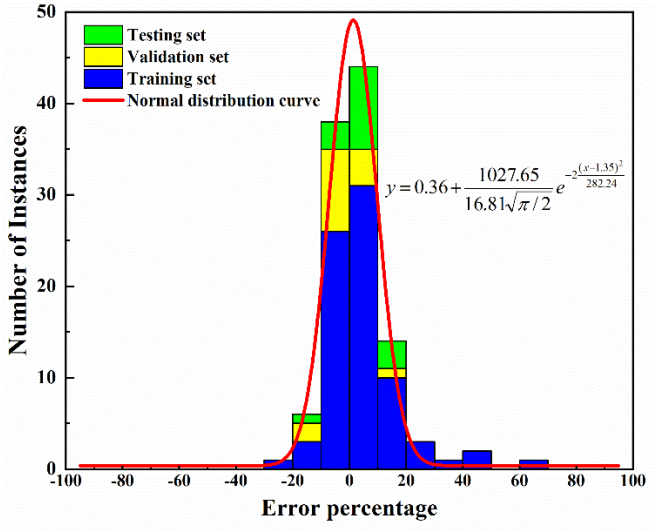


(c) ELM

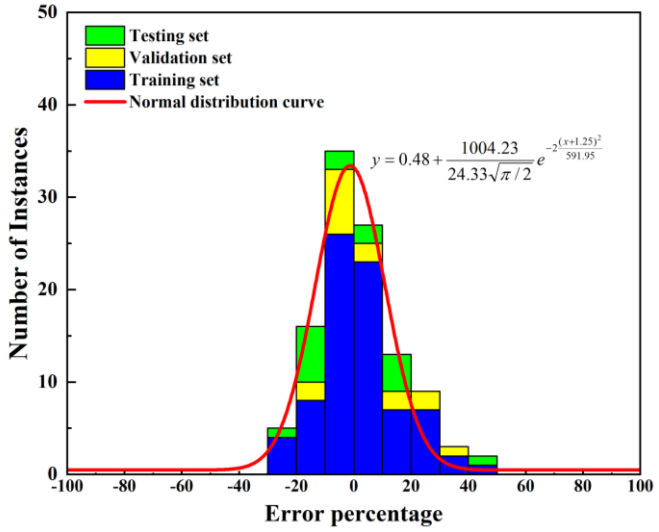
Fig. 7 Measured value vs. predicted value of three machine learning algorithms



(a) SVM



(b) BPNN

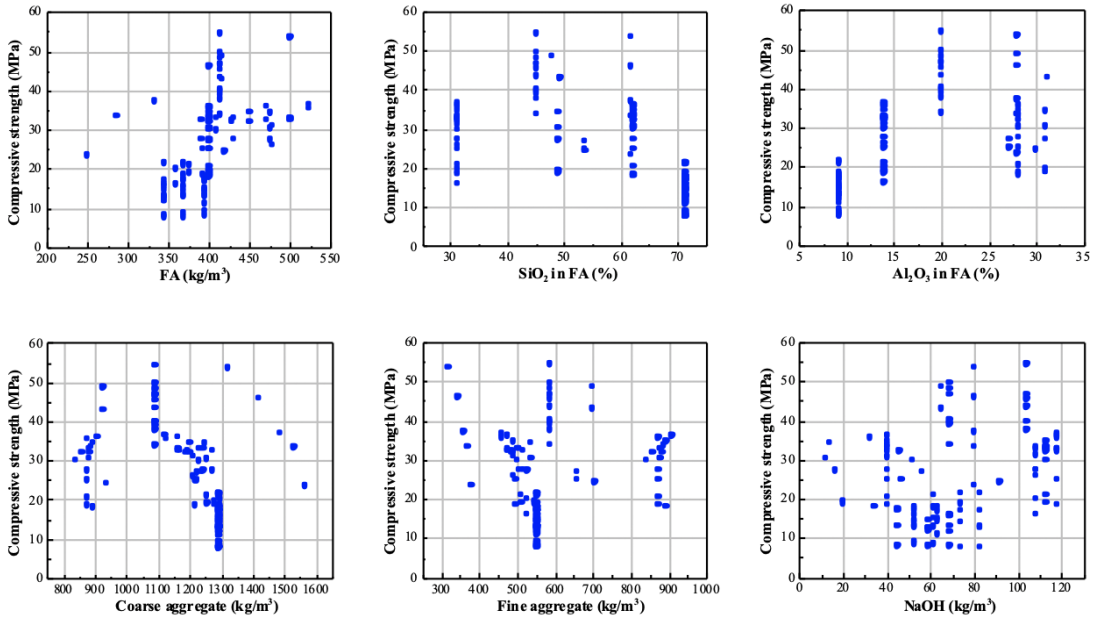


(c) ELM

Fig. 8 Error percentage distribution and normal distribution fitting of models

4.3 Effects of different input parameters

Fig. 9 shows the effect of different mix design variables on the compressive strength of FA-based geopolymer concrete. The presented results revealed the lack of any significant positive or negative relationship between any of the variables and compressive strength, which may be due to the large number of variables and their wide ranges adopted in this study, resulting in a small influence weight of any single factor on the results. However, some potential trends could still be identified through the relationship between some independent variables and dependent variables. For instance, the results of SiO_2 and Al_2O_3 contents highlighted the existence of an optimal SiO_2/Al_2O_3 ratio, which enhanced the contribution of FA to compressive strength. The content and molar concentration of NaOH and Na_2SiO_3 also affected compressive strength. Accordingly, an increase in the molar concentration of NaOH solution generally improved the overall compressive strength, especially when the NaOH molar concentration was over 12 M. Alternatively, an increase in the water content beyond 100 kg/m^3 reduced strength. The contents of coarse and fine aggregates did not seem to have a significant effect on strength. This trend was in line with the findings of previous studies, where it was reported that the interface transition zone (ITZ), morphology and particle size distribution of the aggregates was more influential in strength development than their contents [30, 54]. Finally, Table 7 elucidates the effect of pre-curing on the compressive strength of FA-based geopolymer concrete, revealing the role pre-curing plays in improving the minimum, average and maximum strength values of all data sets.



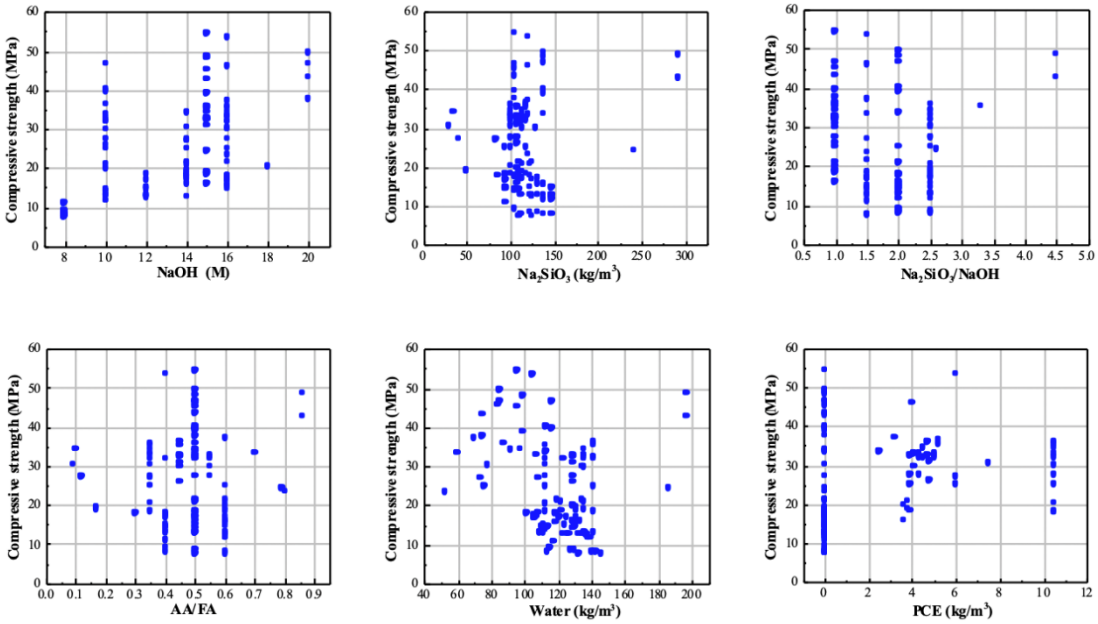


Fig. 9 Correlation between mix design variables and compressive strength

Table 7 Effect of pre-curing on the 28-day compressive strength of FA-based geopolymer concrete

Curing condition	No pre-curing			With pre-curing (up to 80°C and 48h)		
Compressive strength (MPa)	Min	Max	Ave.	Min	Max	Ave.
	7.6	46.7	21.4	18.0	54.4	35.7

In addition to understanding the potential impact of each input parameter (i.e. feature) on the prediction results, it is also necessary to compare their importance for the result influence. This approach can enable the identification of certain features that have a greater impact on the final results, and also reveal those with relatively small impacts. A common method to measure the influence of different features on the results is permutation feature importance (*PFI*) [55]. The core idea of *PFI* is that if a certain input variable (X_i) has a great influence on the result, the prediction accuracy will significantly decrease by randomly arranging X_i, during which the order of other variables is unchanged. According to this definition, the calculation expression of the *PFI* value for a specific variable obtained by using *MAE* as error measurement function is shown in Eq. (9), where MAE_{perm} and MAE_{orig} are mean absolute error before and after randomly adjusting X_i sequence, respectively.

$$PFI = MAE_{perm} - MAE_{orig} \quad (9)$$

In line with this approach, when the *PFI* value was close to zero, changing a certain feature had a lower influence on the output value; whereas when *PFI* was large, the influence of the feature on the output value was significant. Table 8 shows the feature

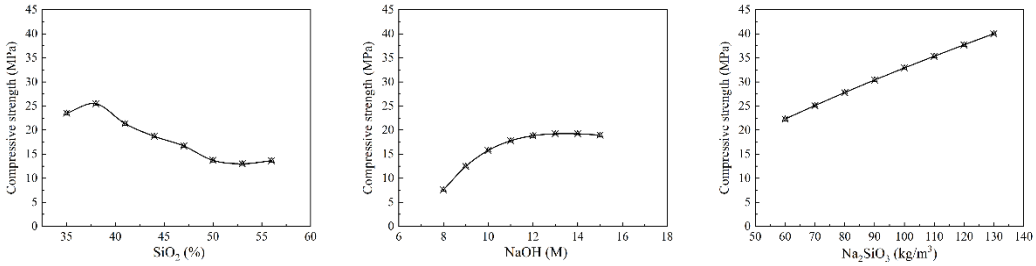
importance of 14 input variables obtained by the SVM algorithm. The *PFI* values presented in the last column indicated that the SiO_2 and Al_2O_3 contents in FA has significant effects on the strength results. This is due to the fact that the overall degree and rate of hydration determines the mechanical properties of cement-based mixes. In this respect, the chemical composition of FA will directly affect the hydration process and the formation of hydration products, thus playing a key role in strength development [56, 57]. These were followed by the activator contents/ratios and PCE content introduced in the initial mix design. In contrast, other parameters such as FA, coarse and fine aggregate and water contents; as well as pre-curing conditions revealed lower *PFI* values. Although *PFI* index can represent the influence of a single influencing factor on the final result, it can only reveal the relationship between input parameters and output parameters from the perspective of single correspondence because the final compressive strength of concrete is often caused by the comprehensive action of several factors.

Table 8 *PFI* values of different input parameters

Notation	Input parameter	MAE_{perm}	*PFI*
X_1	FA (kg/m^3)	2.33	0.03
X_2	SiO_2 (%)	3.01	0.71
X_3	Al_2O_3 (%)	2.31	0.21
X_4	Coarse aggregate (kg/m^3)	2.33	0.03
X_5	Fine aggregate (kg/m^3)	2.32	0.02
X_6	NaOH (kg/m^3)	2.38	0.08
X_7	NaOH (M)	2.55	0.25
X_8	Na_2SiO_3 (kg/m^3)	2.6	0.30
X_9	Na_2SiO_3/NaOH	2.42	0.12
X_{10}	AA/FA	2.36	0.06
X_{11}	Water (kg/m^3)	2.39	0.09
X_{12}	PCE (kg/m^3)	2.56	0.26
X_{13}	Temperature (°C)	2.37	0.07
X_{14}	Duration (h)	2.39	0.09

According to the findings of the preceding investigation, the impact of different input factors on the ultimate compressive strength varies. Accordingly, changing some of these parameters may considerably increase or decrease the strength of concrete, while altering other parameters has minimal influence. Based on the PFI parameter analysis presented above, it is clear that the chemical composition of FA (particularly the SiO_2 content), molar concentration of NaOH and mass fraction of Na_2SiO_3 solution all have significant effect weights. Considering that machine learning methods cannot provide an exact equation between input and output variables, a sensitivity analysis can be utilized to analyze and evaluate the probable changes in compressive strength when ideally only one single parameter value is changed [58, 59]. Fig. 10 depicts the variance in compressive strength derived by using the BPNN model

and the abovementioned three input parameters. It can be seen from Fig. 10 that when the SiO_2 content in FA increases over a certain value, the compressive strength steadily decreases. Alternatively, increasing the concentration of NaOH and the content of Na_2SiO_3 solution improves the compressive strength, within the ranges presented in this study. These observed trends are consistent with the results proposed by previous relevant studies [44, 47].



(a) SiO_2 content in FA (b) NaOH (M) (c) Na_2SiO_3 (kg/m^3)

Fig. 10 Sensitivity analysis of feature parameters based on BPNN model

5. Conclusions

In this study, three different machine learning algorithms (BPNN, SVM and ELM) were used to predict the 28-day compressive strength of FA-based geopolymer concrete. Through a detailed literature search, 110 groups of mix proportion data were collected. The collected data included 14 different variables such as FA content and composition (i.e. SiO_2 and Al_2O_3 contents), coarse and fine aggregate contents, NaOH and Na_2SiO_3 content and ratio, AA/FA ratio, water and PCE contents, pre-curing temperature and duration. The performance of each algorithm was evaluated by R^2, *MSE*, *RMSE* and *MAE*. The prediction performance of the three models and the influence weight of different input parameters on the results were compared. The following conclusions were drawn from the obtained results:

1. BPNN, SVM and ELM all revealed good prediction performance. The predicted values based on the three models were close to the actual values, and their error ranges were generally within 20%.

2. For the total data sets, the R^2 order of three algorithms from high to low was BPNN > ELM > SVM. However, the analysis of different sets (i.e. training set, validation set and testing set) revealed that the R^2 of validation set calculated by BPNN was 0.8221, while the R^2 of testing set obtained by ELM was only 0.7032, indicating that the generalization abilities of BPNN and ELM were weak. In all the three sets, SVM achieved the most stable prediction ability. Compared with the other two models, the error percentage of BPNN was more concentrated in the range of -20% to 20%, supported by a steep fitted normal distribution curve.

3. Through the influence analysis of single feature on compressive strength, it was found that the SiO_2 and Al_2O_3 contents in FA had a direct influence on the compressive strength of geopolymer concrete. The content and concentration of the alkaline activator also contributed to the enhancement of compressive strength, while the increase of water content beyond a certain value reduced the strength, and the aggregate contents were not as influential on strength development as these other mix design variables. Furthermore, the use of *PFI* index revealed the influence of each individual feature on the output value. Among them, the SiO_2 content in FA had the most significant effect on the strength prediction results, followed by Al_2O_3 content in FA, and activator and PCE contents, which further verified the results of single feature analysis.

Using machine learning to predict the mechanical properties of FA-based geopolymer concrete can support the traditional empirical models by revealing the prediction results more quickly and conveniently, enabling the prediction of the performance of concrete with an unknown mix proportion through deep learning. This approach can facilitate reductions in experimental workload, and labor and material consumption; improve time efficiency; and provide a theoretical and practical guidance for optimizing concrete mix proportions. These initiatives can accelerate the development and application of sustainable concrete formulations involving alternative binders (e.g. geopolymers) with minimum operational costs and environmental impact.

Acknowledgement:

This work was funded by The Royal Society (project ref: ICA\R1\201310) and China Scholarship Council (grant number: 202006370082).

References:

[1] M.O. Yusuf, M.A.M. Johari, Z.A. Ahmad, M. Maslehuddin, Strength and microstructure of alkali-activated binary blended binder containing palm oil fuel ash and ground blast-furnace slag, Constr. Build. Mater. 52 (2014) 504-510.

[2] W.K. Part, M. Ramli, C.B. Cheah, An overview on the influence of various factors on the properties of geopolymer concrete derived from industrial by-products, Constr. Build. Mater. 77 (2015) 370-395.

[3] B.C. McLellan, R.P. Williams, J. Lay, A. Van Riessen, G.D. Corder, Costs and carbon emissions for geopolymer pastes in comparison to ordinary portland cement, J. Clean. Prod. 19(9-10) (2011) 1080-1090.

[4] H. Biricik, M.S. Kırgız, A.G. de Sousa Galdino, S. Kenai, J. Mirza, J. Kinuthia, A. Ashteyat, A. Khitab, J. Khatib, Activation of slag through a combination of NaOH/NaS alkali for transforming it into geopolymer slag binder mortar–assessment the effects of two different Blaine fines and three different curing conditions, J. Mater. Res. Technol. 14 (2021) 1569-1584.

[5] B. Lothenbach, K. Scrivener, R. Hooton, Supplementary cementitious materials, Cem. Concr. Res. 41(12) (2011) 1244-1256.

[6] M.C. Juenger, R. Siddique, Recent advances in understanding the role of supplementary cementitious materials in concrete, Cem. Concr. Res. 78 (2015) 71-80.

[7] Y. Du, W. Yang, Y. Ge, S. Wang, P. Liu, Thermal conductivity of cement paste containing waste glass powder, metakaolin and limestone filler as supplementary cementitious material, J. Clean. Prod. 287 (2021) 125018.

[8] U.K. Sevim, H.H. Bilgic, O.F. Cansiz, M. Ozturk, C.D. Atis, Compressive strength prediction models for cementitious composites with fly ash using machine learning techniques, Constr. Build. Mater. 271 (2021) 121584.

[9] M.S. Kirgiz, Advance treatment by nanographite for Portland pulverised fly ash cement (the class F) systems, Compos. B. Eng. 82 (2015) 59-71.

[10] E. Crossin, The greenhouse gas implications of using ground granulated blast furnace slag as a cement substitute, J. Clean. Prod. 95 (2015) 101-108.

[11] J. Yang, J. Huang, Y. Su, X. He, H. Tan, W. Yang, B. Strnadel, Eco-friendly treatment of low-calcium coal fly ash for high pozzolanic reactivity: A step towards waste utilization in sustainable building material, J. Clean. Prod. 238 (2019) 117962.

[12] I. Navarrete, F. Vargas, P. Martinez, A. Paul, M. Lopez, Flue gas desulfurization (FGD) fly ash as a sustainable, safe alternative for cement-based materials, J. Clean. Prod. 283 (2021) 124646.

[13] M.S. Kirgiz, Green cement composite concept reinforced by graphite nano-engineered particle suspension for infrastructure renewal material, Compos. B. Eng. 154 (2018) 423-429.

[14] X. Han, J. Yang, J. Feng, C. Zhou, X. Wang, Research on hydration mechanism of ultrafine fly ash and cement composite, Constr. Build. Mater. 227 (2019) 116697.

[15] M. Wu, S. Sui, Y. Zhang, Y. Jia, W. She, Z. Liu, Y. Yang, Analyzing the filler and activity effect of fly ash and slag on the early hydration of blended cement based on calorimetric test, Constr. Build. Mater. 276 (2021) 122201.

[16] Z. Zhou, M. Sofi, J. Liu, S. Li, A. Zhong, P. Mendis, Nano-CSH modified high volume fly ash concrete: Early-age properties and environmental impact analysis, J. Clean. Prod. 286 (2021) 124924.

[17] F. Puertas, S. Martínez-Ramírez, S. Alonso, T. Vazquez, Alkali-activated fly ash/slag cements: strength behaviour and hydration products, Cem. Concr. Res. 30(10) (2000) 1625-1632.

[18] A.A. Shubbar, M. Sadique, P. Kot, W. Atherton, Future of clay-based construction materials–A review, Constr. Build. Mater. 210 (2019) 172-187.
[19] A.M. Rashad, A.A. Hassan, S.R. Zeedan, An investigation on alkali-activated Egyptian metakaolin pastes blended with quartz powder subjected to elevated temperatures, Appl. Clay Sci. 132 (2016) 366-376.
[20] M. Komljenović, Z. Baščarević, V. Bradić, Mechanical and microstructural properties of alkali-activated fly ash geopolymers, J. Hazard. Mater. 181(1-3) (2010) 35-42.
[21] B. Singh, G. Ishwarya, M. Gupta, S.K. Bhattacharyya, Geopolymer concrete: A review of some recent developments, Constr. Build. Mater. 85 (2015) 78-90.
[22] X.Y. Zhuang, L. Chen, S. Komarneni, C.H. Zhou, D.S. Tong, H.M. Yang, W.H. Yu, H. Wang, Fly ash-based geopolymer: clean production, properties and applications, J. Clean. Prod. 125 (2016) 253-267.
[23] P. Chindaprasirt, T. Chareerat, S. Hatanaka, T. Cao, High-strength geopolymer using fine high-calcium fly ash, J. Mater. Civ. Eng. 23(3) (2011) 264-270.
[24] M. Soutsos, A.P. Boyle, R. Vinai, A. Hadjierakleous, S.J. Barnett, Factors influencing the compressive strength of fly ash based geopolymers, Constr. Build. Mater. 110 (2016) 355-368.
[25] D.L. Kong, J.G. Sanjayan, Effect of elevated temperatures on geopolymer paste, mortar and concrete, Cem. Concr. Res. 40(2) (2010) 334-339.
[26] P. Chindaprasirt, U. Rattanasak, S. Taebuanhuad, Resistance to acid and sulfate solutions of microwave-assisted high calcium fly ash geopolymer, Mater. Struct. 46(3) (2013) 375-381.
[27] P. Chindaprasirt, W. Chalee, Effect of sodium hydroxide concentration on chloride penetration and steel corrosion of fly ash-based geopolymer concrete under marine site, Constr. Build. Mater. 63 (2014) 303-310.
[28] H. Zhang, L. Li, C. Yuan, Q. Wang, P.K. Sarker, X. Shi, Deterioration of ambient-cured and heat-cured fly ash geopolymer concrete by high temperature exposure and prediction of its residual compressive strength, Constr. Build. Mater. 262 (2020) 120924.
[29] X. Cong, W. Zhou, M. Elchalakani, Experimental study on the engineering properties of alkali-activated GGBFS/FA concrete and constitutive models for performance prediction, Constr. Build. Mater. 240 (2020) 117977.
[30] H.-B. Le, Q.-B. Bui, L. Tang, Geopolymer Recycled Aggregate Concrete: From Experiments to Empirical Models, Materials 14(5) (2021) 1180.
[31] W.B. Chaabene, M. Flah, M.L. Nehdi, Machine learning prediction of mechanical properties of concrete: Critical review, Constr. Build. Mater. 260 (2020) 119889.
[32] F. Farooq, W. Ahmed, A. Akbar, F. Aslam, R. Alyousef, Predictive modeling for sustainable high-performance concrete from industrial wastes: A comparison and optimization of models using ensemble learners, J. Clean. Prod. 292 (2021) 126032.
[33] A.-T. Akono, J. Chen, M. Zhan, S.P. Shah, Basic creep and fracture response of fine recycled aggregate concrete, Constr. Build. Mater. 266 (2021) 121107.
[34] A. Karakoç, Ö. Keleş, A predictive failure framework for brittle porous materials via machine learning and geometric matching methods, J. Mater. Sci. 55(11) (2020) 4734-4747.
[35] D.-C. Feng, Z.-T. Liu, X.-D. Wang, Y. Chen, J.-Q. Chang, D.-F. Wei, Z.-M. Jiang, Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach, Constr. Build. Mater. 230 (2020) 117000.
[36] J. Xu, Y. Chen, T. Xie, X. Zhao, B. Xiong, Z. Chen, Prediction of triaxial behavior of recycled

aggregate concrete using multivariable regression and artificial neural network techniques, Constr. Build. Mater. 226 (2019) 534-554.
[37] A.P. Marugán, F.P.G. Márquez, J.M.P. Perez, D. Ruiz-Hernández, A survey of artificial neural network in wind energy systems, Appl. Energy 228 (2018) 1822-1836.
[38] O.B. Douma, B. Boukhatem, M. Ghrici, A. Tagnit-Hamou, Prediction of properties of self-compacting concrete containing fly ash using artificial neural network, Neural. Comput. Appl. 28(1) (2017) 707-718.
[39] P.O. Awoyera, M.S. Kirgiz, A. Viloria, D. Ovallos-Gazabon, Estimating strength properties of geopolymer self-compacting concrete using machine learning techniques, J. Mater. Res. Technol. 9(4) (2020) 9016-9028.
[40] A. Nazari, J.G. Sanjayan, Modelling of compressive strength of geopolymer paste, mortar and concrete by optimized support vector machine, Ceram. Int. 41(9) (2015) 12164-12177.
[41] K. Yan, C. Shi, Prediction of elastic modulus of normal and high strength concrete by support vector machine, Constr. Build. Mater. 24(8) (2010) 1479-1485.
[42] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: a new learning scheme of feedforward neural networks, 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541), Ieee, 2004, pp. 985-990.
[43] J.S. Tu, Y.Z. Liu, M. Zhou, R.X. Li, Prediction and analysis of compressive strength of recycled aggregate thermal insulation concrete based on GA-BP optimization network, J. Eng. Des. Technol. 19(2) (2021) 412-422.
[44] M.T. Ghafoor, Q.S. Khan, A.U. Qazi, M.N. Sheikh, M. Hadi, Influence of alkaline activators on the mechanical properties of fly ash based geopolymer concrete cured at ambient temperature, Constr. Build. Mater. 273 (2021) 121752.
[45] P.e. Pavithra, M.S. Reddy, P. Dinakar, B.H. Rao, B. Satpathy, A. Mohanty, A mix design procedure for geopolymer concrete with fly ash, J. Clean. Prod. 133 (2016) 117-125.
[46] T. Xie, T. Ozbakkaloglu, Behavior of low-calcium fly and bottom ash-based geopolymer concrete cured at ambient temperature, Ceram. Int. 41(4) (2015) 5945-5958.
[47] A.A. Aliabdo, M. Abd Elmoaty, H.A. Salem, Effect of water addition, plasticizer and alkaline solution constitution on fly ash based geopolymer concrete performance, Constr. Build. Mater. 121 (2016) 694-703.
[48] C. Gunasekara, D.W. Law, S. Setunge, Long term permeation properties of different fly ash geopolymer concretes, Constr. Build. Mater. 124 (2016) 352-362.
[49] P.S. Deb, P. Nath, P.K. Sarker, The effects of ground granulated blast-furnace slag blending with fly ash and activator content on the workability and strength properties of geopolymer concrete cured at ambient temperature, Mater. Des. 62 (2014) 32-39.
[50] P. Topark-Ngarm, P. Chindaprasirt, V. Sata, Setting time, strength, and bond of high-calcium fly ash geopolymer concrete, J. Mater. Civ. Eng. 27(7) (2015) 04014198.
[51] T. Phoo-Ngernkham, C. Phiangphimai, N. Damrongwiriyanupap, S. Hanjitsuwan, J. Thumrongvut, P. Chindaprasirt, A mix design procedure for alkali-activated high-calcium fly ash concrete cured at ambient temperature, Adv. Mater. Sci. Eng. 2018 (2018).
[52] Y. Wang, X. Liu, W. Zhang, Z. Li, Y. Zhang, Y. Li, Y. Ren, Effects of Si/Al ratio on the efflorescence and properties of fly ash based geopolymer, J. Clean. Prod. 244 (2020) 118852.
[53] B. Lee, G. Kim, R. Kim, B. Cho, S. Lee, C.-M. Chon, Strength development properties of geopolymer paste and mortar with respect to amorphous Si/Al ratio of fly ash, Constr. Build. Mater. 151

(2017) 512-519.
[54] A. Hassan, M. Arif, M. Shariq, Use of geopolymer concrete for a cleaner and sustainable environment–A review of mechanical properties and microstructure, J. Clean. Prod. 223 (2019) 704-728.
[55] M. Almustafa, M. Nehdi, Machine learning model for predicting structural response of RC slabs exposed to blast loading, Eng. Struct. 221 (2020) 111109.
[56] G.D. Moon, S. Oh, Y.C. Choi, Effects of the physicochemical properties of fly ash on the compressive strength of high-volume fly ash mortar, Constr. Build. Mater. 124 (2016) 1072-1080.
[57] Y.K. Cho, S.H. Jung, Y.C. Choi, Effects of chemical composition of fly ash on compressive strength of fly ash cement mortar, Constr. Build. Mater. 204 (2019) 255-264.
[58] H. Naseri, H. Jahanbakhsh, P. Hosseini, F.M. Nejad, Designing sustainable concrete mixture by developing a new machine learning technique, J. Clean. Prod. 258 (2020) 120578.
[59] P.G. Asteris, A.D. Skentou, A. Bardhan, P. Samui, K. Pilakoutas, Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models, Cem. Concr. Res. 145 (2021) 106449.

List of Tables:

Table 1 Hyperparameters of machine learning models

Machine learning algorithms	Hyperparameters	Values/categories
BPNN	Activation function	Sigmoid
	Hidden layer numbers	1
	Number of hidden layer nodes	18
SVM	SVM type	e-SVR
	Kernel function	RBF
	Tolerance	0.001
	epsilon	0.1
	Shrinking	True
ELM	Hidden layer nodes	20
	Activation function	Sigmoid
	Mode	Regression

Table 2 A sample mix proportion of FA-based geopolymer concrete representing those used this study

FA (kg/m^3)	Coarse aggregate (kg/m^3)	Fine aggregate (kg/m^3)	NaOH solution (kg/m^3)	NaOH (M)	Na_2SiO_3 solution (kg/m^3)	Na_2SiO_3/NaOH	AA/FA	Water (kg/m^3)	PCE (kg/m^3)
400	875	875	40	16	100	2.5	0.35	112.4	10.5

Table 3 Chemical composition of FA obtained from different sources

Reference		SiO_2	Al_2O_3	CaO	SO_3	Fe_2O_3	MgO	LOI
[44]		71.50	9.20	6.72	2.40	2.37	0.60	3.67
[45]		61.89	28.05	0.87	1.32	4.11	0.38	0.49
[46]		49	31	5	-	3	3	-
[47]		62.30	28.10	0.5	0.40	2.10	1.00	2.50
[48]	Type I	47.87	28.0	3.81	0.27	14.09	0.93	0.43
	Type II	49.37	31.25	4.80	0.24	4.47	1.28	0.51
	Type III	53.82	29.95	1.03	0.34	9.24	0.58	0.63
[49]		53.71	27.20	1.90	0.30	11.17	-	0.68
[50]		45.23	19.95	15.51	-	13.15	-	-
[51]		31.32	13.96	25.79	3.29	15.64	2.94	1.30

LOI = Loss on ignition

Table 4 Particle size parameters of coarse and fine aggregates obtained from different sources

Reference	Coarse aggregates		Fine aggregates		
	Raw materials	Size range (mm)	Raw materials	Size range (mm)	Fineness modulus
[44]	-	9.5-12.5	River sand	-	2.35
[45]	-	6.3-20	Graded river sand	< 4.75	-
[46]	Crushed bluestone gravel	7	Graded sand	< 0.4	-
[47]	Pink limestone	1.18-9.5	Natural siliceous sand	0.075-4.75	2.45
[48]	Crushed basalt aggregate	7-10	Uncrushed river sand	-	3.0
[49]	Crushed granite	7-20	Sand	-	1.97
[50]	Limestone	20	River sand	-	2.90
[51]	Crushed limestone	7-16	River sand	-	2.20

Table 5 Mean, standard deviation and range of input/output parameters

Notation	Input parameter	Mean/count	Standard deviation	Range/categories
X_1	FA (kg/m^3)	396.63	44.29	250-523
X_2	SiO_2 (%)	56.34	15.93	31.32-71.5
X_3	Al_2O_3 (%)	16.55	8.01	9.2-31.25
X_4	Coarse aggregate (kg/m^3)	1196.76	155.86	840-1567
X_5	Fine aggregate (kg/m^3)	582.25	127.59	320-910
X_6	NaOH (kg/m^3)	71.96	28.35	11.78-118
X_7	NaOH (M)	13.32	3.01	8-20
X_8	Na_2SiO_3 (kg/m^3)	116.72	34.12	29.51-292
X_9	Na_2SiO_3/NaOH	1.84	0.71	1-4.5
X_{10}	AA/FA	0.48	0.13	0.09-0.86
X_{11}	Water (kg/m^3)	119.76	23.29	52.5-197.23
X_{12}	PCE (kg/m^3)	2.36	3.41	0-10.5
X_{13}	Temperature (°C)	0/50/60/80	-	4
X_{14}	Duration (h)	0/24/48	-	3

Table 6 Precision of three machine learning algorithms for the 28-day compressive strength of FA-based geopolymer concrete

	SVM				BPNN				ELM			
	Training set	Validation set	Testing set	All	Training set	Validation set	Testing set	All	Training set	Validation set	Testing set	All
R^2	0.9038	0.9200	0.9559	0.9148	0.9472	0.8221	0.9156	0.9323	0.9224	0.8936	0.9042	0.9146
MSE	12.16	12.60	6.41	11.39	6.26	13.54	3.80	6.83	9.81	16.75	13.91	11.41
RMSE	3.49	3.55	2.53	3.37	2.5	3.68	1.95	2.61	3.13	4.09	3.73	3.38
MAE	2.28	2.68	2.04	2.30	1.48	2.92	1.59	1.61	2.41	3.07	2.85	2.57

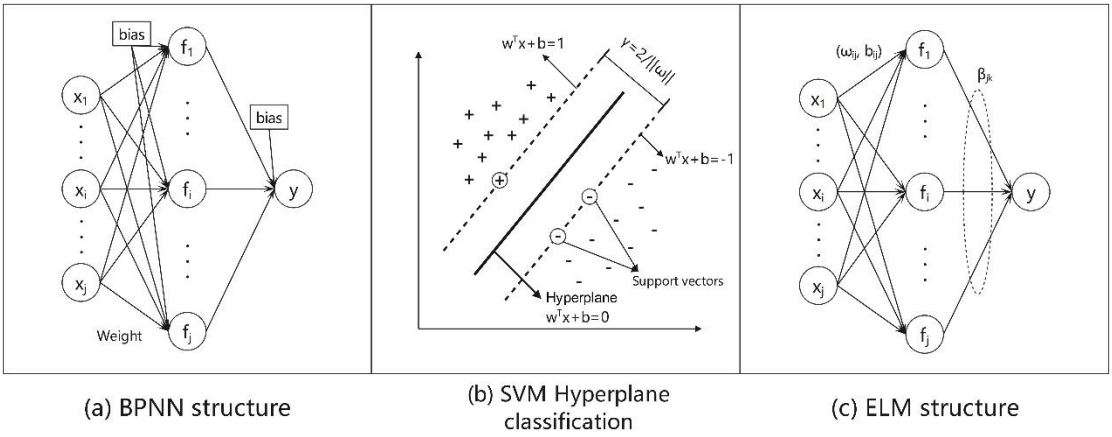
Table 7 Effect of pre-curing on the 28-day compressive strength of FA-based geopolymer concrete

Curing condition	No pre-curing			With pre-curing (up to 80℃ and 48h)		
Compressive strength (MPa)	Min	Max	Ave.	Min	Max	Ave.
	7.6	46.7	21.4	18.0	54.4	35.7

Table 8 *PFI* values of different input parameters

Notation	Input parameter	MAE_{perm}	*PFI*
X_1	FA (kg/m^3)	2.33	0.03
X_2	SiO_2 (%)	3.01	0.71
X_3	Al_2O_3 (%)	2.31	0.21
X_4	Coarse aggregate (kg/m^3)	2.33	0.03
X_5	Fine aggregate (kg/m^3)	2.32	0.02
X_6	NaOH (kg/m^3)	2.38	0.08
X_7	NaOH (M)	2.55	0.25
X_8	Na_2SiO_3 (kg/m^3)	2.6	0.30
X_9	Na_2SiO_3/NaOH	2.42	0.12
X_{10}	AA/FA	2.36	0.06
X_{11}	Water (kg/m^3)	2.39	0.09
X_{12}	PCE (kg/m^3)	2.56	0.26
X_{13}	Temperature (°C)	2.37	0.07
X_{14}	Duration (h)	2.39	0.09

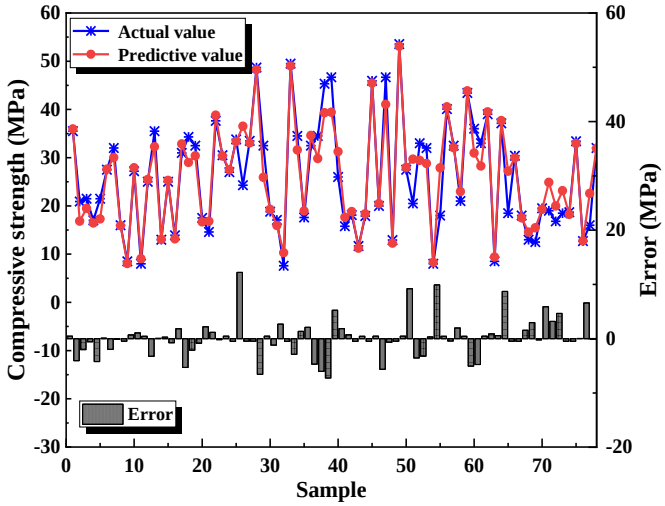
List of Figures:



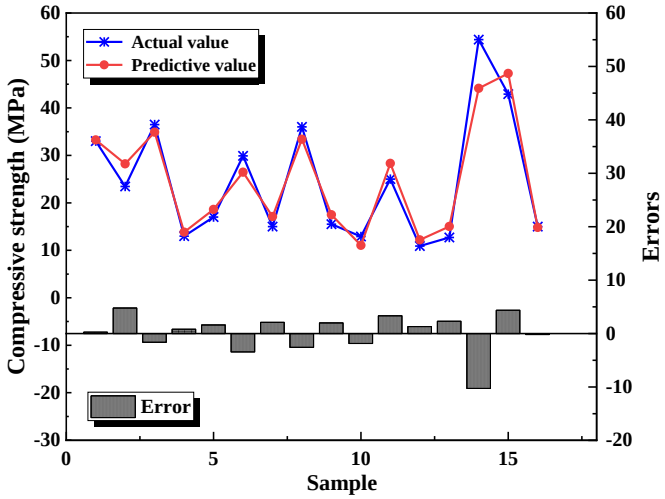
(a) BPNN structure (b) SVM Hyperplane classification (c) ELM structure

Fig. 1 Graphical representation of different machine learning models (modified from [31] and [8])

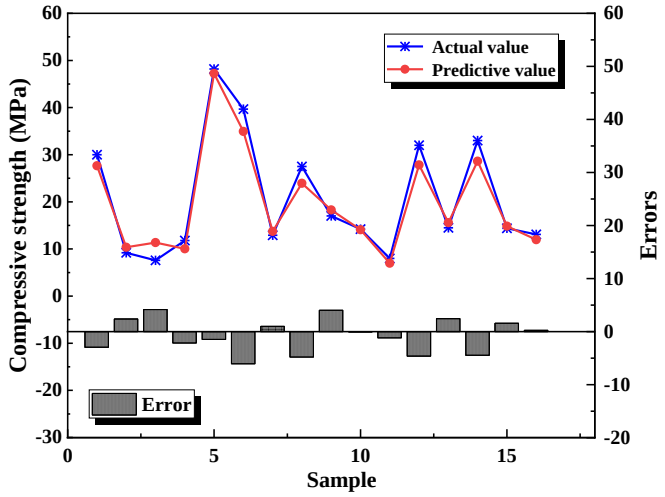
Fig. 2 Construction and data prediction process of machine learning models



(a) Training set

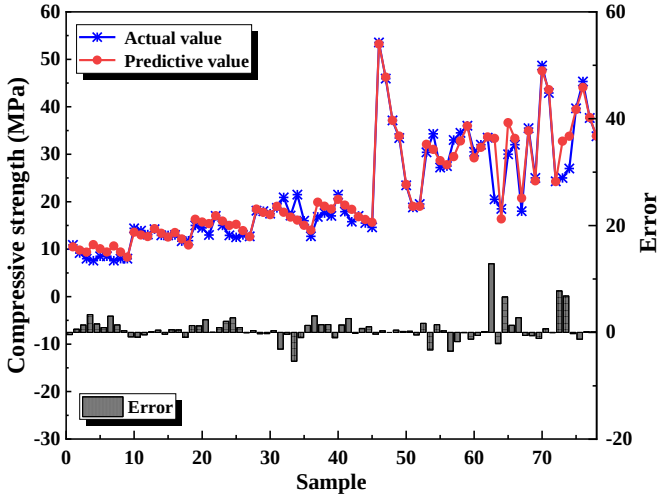


(b) Validation set

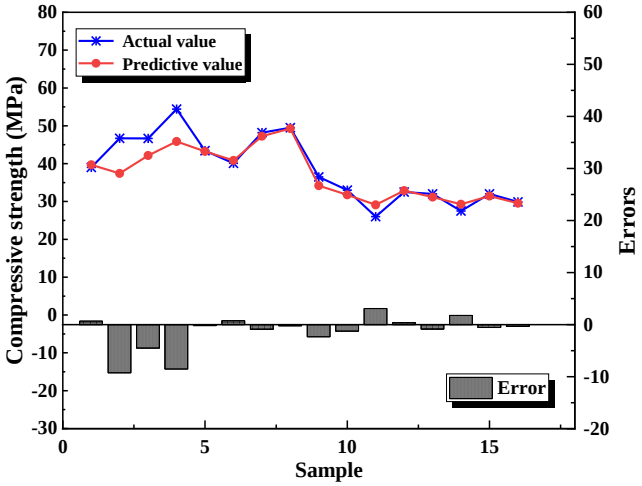


(c) Testing set

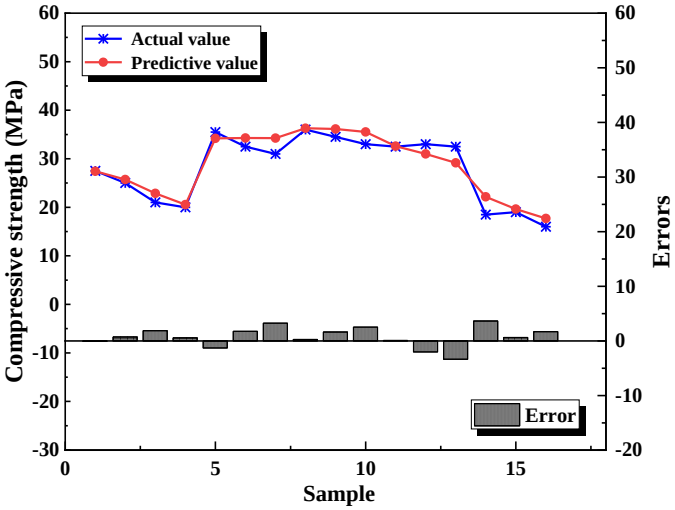
Fig. 3 Target and predicted compressive strength calculated by SVM



(a) Training set

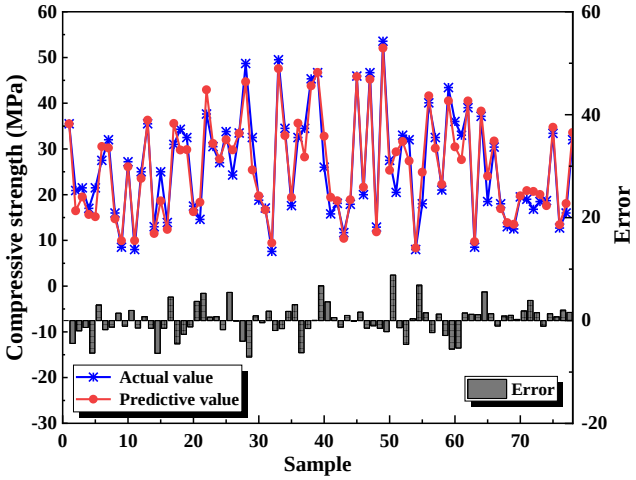


(b) Validation set

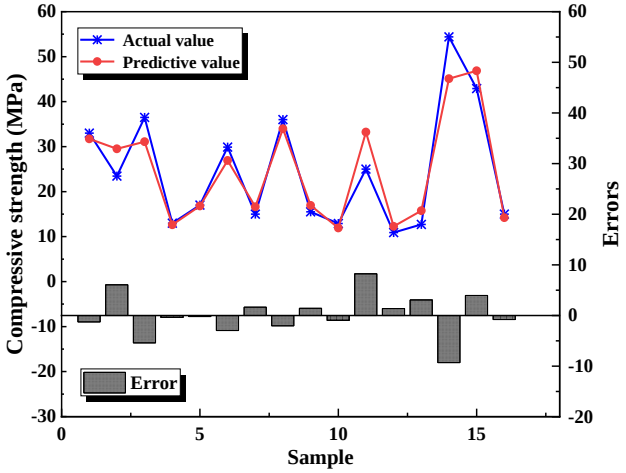


(c) Testing set

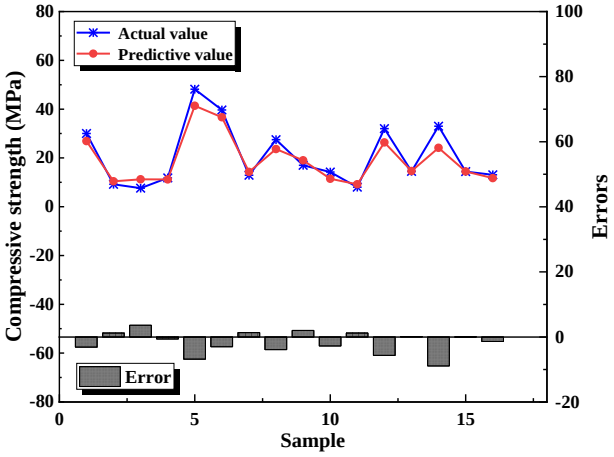
Fig. 4 Target and predicted compressive strength calculated by BPNN



(a) Training set



(b) Validation set



(c) Testing set

Fig. 5 Target and predicted compressive strength calculated by ELM

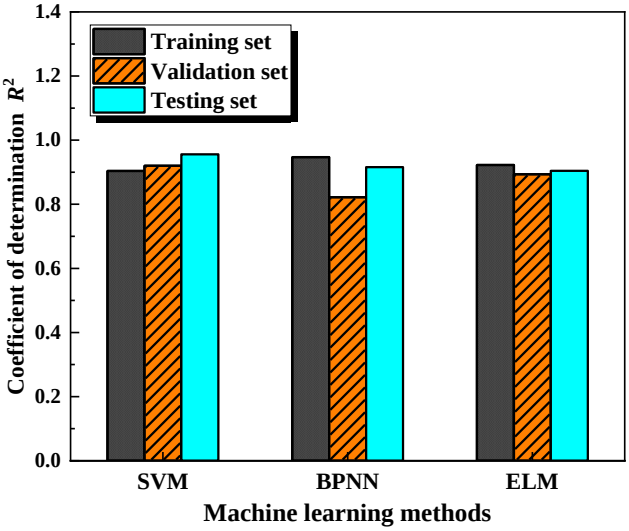
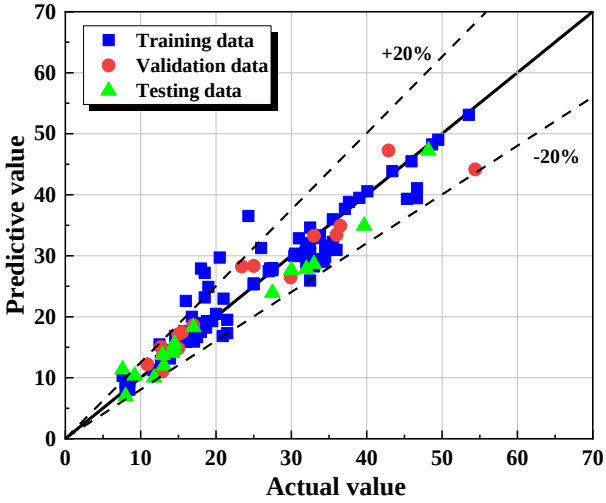
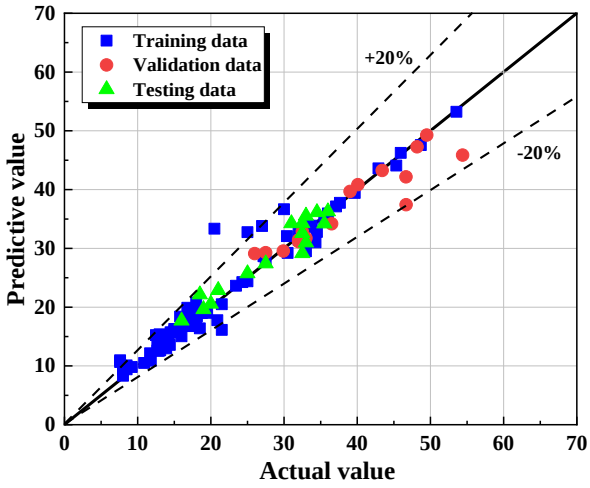


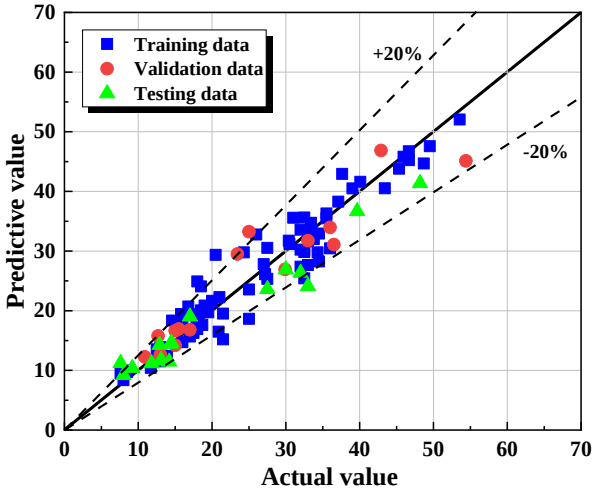
Fig. 6 Coefficient of determination (R^2) of three machine learning algorithms



(a) SVM

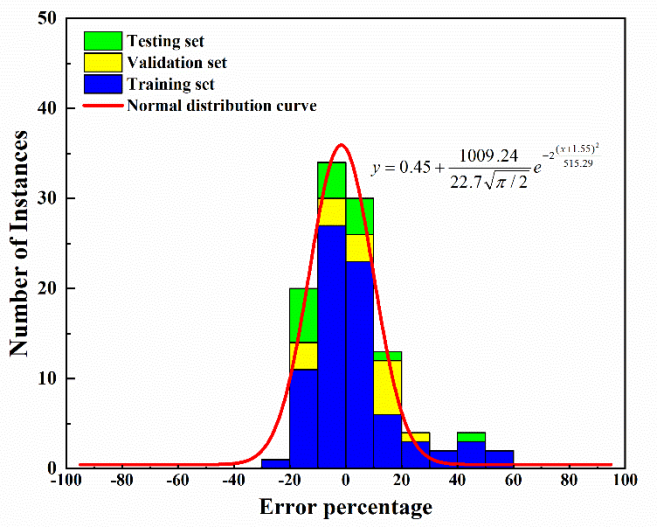


(b) BPNN

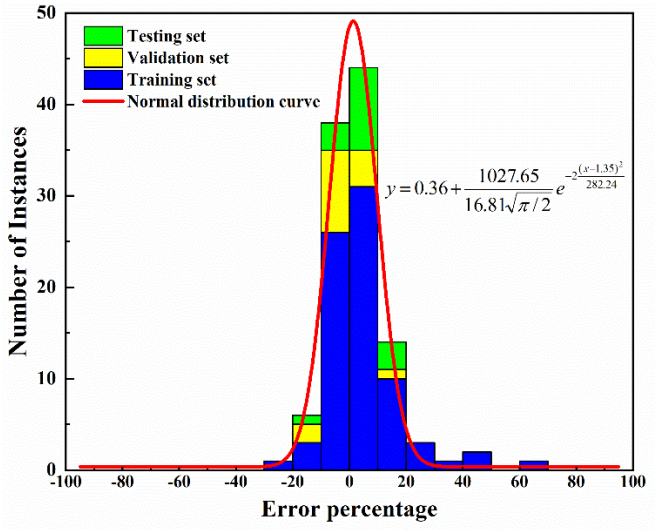


(c) ELM

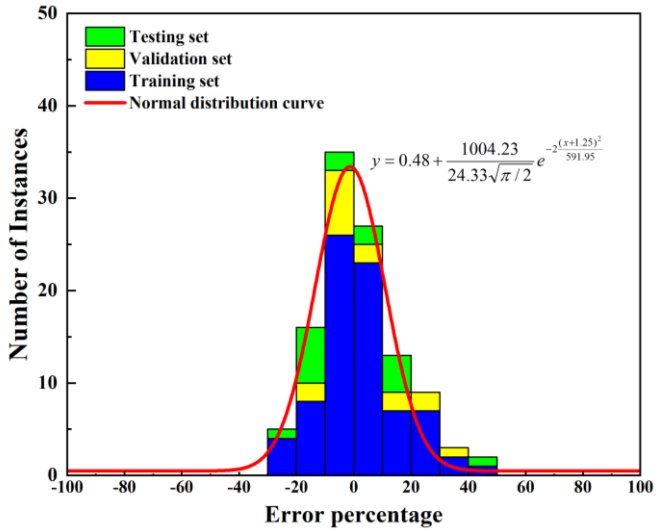
Fig. 7 Measured value vs. predicted value of three machine learning algorithms



(a) SVM



(b) BPNN



(c) ELM

Fig. 8 Error percentage distribution and normal distribution fitting of models

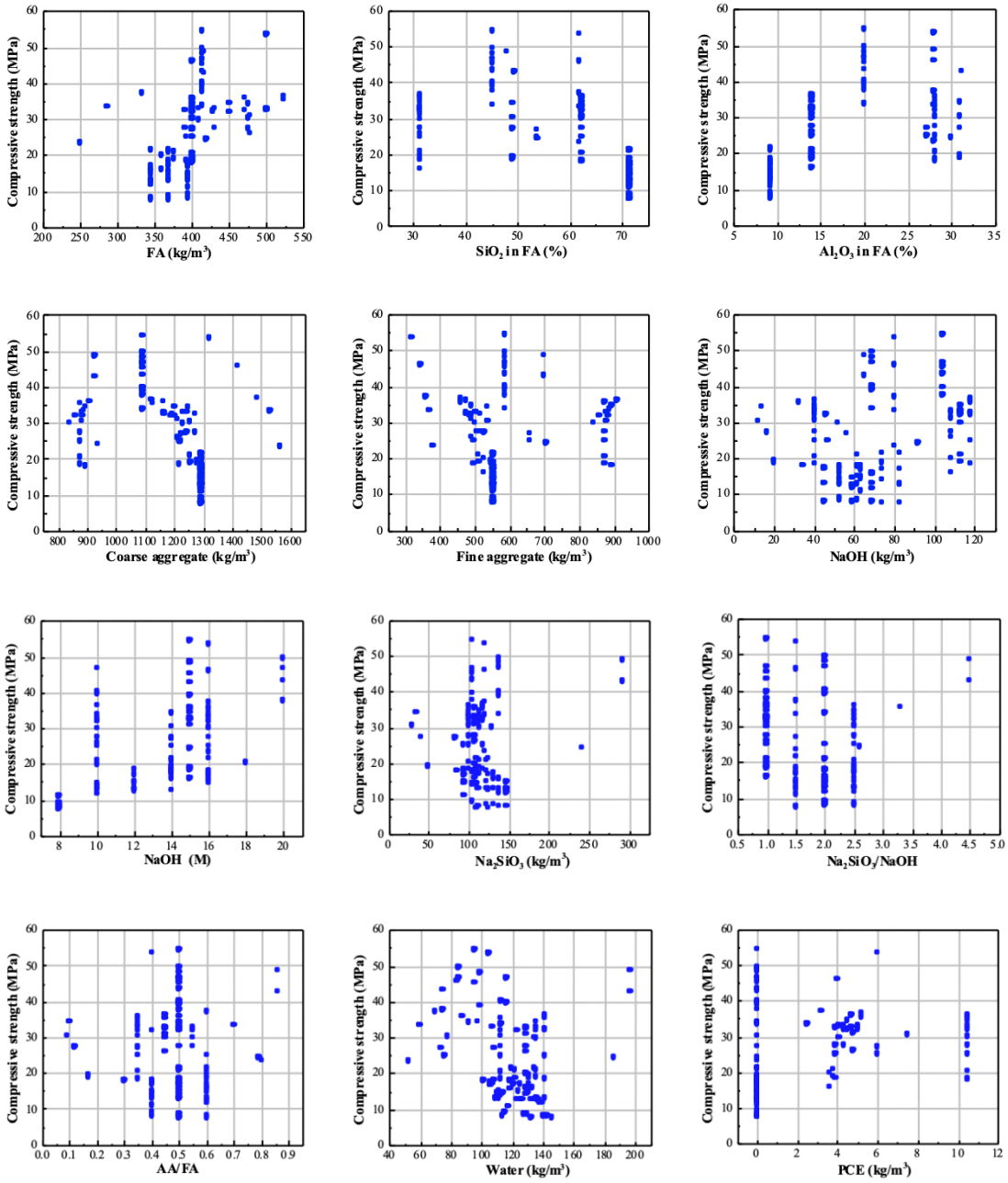


Fig. 9 Correlation between mix design variables and compressive strength

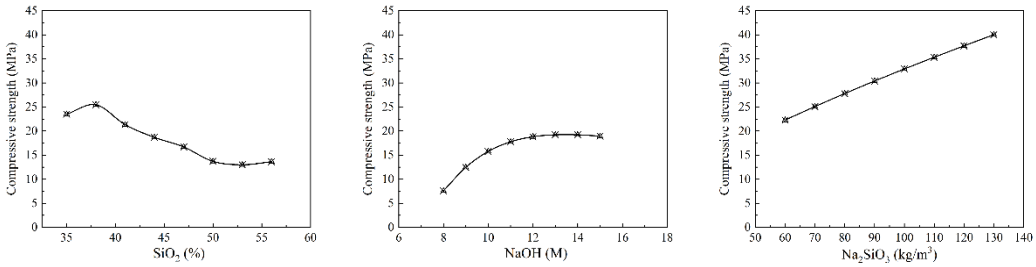


Fig. 10 Sensitivity analysis of feature parameters based on BPNN model