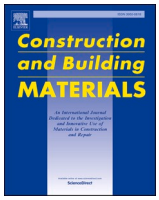
Construction and Building Materials 419 (2024) 135519

Contents lists available at ScienceDirect

Construction and Building Materials

journal homepage: www.elsevier.com/locate/conbuildmat



Review

Machine learning approaches to predict compressive strength of fly ash-based geopolymer concrete: A comprehensive review

Madushan Rathnayaka [a,b], Dulakshi Karunasinghe [b], Chamila Gunasekara [a,*], Kushan Wijesundara [b], Weena Lokuge [c], David W. Law [a]

[a] School of Engineering, Royal Melbourne Institute of Technology (RMIT) University, Melbourne, VIC 3000, Australia
[b] Faculty of Engineering, University of Peradeniya, Sri Lanka
[c] School of Civil Engineering and Surveying, University of Southern Queensland, Springfield, QSL 4300, Australia

ARTICLE INFO

Keywords:
Fly ash
Geopolymer concrete
Machine leaning
Regression analysis
Compressive strength

ABSTRACT

Geopolymer concrete is a sustainable replacement to the Ordinary Portland Cement (OPC) concrete as it mitigates some of the associated problems of OPC manufacturing such as greenhouse gas emission and natural resource depletion. There has been significant recent research in the design of fly ash-based geopolymer concrete using advanced machine learning techniques which can address some of the problems with classical mix design approaches. However, practical application of geopolymer concrete is limited due to lack of standard mix design procedure. This comprehensive review summarizes the current literature on machine learning methodologies to predict the compressive strength of fly ash-based geopolymer concrete. Firstly, the input parameters used for the machine learning model development are categorized based on feature selection or feature extraction. Secondly, available machine learning approaches are categorized based on analysis methods namely, nonlinear regression, ensemble learning, and evolutionary programming. The effect of hyperparameters on the individual model performance, and model comparison based on the prediction performance are also discussed to identify potentially more suitable model type and hyper parameter ranges. Further, the paper discusses the input variable's sensitivity towards the model performance which provides guidance towards future model developments. Overall, this paper will provide an understanding of the current state of machine learning approaches to predict the compressive strength of geopolymer concrete and the gaps in research for the development of models and achieving the required performance. Hence, the summarized knowledge will be highly beneficial to design prospective research towards sustainable cement-free concrete using fly ash.

1. Introduction

About 4 billion tons of cement is manufactured annually and this figure is expected to be increased by 80 million tons each year [1]. Each ton of cement manufactured produces approximately 0.7–1.0 ton of CO_2 [2,3], which represent up to 5–8% of global anthropogenic CO_2 emission [4–8], and pertaining to prevailing conditions, it is expected to rise by 7% in 2050 [9]. Furthermore, 110–120 kWh of energy is invested on one ton of cement produced and collectively [10], it accounts for 10% of the annual industrial energy usage [11]. Utilization of alternatives like alkali activated binders have not only proven a sustainable solution for the aforesaid problems but also provide a way to dispose of industrial waste like silica fume, ground granulated blast-furnace slag (GGBFS), and fly ash [12–14]. Fly ash, the most abundantly used and least utilized aluminosilicate material, is the dust collected from flue gas following coal combustion, and it is one of the most significant solid wastes produced by the coal burning sector. Fly ash particles have a smooth surface and are usually spherical [15]. SiO_2, Al_2O_3, CaO, Na_2O, Fe_2O_3, MgO, K_2O and other metal oxides make up the majority of fly ash. Fly ash also includes trace elements including Cd, Cr, Pb, and Hg [16], as well as trace levels of radioactive elements like ^{226}Ra, ^{232}Th, and ^{40}K [17]. In terms of phase composition, fly ash comprises mullite, quartz, calcite, magnetite, and hematite in addition to a considerable quantity of amorphous glass. Every year, the globe produces more over 800 million tonnes of fly ash [18]. However, a large amount of fly ash is discharged, which causes serious damage to the environment and threatens human health [15]. On the other hand, fly ash-based geopolymer concrete results high early strength and improved durability properties (such as

* Corresponding author.
E-mail address: chamila.gunasekara@rmit.edu.au (C. Gunasekara).

https://doi.org/10.1016/j.conbuildmat.2024.135519
Received 10 June 2023; Received in revised form 3 December 2023; Accepted 18 February 2024
Available online 26 February 2024

reduced permeability, resistance to acid and sulfate attack, and fire resistance) against aggressive environments [19,20]. Thus, the paper primarily focuses on fly ash-based geopolymer concrete in this review.

The reactivity of these fly ash sources is contingent upon their chemical composition [19,21], mineralogical composition [22], morphology [19,23], fineness [24,25], and glassy phase content [26]. Alkali activators such as sodium or potassium hydroxides, silicates, carbonates or a mixture of these influence the geopolymer properties and performance [22]. Due to inconsistency in raw material and alkaline activators, the same formulation can result in a wide range of compressive strength [27]. To date many attempts have been made to design geopolymer mix proportions using classical approaches, which can be classified into four groups, namely the target strength method, factorial model method, performance-based method, and statistical methods. Table 1 shows an updated version of the table found in the review conducted by Li, et al. [27]. According to Li, et al. [27] the target strength method can be further classified based on the fixity of water content, binder content and paste content and used an iterative process of laboratory experiments to acquire the required properties of the geopolymer concrete. The factorial model has utilized the Taguchi method matrix to identify the best combination of input parameters whereas the performance-based method involves extensive laboratory experiments. Furthermore, statistical approaches have been adopted to develop linear and non-linear relationships using the data extracted from open literature to predict the compressive strength of geopolymer concrete, Eq.1-Eq.3. However, extensive laboratory experimentation requires an extensive amount of time and cost [28,29]. The traditional statistical models are inefficient when it comes to evaluating actual scenarios of concrete with diverse constituents, while the findings become inaccurate when new data that differs from the original data is utilized [30,31]. This is primarily due to traditional statistical models being constructed using fixed equations with limited inputs [32].

The application of machine learning in geopolymer concrete to predict compressive strength has emerged recently. The advantage of using machine learning as opposed to empirical models 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 [5,12]. It further offers accurate and generalized predictions, less computational complexity and convenient reproducibility [30,33]. Fig. 1 illustrates the range of machine-learning techniques applied in geopolymer research up to October 2022. The study focusses on machine learning models developed to predict the compressive strength of 100% class F fly ash-based geopolymer concrete and published in peer-reviewed literature articles. A few studies have used multiple machine-learning approaches and selected the best model based on their performance while others have relied on a single technique. Whilst individual efforts have been made to predict the compressive strength using a range of machine learning techniques, there is no in-depth analytical comparison conducted between them with respect to the geopolymer performance. Hence, this comprehensive review will provide clear insight and detailed analysis of existing machine learning approaches used in fly ash-based geopolymers and identify the current research gaps and future development required. The objective of this study is to identify the most influential input variables for the machine learning model development, and to assess the range of machine learning algorithms in terms of their hyperparameters along with their optimum values. Moreover, this will consider various methods of model performance evaluation and the sensitivity of most utilized input variables toward the final model predictions.

The machine learning models shown in Fig. 1 are classified into three groups, namely nonlinear regression analysis, ensemble learning, and evolutionary programming for discussion purposes in this study. It should also be noted that 78% of the models fall into the nonlinear regression category while only 22% fall in to the other two combined.

2. Model inputs and outputs

The quality of the inputs has a significant influence on output accuracy, in addition to decreasing the computation cost of a machine learning model [65]. Two methods namely feature selection, where the influential mix variables were used as direct inputs to the model, and feature extraction, where new features were derived from the influential mix variables, have been adopted for the selection of inputs for the 19 machine learning models discussed in this review [4,5,13,14,33,58–64]. The current review identifies 16 inputs based on feature selection such as NaOH concentration, curing temperature, curing time, etc. and 6 inputs, namely activator / fly ash ratio, Na_2SiO_3 / NaOH ratio, fine aggregate / total aggregate ratio, SiO_2 / water ratio, fly ash/aggregates ratio and water / solid ratio, based on feature extraction, Fig. 2. The mechanism of geopolymer synthesis includes the dissolution of Al and Si in the alkali medium, transportation (orientation) of dissolved species, followed by a polycondensation, forming a 3D network of silico-aluminate structures. The dissolution process is governed by the OH^- provided by the NaOH solution, and thus the concentration of the NaOH solution has been used as a key input variable in more than 80% of machine learning models. Furthermore, Na^+ resulting from NaOH and Na_2SiO_3 balance the negative charge of the Al^{3+} in IV-fold coordination, while SiO_2 from the Na_2SiO_3 controls the Si/Al ratio in the sialates network, and hence the end product [66]. NaOH and Na_2SiO_3 have been used either as weight per meter cube or as a Na_2SiO_3 / NaOH ratio. However, Chu, et al. [62] used SiO_2 resulting from the Na_2SiO_3 and Gomaa, et al. [64] utilized Na_2O derived from the Na_2SiO_3 respectively. Furthermore, Chu, et al. [62] identified both NaOH concentration and SiO_2 resulting from the Na_2SiO_3 has an equal contribution towards the model predictions which indicates that the properties of the activator solution have a great influence on the model predictions.

Fly ash, the sole precursor material in the geopolymer mix controls the availability of SiO_2 and Al_2O_3 for the dissolution. However, geopolymer gel forms on the surface of fly ash particles thereby encapsulating these particles, causing the reaction to cease before it comes to completion [67]. This indicates the importance of the size of the fly ash particle to enable full utilization of SiO_2 and Al_2O_3 in the reaction process. Furthermore, degree of reaction increases with the increased amorphous percentage in the fly ash, because dissolution of amorphous SiO_2 and Al_2O_3 requires less energy due to weaker bonds. As Li et al. noted that the bond energy of Si-O, Al-O and Ca-O are 444, 221–280 and 134 kJ, respectively [68]. Thus, the direct content of fly ash as used in most of the models does not truly represent the actual amount of SiO_2 and Al_2O_3 available for the polymerization. Several attempts have been made to incorporate SiO_2 and Al_2O_3 as separate entities which enables improved ANN model performance [5,64] and further study is required.

High temperature, ranging from 60 ^{0}C to 100 ^{0}C, curing can significantly improve the compressive strength of fly ash geopolymer concrete, especially early strength, by promoting the polymerization reaction of fly ash which clearly illustrates the reason more than 70% of the models used temperature as an input variable. Even though the studies identified that water content in geopolymer mix only contributes to the workability during the handling process [66] excess volume of water cause segregation and strength reduction. Furthermore, it has been identified that sensitivity of water content towards the compressive strength of geopolymer concrete is crucial, means one unit of water can completely change the resultant compressive strength. More than 77% of the models used the water content either as a direct input variable or a ratio with another solid content illustrating its importance towards model predictions.

3. Nonlinear regression analysis

3.1. Artificial neural network (ANN)

The structure of an ANN is generally composed of three types of

Table 1
Summary of mixture design methods for geopolymer concrete in the literature.

Classification	Comments
Target strength method	
1. Fixed water content	• A Ukraine standard RSN 336–84 1984, the first ever standardized method [34] • Fix alkaline activator dosage. Then relationship between the concrete strength and activator dosage was established based on design strength requirement. [35] • Develop a statistical geopolymer mix design method based on ACI 211.1 Fix the alkaline solution content. Then relationship between GPC strength and AL/B was established. [36] • Based on the JGJ 55–2011 method for proportioning PCC with some modification for regression coefficient in Paoluomi formula. [37] • Propose the mixture design by combining concepts of ACI strength versus W/B curve of PCC, absolute volume method and combined grading. [38] • Modify mixture design of the GPC mixtures by updating the AL/B from laboratory experiments based on the ACI method. [39]
2. Fixed binder content	• Based on the IS 10262–2009 method for proportioning PCC with some modification for different grades of GPC. [40] • Variable concrete densities, the ingredients' specific gravities, air volume were considered.[41] • Further simplified the mix design method proposed by Ferdous et al., in 2013. [42] • W/B of 0.35, AL/B of 0.35 and SS/SH ratio of 1.0 by mass were selected based on workability and strength. [43] • Various input variables such as AL/B, W/S, concentration of alkaline solution, curing time and temperature were considered to meet required strength. [44]
3. Fixed paste content	• The influences of specific gravities of ingredients and the air content on GPC densities were not considered in mixture design. [45] • Use discrete element modelling (DEM), the close packing density of aggregate was determined. [46] • W/S was determined by required strength and satisfying setting time. [47] • Combine the close packing of aggregates, excess paste thickness theory and Taguchi method, GPC that meets workability and strength were prepared. [48] • Determine aggregates composition according to packing fraction and paste thickness, while binder content was determined by wet packing fraction. [49]
Performance-based method	• Chloride transport and binding properties of GPC were considered in mix design method. [48]
Factorial method-based models: Taguchi methods	• Aggregate content (1800, 1848, and 1896 kg/m3); AL/B ratio (0.30, 0.35, and 0.40); SS/SH ratio (1.5, 2, and 2.5); Curing method (24 h–60^{0}C, 12 h to 70^{0}C, and 24 h to 75^{0}C). [50] • Oven curing temperature (25, 70, and 90^{0}C); Oven curing time (2, 4, and 8 h); SH concentration (5, 8, and 12 M) [51] • NaOH concentration (5, 8, 12, and 14 M); SS/SH ratio (1, 1.5, 2, and 2.5); AL/B ratio (0.35, 0.38. 0.42, and 0.45); Oven curing time (2, 4, 6, 10 h); Oven curing temperature (25, 40, 70, and 90^{0}C); Water curing regime (1, 3, 7, and 28 d). [52] • OPC content (5, 10, 15, and 20%); NaOH concentration (5, 10, 15, and 20 M); Curing temperature (60, 70, 80, and 90^{0}C). [53] • W/B ratio (0.32, 0.35, 0.38, and 0.41); F/A ratio (35, 37, 39, and 41%); SL/FA ratio (3:7, 4:6, 5:5, and 6:4); Water content (170, 180, 190, and 200 kg/ m3). [54] • Binder content (400,450, and 500 kg/m3); AL/B ratio (0.35, 0.45,and 0.55);SS/SH ratio (1.5, 2, and 2.5); [55] • SH concentration (10, 12, and 14 M) W/B ratio (0.5, 0.45, 0.4, and 0.35); FA/B ratio (0, 15, 30, and 45%); Excess paste thickness (32, 40, 48, and 56 μm); Na2O/binder ratio (4, 5, 6, and 7%); Ms (0.5, 1.0, 1.5, and 2.0). [56]
Statistical methods	• Data extracted from open literature and developed regression models for both low calcium and high calcium fly ash-based alkali activated concrete.[57] $CS = 157.09 - 406.95x\frac{AC}{B} + 51.27x\frac{SiO2_{Fly\ ash}}{Al2O3_{Fly\ ash}} + 28.129 \times \frac{SS}{SH} - 78.811 \times \frac{W}{S} - 0.210x\ C + 1.821x\ M - 9.125x\frac{AC}{B} \times \frac{Si}{Al_2} + 0.345x\frac{AC\ x\ C}{B} + 0.7078x\frac{AC\ x\ Curing\ time}{B}$ Eq.1 • Data extracted from open literature and developed liner and no liner Regression models for low calcium fly ash-based geopolymer concrete [11] $CS = -66.8 - 1.697\frac{Si}{Al} + 187.75\frac{L}{B} + 0.246(FA) - 0.016(F) - 0.012(C) - 0.334(SH) - 0.538(SS) + 0.942\frac{SS}{SH} + 0.179(M) + 0.228(A) + 0.342(T) + 0.01(CD)$Eq. 2 $CS = -1997208 \times \left(\frac{Si}{Al}\right)^{-0.508} \times \left(\frac{L}{B}\right)^{-1.606} \times FA^{-2.134} \times F^{0.016} \times C^{0.089} \times SH^{-0.27} \times SS^{0.274} \times \left(\frac{SS}{SH}\right)^{-0.533} \times M^{0.117} \times A^{-0.305} + 9993.13 \times \left(\frac{Si}{Al}\right)^{-0.423} \times \left(\frac{L}{B}\right)^{-0.068} \times FA^{-0.368} \times F^{-0.151} \times C^{-0.184} \times SH^{-0.426} \times SS^{-0.0007} \times \left(\frac{SS}{SH}\right)^{-0.453} \times M^{0.134} \times A^{-0.022} \times T^{0.352} \times CD^{-0.064}$Eq.3

Note: AL=Alkaline liquid, B= Binder, PCC= Portland cement concrete, W= Water, SS= Sodium silicate, SH= Sodium hydroxide, S= Solids, GPC= Geopolymer concrete, FA= Fly ash, L=Liquid, F = Fine aggregates, C=Coarse aggregates, M= NaOH Molarity, A= Sample age, AC=Activator

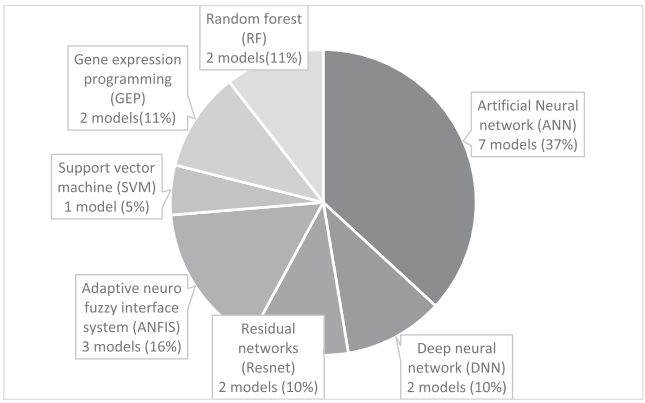


Fig. 1. Machine learning techniques used for the prediction of compressive strength of geopolymer concrete [4,5,13,14,33,58–64].

layers: an input layer, hidden layer(s) an output layer. The input layer conveys input variables for model training validation and testing. The hidden layer(s) is/are responsible for linking between the input layer and the output layer that delivers the result of the model [69]. Activation functions make backpropagation possible as the gradients are supplied along with the error to update the weights and biases. Nonlinear activation functions introduce non-linearity into the output of a neuron [70]. Furthermore, ANN training is achieved via learning algorithms, which enable the model to understand the concept of the problem [71]. Hence, the behavior of ANN changes according to the type of learning algorithm [31]. Table 2 outlines the various ANN approaches employed for estimating compressive strength of geopolymer concrete, which are discussed later.

It can be observed in the Table 2 that ANN has been widely used by researchers to predict the compressive strength of fly ash-based geopolymer concrete. The size of the databases varied between 110 and 298 datasets. All the databases have been developed using the data extracted from the literature except for two models which used data from their own set of laboratory experiments [4,58]. Generally, in making ANN models, the database is divided into training (TR), validation (VAL), and testing (TS) subsets. The training set is used for model training. Validation data provides unbiased evaluation of the model fit on training data and prevents model overfitting by stopping the training process when the error increases. The model is finally applied on the testing data to assess the predictive performance. The most common data division scheme used is 70% of the data for the training dataset while the remaining is allocated to the validation and testing. Models having both validation and testing datasets used equal splits of 15% for each, while models having either validation or test set used the remaining 30%. However, two studies have used 90% [4] and 85% [60] data allocation for the training subset while the remaining 10% and 15% were allocated to the validation subset. Furthermore, models have used data normalization which enables them to overcome distortion effect due to different scales of data. Random sorting has been used to ensure unbiased data splitting between training, validation and test subsets. Peng and Unluer [5] sorted the data randomly and normalized between −1 and 1 while Gunasekara, et al. [13] used 0–1 range for the normalization. Other studies have not reported any data preprocessing before feeding data into the ANN models.

The number of hidden layers and the number of neurons in each hidden layer governs the size of the weight and bias matrixes associated with the ANN model. Each neuron contains its own bias and each connection between neurons has a dedicated weight. While most of the studies including Thanh Pham et al. [60] and Gunasekara et al. [13] have used a single hidden layer in the neural network, Bhogayata et al. [61], Dao et al. [58] and Khan et al. [14] used 3, 5 and 10 hidden layers respectively. Furthermore, the majority of the studies have used less than 20 hidden neurons whereas Bhogayata et al. [61] and Huynh et al. [4] used 60 and 384 hidden neurons respectively. As a remedy for co-adaptation caused by large number of hidden neurons in the network, the drop out method, which removes hidden and visible units from a neural network in a random manner, has been applied [4]. The dropout process disregards neurons during the training phase. At each training step, nodes are dropped out of the network with a probability of 1-p or maintained with a probability of p, leaving a smaller network. Huynh, et al. [4] used a drop out ratio of 0.2 which retained 77 neurons in the trained model out of 384 neurons originally used. By utilizing this dropout, the network overfitting caused by large number of hidden neurons is expected to be avoided.

The effect of learning rate on the ANN model performance has been studied by using two learning rates 0.1 and 0.5 with a 3 hidden layered neural network, each having varied the number of hidden neurons

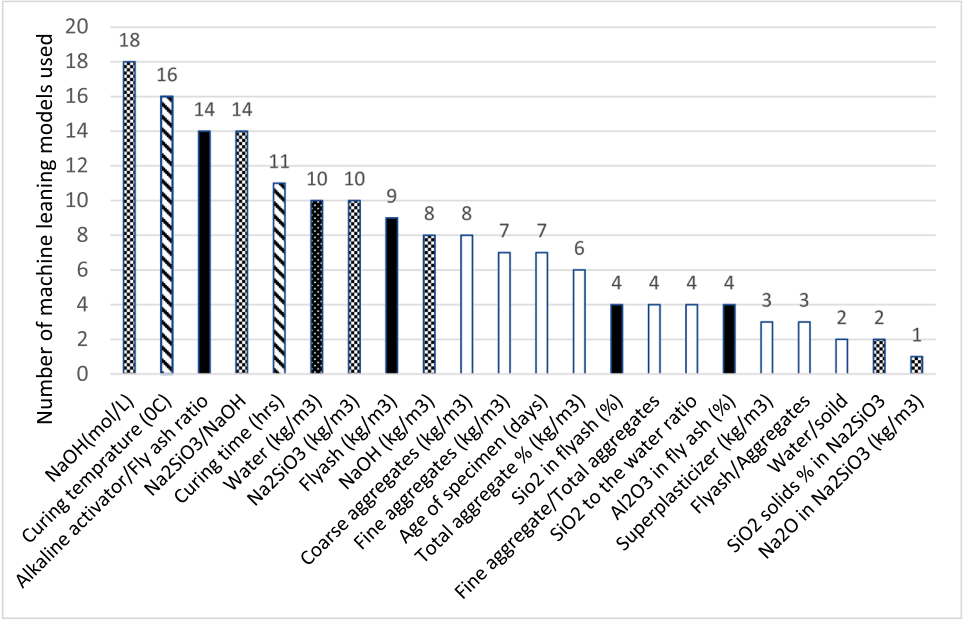


Fig. 2. Input variables used in machine learning models [4,5,13,14,33,58–64].

Table 2
Summary of ANN models.

Data	Model parameters	Performance	Ref
Database size: 150, **Data source**: Literature, **Data division:** Train: 70%, Validation: 15% Test: 15%, **Number of inputs**: 6, **Output**: Compressive strength (20.18–77.00 MPa),	Sigmoidal activation function (SAF). Maximum epochs 10,000. **Details of 8 models** HL 1 HL 2 HL 3 LR 2 2 2 0.1 2 2 2 0.5 6 10 2 0.1 6 10 2 0.5 30 30 30 0.1 30 30 30 0.5 1 30 60 0.1 1 30 60 0.5	Model performance was evaluated based on the laboratory experiments.	[61]
Database size: 263, **Data source**: Experimental, **Data division:** Train: 90%, Validation: 10% Test: 0%, **Number of inputs**: 6, **Output**: Compressive strength (5.44–67.86 MPa)	Rectified liner unit (ReLu). Hidden neurons: 384 **Model 1** Layer normalization: Yes Drop out ratio (p): 0.2 **Model 2** Layer normalization: No Drop out ratio (p): 0.2	Coefficient of determination (R^2) TR: 0.921 VAL: 0.889 Root means square error (RMSE) TR: 3.153 VAL: 4.711 Coefficient of determination (R^2) TR: 0.827 VAL: 0.798 Root means square error (RMSE) TR: 4.802 VAL: 5.059	[4]
Database size: 190, **Data source**: Literature, **Data division:** Train: 85%, Validation: 15% Test: via lab experiments **Number of inputs**: 6, **Output**: Compressive strength (20 – 89 MPa)	Hidden layers: 1 Hidden neurons: 19 Training function: Levenberg-Marquardt	Coefficient of determination (R^2) TR: 0.720 VAL: 0.619 Root means square error (RMSE) TR: 9.04 VAL: 8.97	[60]
Database size: 210, **Data source**: Experimental, **Data division:** Train: 70%, Validation: 0% Test: 30% **Number of inputs**: 4, **Output**: Compressive strength (18.92–74.12 MPa)	Sigmoidal activation function (SAF). Training function: Scaled Conjugate Gradient method	Coefficient of determination (R^2) TR: 0.693 TS: 0.851 Root means square error (RMSE) TS: 2.265	[58]
Database size: 110, **Data source**: Literature, **Data division:** Train: 70%, Validation: 15% Test: 15% **Number of inputs**: 14, **Output**: Compressive strength (min, max)	Hidden layers: 1 Hidden neurons:18 Sigmoidal activation function (SAF).	Coefficient of determination (R^2) TR: 0.947 VAL: 0.822 TS: 0.915 Root means square error (RMSE) TR: 2.5 VAL: 3.68 TS: 1.95	[5]
Database size: 298, **Data source**: Literature, **Data division:** Train: 70%, Validation: 30% Test: 0% **Number of inputs**: 10, **Output**: Compressive strength (8.2 – 63 MPa)	Hidden layers: 10 Hidden neurons: Each layer maximum up to 10 neurons. Training function: Levenberg-Marquardt Maximum epochs: 100	Coefficient of correlation (R) TR: 0.849 VAL: 0.931 Root means square error (RMSE) TR: 6.02 VAL: 4.93	[14]
Database size: 166, **Data source**: Literature, **Data division:** Train: 70%, Validation: 0% Test: 30% **Number of inputs**: 4, **Output**: Compressive strength (12.0–80.4 MPa)	Hidden layers: 1 Hidden neurons: 8 Training function: Bayesian Regularization	Coefficient of correlation (R) TR: 0.86 TS: 0.86 Root means square error (RMSE) TR: 0.006	[13]

Note: HL1: Hidden layer 1, HL2: Hidden layer 2, HL3: Hidden layer 3, TR: Training, VAL: Validation, TS: Test

between 1 and 60 [61], Table 2. Furthermore, Thanh Pham [60] and Gunasekara, et al. [13] studied the effect of training function by developing multiple models with distinct training functions. Thanh Pham [60] identified the Levenberg-Marquardt training function outperforms the other seven training functions used, namely, resilient backpropagation, Polak-Ribiere conjugate gradient, conjugate gradient with Powell / Beale restarts, BFGS quasi-newton, one step secant, Fletcher-powell conjugate gradient, adaptive learning rate backpropagation, Bayesian regularization and scaled conjugate gradient, while Gunasekara, et al. [13] identified Bayesian regularization as the best training function compared to Levenberg-Marquardt and scaled conjugate gradient. Furthermore, the selection of the layer activation function plays a crucial role in ANN model as the weighted inputs are passed to the next level through activation functions. The different activation functions used in ANN models are summarized in the Table 2. Sigmoid activation, linear activation and rectified linear unit functions have been used, where the linear activation function was used in the output layer of all the models while either sigmoid activation or rectified liner unit were used in the hidden layers.

Performance assessment of machine learning algorithms has been carried out using statistical methods that describe the model fitting. Table 3 gives statistical metrics employed for evaluating ML models with their corresponding mathematical expressions. These methods indicate how well the predicted values fit with actual data. Moreover, they can be adopted in sensitivity analysis, pointing out the weight of each input variable in the prediction process. Not only can statistical metrics assess the performance of machine learning techniques, but they may also be used as reference for comparing the effectiveness of several algorithms. Huynh, et al. [4] used coefficient of determination (R^2) to determine the best model out of the two models developed while Gunasekara, et al. [13] used mean squared error for the evaluation of best training algorithm. Coefficient of determination (R^2), root mean squared error (RMSE) and mean squared error (MSE) have been commonly used to assess the model performance [4,5,13,58,60] while other statistical parameters [14] were seldomly used between the models. Further to statistical measures, some studies performed laboratory experiments and compared the predicted compressive strength with the actual strength resulted from the laboratory experiments [61].

Table 3
Statistical metrics used for the performance evaluation.

Statistical parameter	Formula	Ref
Coefficient of determination	$R^2 = \frac{\left(n\sum_i y_i y_i' - \sum_i y_i' \sum_i y_i\right)^2}{\left(n\sum_i y_i'^2 - \sum_i y_i'\right)^2 \left(n\sum_i y_i^2 - \sum_i y_i\right)^2}$	[4]
	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y_i')^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$	[5,58, 60]
Coefficient of correlation	$R = \frac{\sum_{i=1}^{n}(y_i - \bar{y_i})(y_i' - \bar{y_i'})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y_i})^2 \sum_{i=1}^{n}(y_i' - \bar{y_i'})^2}}$	[13, 14]
Mean absolute percentage error	$MAPE = \frac{1}{n}\sum\left\lvert\frac{y_j - y_j'}{y_j}\right\rvert \times 100$	[4]
Root mean square error	$RMSE = \sqrt{\frac{1}{n} \times \sum_{j=1}^{n}\left(y_j - y_j'\right)^2}$	[4,5, 58,60]
Mean square error	$MSE = \frac{1}{n} \times \sum_{j=1}^{n}\left(y_j - y_j'\right)^2$	[5,13, 60]
Mean absolute error	$MAE = \frac{1}{n} \times \sum_{i=1}^{n}\lvert y_i - y_i'\rvert$	[5,58]
Relative squared error	$RSE = \frac{\sum_{i=1}^{n}(y_i' - y_i)^2}{\sum_{i=1}^{n}(\bar{y_i} - y_i)^2}$	[14]
Relative root mean squared error	$RRMSE = \frac{1}{\hat{y}} \times \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_i')^2}{n}}$ $where \hat{y} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y_i})^2}{n}}$	[14]
Performance index	$\rho = \frac{RRMSE}{1 + R}$	[14]

Note: y_i: Actual compressive strength, y_i' :Predicted compressive strength, $\bar{y_i}$: Mean of actual compressive strength, $\bar{y_i'}$: Mean of predicted compressive strength.

3.2. Deep neural networks (DNN) and residual networks (Resnet)

DNN is another approach used to predict the compressive strength of geopolymer concrete which consists of more layers and neurons compared to the ANN, leading to its ability to learn functions with a high degree of complexity. However, according to the authors [4,63] DNN has its own limitations in training a very deep network to further improve its accuracy. DNN are trained through a backpropagation process that relies on gradient descent, shifting down the loss function and finding the weights that minimize the output. If there are too many layers, repeated multiplications will eventually reduce the gradient until it "disappears", and performance saturates or deteriorates with each layer added. To overcome the problems associated with DNN, ResNet was developed [72]. ResNet models with modified architectures have been empirically demonstrated to improve the learnability of neural networks with less error on predefined tasks. Fig. 3 depicts ResNet's residual blocks with shortcut connections, where H(x) is the desired mapping output of a certain layer and x is the input data. Skip connections behave as shortcuts that skip 2 or 3 layers (Fig. 3). This shortcut will change how the gradient is calculated at each layer. This solves the problem of vanishing gradient in DNN by allowing this alternate shortcut path for the gradient to flow through. Furthermore, it allows the model to learn the identity functions which ensures that the higher layer will perform at least as good as the lower layer, and not worse [4].

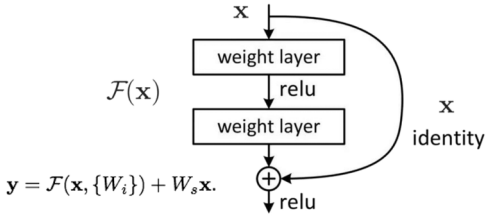


Fig. 3. A block in a deep residual network [72].

Two studies have used both DNN and Resnet to predict the compressive strength of geopolymer concrete. It can be observed from Table 4 that both the studies have performed a set of laboratory experiments to develop their databases of sizes 335 [63] and 263 [4] and used 80:20 and 90:10 training to validation ratios respectively. However, none of the studies have used test data set to assess the generalization capability of the models. Furthermore, models have used 9 and 6 input variables while their compressive strength had same ranges of 5.44 MPa to 67.86 MPa.

The number of hidden layers in DNN were limited to two, whereas Resnet used an additional weight which make total number of hidden layers in to three to make the element-wise addition operation at the end of the architecture. The hidden neurons in each hidden layer varied between 128 and 300 in DNN while Resnet repeat the same number of hidden neurons in the layer 2 and the layer 3. Rectified linear unit (ReLu) has been used as the layer activation function in all both DNN and Resnet models. Furthermore, the effect of layer normalization has been assessed for both DNN and Resnet models by switching layer normalization between on and off state. The effect of drop out was also assesd for both DNN and Resnet models using a dropout ratio 0.7 and comparing it to the model without applying dropout [63], while another study used a constant dropout value of 0.2 for all the models [4]. Both studies [4,63] used Adam as the optimizer for both DNN and Resnet models, as the authors stated it inherited various good features from optimizers such as AdaGrad and RMSProp. With Adam, an update is ensured in every step, to modify network parameters in the direction influenced by historical gradients. This makes it less biased by the current learning sample or minibatch and keep the learning effective when there are sparse gradients.

Layer normalization enables the DNN to overcome issues such as longer training time and instability caused by accumulation of large error gradients which is termed as exploding gradient. Nguyen, et al. [63] observed an increase in coefficient of correlation of the training dataset from 0.873 to 0.951 while validation performance increased from 0.867 to 0.954 as layer normalization is introduced to the model. Huynh, et al. [4] observed similar improvement in a layer normalized model with a 12% increment training results and 7% increase in validation results, Table 4. Dropout reduces both training and validation model performance. Nguyen, et al. [63] observed a 4% and 3% reduction in the training and validation performance for the model with both layer normalization and dropout, while a higher reduction was observed for the same model without the layer normalization and with dropout, in range of 10% for both training and validation.

Unlike the DNN, the effect of layer normalization on the performance of the Resnet model seemed to be ineffective. According to the observations of Nguyen, et al. [63] training performance of the ResNet model was slightly increased by 0.5% with layer normalization while validation performance remained unchanged, Table 4. However, Huynh, et al. [4] noticed about 2.5% increment in both training and validation performance with the layer normalization. Dropout in the Resnet behaved in a similar manner as in the DNN, where the training and validation performance of the model decrease by 4% for both models either with or without the layer normalization.

3.3. Adaptive neuro-fuzzy interface system (ANFIS)

ANFIS is an artificial intelligence technique which combines ANN with fuzzy logic [58] and consists of five layers, namely fuzzification layer, rule layer, normalization layer defuzzification layer and summation layer [73]. ANFIS models have been modeled using databases of size 90 [33], 298 [14] and 210 [58]. Different data division schemes have been adopted for the selection of training, validation and testing datasets. Ahmad, et al. [33] used 80:20 train to test data splitting while

Table 4
Summary of DNN and Resnet models.

Data	Model parameters	Performance	Ref
DNN			
Database size: 335, **Data source**: Experimental, **Data division:** Train: 80%, Validation: 20% Test: 0%, **Number of inputs**: 9, **Output**: Compressive strength (5.44–67.86 MPa)	Hidden layers: 2 Hidden neurons: L1 300, L2: 200 Rectified liner unit (ReLu).		[63]
	Model 1 Layer normalization: Yes Drop out ratio: 0.7	Coefficient of correlation (R) TR: 0.951 VAL: 0.954 Root means square error (RMSE) TR: 2.39 VAL: 3.14	
	Model 2 Layer normalization: Yes Drop out ratio: 0	Coefficient of correlation (R) TR: 0.993 VAL: 0.985 Root means square error (RMSE) TR: VAL: 1.82 1.47	
	Model 3 Layer normalization: No Drop out ratio: 0.7	Coefficient of correlation (R) TR: 0.873 VAL: 0.867 Root means square error (RMSE) TR: VAL: 4.80 4.96	
	Model 4 Layer normalization: No Drop out ratio: 0	Coefficient of correlation (R) TR: 0.968 VAL: 0.960 Root means square error (RMSE) TR: VAL: 2.71 2.43	
Database size: 263, **Data source**: Literature, **Data division:** Train: 90%, Validation: 10% Test: 0% **Number of inputs**: 6, **Output**: Compressive strength (5.44–67.86 MPa)	Hidden layers:2 Hidden neurons: L1:128 L2: 256 Rectified liner unit (ReLu).		[4]
	Model 1 Layer normalization: Yes Drop out ratio: 0.2	Coefficient of Determination (R^2) TR: 0.923 VAL: 0.898 Root means square error (RMSE) TR: VAL: 2.52 3.27	
	Model 2 Layer normalization: No Drop out ratio: 0.2	Coefficient of Determination (R^2) TR: 0.821 VAL: 0.840 Root means square error (RMSE) TR: VAL: 3.16 5.01	
Resnet			
Database size: 335, **Data source**: Experimental, **Data division:** Train: 80%, Validation: 20% Test: 0%, **Number of inputs**: 9, **Output**: Compressive strength (5.44–67.86 MPa)	Hidden layers: 3 Hidden neurons: L1 300, L2: 200 L3 200 Rectified liner unit (ReLu).		[63]
	Model 1 Layer normalization: Yes Drop out ratio: 0.7	Coefficient of correlation (R) TR: 0.958 VAL: 0.942 Root means square error (RMSE)	

Table 4 (*continued*)

Data	Model parameters	Performance	Ref
		TR: 2.66 VAL: 2.18	
	Model 2 Layer normalization: Yes Drop out ratio: 0	Coefficient of correlation (R) TR: 0.998 VAL: 0.986 Root means square error (RMSE) TR: 1.30 VAL: 1.28	
	Model 3 Layer normalization: No Drop out ratio: 0.7	Coefficient of correlation (R) TR: 0.954 VAL: 0.942 Root means square error (RMSE) TR: 2.61 VAL: 1.95	
	Model 4 Layer normalization: No Drop out ratio: 0	Coefficient of correlation (R) TR: 0.994 VAL: 0.992 Root means square error (RMSE) TR: 1.31 VAL: 1.26	
Database size: 263, **Data source**: Literature, **Data division:** Train: 90%, Validation: 10% Test: 0% **Number of inputs**: 6, **Output**: Compressive strength (5.44–67.86 MPa)	Hidden layers: 3 Hidden neurons: L1:128 L2: 256 L3: 256 Rectified liner unit (ReLu).		[4]
	Model 1 Layer normalization: Yes Drop out ratio: 0.2	Coefficient of Determination (R^2) TR: 0.921 VAL: 0.915 Root means square error (RMSE) TR: 3.24 VAL: 3.28	
	Model 2 Layer normalization: No Drop out ratio: 0.2	Coefficient of Determination (R^2) TR: 0.896 VAL: 0.937 Root means square error (RMSE) TR: 3.81 VAL: 1.98	

Note: TR: Train, VAL: Validation

Dao, et al. [58] used 70:30. However, Khan, et al. [14] used the same ratio of Dao, et al. [58], (70:30) data division scheme for training and validation data splitting. Furthermore, the number of model input variables ranged from 4 to 10 while their compressive strength values in the databases ranged between 8.2 MPa to 89 MPa, Table 5.

Three models [14,33,58] have been developed in the studies using two member functions in the fuzzification layer, Triangular membership function [33] and Gaussian membership function [14,58]. The second layer of all the models have used the rule of IF-THEN.A single study reported details on the output member function as constant function.

Performance of the models were evaluated using either coefficient of correlation (R) [14] or coefficient of determination (R^2) combined with other statistical parameters such as mean absolute error (MAE) and root mean square error (RMSE). Other than these common statistical measure, Khan, et al. [14] have used additional measure such as relative squared error (RSE), relative root mean squared error (RRMSE), and the performance index (ρ), Table 3.

Furthermore, all the ANFIS model performance have been compared with some other machine learning models. According to Ahmad, et al. [33], ANFIS model outperform the Multivariate adaptive spline model

Table 5
ANFIS model details.

Data	Model parameters	Performance	Ref
Database size: 90, **Data source**: Literature, **Data division**: Train: 80%, Validation: 0% Test: 20%, **Number of inputs**: 4, **Output**: Compressive strength (20.0 – 89.00 MPa)	**Membership function**: Input membership function: triangular MF Output MF: Constant. **Rule layer**: Rule of IF-THEN **Optimization method**: Hybrid optimization. **Epochs**: 500	Coefficient of determination (R^2) TR 0.74 TS 0.90 RMSE TR 2.81 TS 5.53 MAE TR 0.35 TS 1.11	[33]
Database size: 298, **Data source**: Literature, **Data division**: Train: 70%, Validation: 30% Test: 0% **Number of inputs**: 10, **Output**: Compressive strength (8.2 – 63 MPa)	**Membership function**: Input membership function: Gaussian MF **Rule layer**: Rule of IF-THEN **Epochs**: 4	R TR 0.925 VAL 0.978 RMSE TR 4.08 VAL 2.59 MAE TR 3.28 VAL 2.08	[14]
Database size: 210, **Data source**: Experimental, **Data division**: Train: 70%, Validation: 0% Test: 30% **Number of inputs**: 4, **Output**: Compressive strength (18.92–74.12 MPa)	**Membership function**: Input membership function: Gaussian MF **Rule layer**: Rule of IF-THEN	Coefficient of determination (R^2) TS 0.879 MAE TS 1.655 RMSE TS 2.265	[58]

Note: TR: Train, VAL: Validation, TS: Test

(MARS) in terms of both training and testing, Table 5. Khan, et al. [14] compared the ANFIS model with ANN and Gene expression programming (GEP) model where he observed similar improvement in model performance in all 5 statistical measures used for both training and validation. Similar observations were made by Dao, et al. [58] where he compared the ANFIS model with ANN model and achieved performance improvement of 3.3% for the test data in terms of coefficient of determination.

3.4. Support vector machine (SVM)

It was observed that SVM has been used in a single study to predict the compressive strength of geopolymer concrete. This model developed by Peng and Unluer [5] used a database with 110 datasets and 14 inputs to develop a SVM model and compared the results with the models developed using back propagation neural network (BPNN) and extreme leaning machine (ELM). Collected data were shuffled randomly, and all input variables were preprocessed by the same normalization algorithm to unify their dimensions. Moreover, a data division scheme of 70%, 15% and 15% was adapted for training, validation, and testing data selection.

Kernel function was used as RBF and e-SVR was used as SVM type, while tolerance and epsilon values were maintained at 0.001 and 0.1. The performance of the developed model was evaluated using four statistical measures, coefficient of determination (R^2), mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE).

It has been observed that the performance of the SVM (R^2= 0.903) had a lower value for its training dataset in terms of R^2 compare to the BPNN (R^2= 0.947) and ELM (R^2= 0.922). However, the resultant R^2 value of validation and testing SVM model (R^2_{Val} = 0.92, R^2_{test} = 0.955) outperformed validation and testing of the other two models (BPNN, R^2_{Val} = 0.822, R^2_{test} = 0.915, ELM, R^2_{Val}0.893, R^2_{test} = 0.914).

4. Evolutionary programming and ensemble learning

4.1. Gene expression programming (GEP)

A strong soft computing technique, namely, genetic programming (GP), is valuable as it ignores the previous forms of established relationships for the development of the model. An extension of GP, namely, GEP, which encodes a small program and uses fixed-length linear chromosomes, was recently introduced to predict the compressive strength of geopolymer concrete [74]. The architectural design of GEP model is determined by the number of genes and size of head, which latter determines the complexity of each expression by summation of sub-expression trees (sub-ETs) of the model. Genes value determines the number sub-ETs developed by manipulating the input variables with the functions in the function set while head size defines the maximum number of starting nodes in a sub-ET. Final model output is achieved by combining the sub-ETs using the linking function [75].

Two GEP models have been developed using databases of size 298 and 311 as shown in Table 6. The datasets were extracted from the literature. Furthermore, the database consists of 10 and 8 input variables while the range of compressive strength in the databases varied between 8.2 and 63 MPa and 10.5–63 MPa respectively. Model training was done using 70% of the total database while the remaining 30% was used for the model validation. Both the models did not consider assessing the model generalization using a test data split. Both models used an identical set of model parameters except for number of genes and function set. Khan, et al. [59] used 4 genes while Chu, et al. [62] used 3 genes for the model development. Furthermore, both models used addition, subtraction multiplication and division as the common functions where Khan, et al. [59] used cubic root as an additional function in their model.

GEP has an advantage over the other machine learning models as it provides a simple mathematical expression which can represent the outcome that is appropriate for practicable usage for better predictive accuracy. Based on 4 genes Ali Khan, et al. [12] developed an equation to predict the compressive strength as a combination of four equations Eq.4 to Eq.8 where T, A,M, %S/W, N_s/N_o, $\%A_G$, $F/A_{G,}$. A_L/F_A, %P, $\%E_W$ denote temperature required for curing of the sample, age of the sample, the molarity of the sodium hydroxide (NaOH) solution used, the percentage of silicon dioxide (SiO_2) to the water ratio for preparing sodium silicate (Na_2SiO_3) solution, the ratio between sodium silicate (Na_2SiO_3) solution to NaOH, the percentage by volume of total aggregates, the ratio between fine aggregate to total aggregates, the ratio between alkali to fly ash, percentage of plasticizer, and percentage of extra addition of water, respectively.

$$f_c^{'}(MPa) = A \times B \times C \times D \tag{4}$$

$$A = \sqrt[3]{\frac{S}{W}}\% - \mathrm{P}\% + \left(M \times \frac{F}{A_G} \times \frac{A_L}{F_A} \times 6.61\right) + E_W\% - A_G\% \tag{5}$$

$$B = -\sqrt[3]{\frac{A+80}{0.083(T-17.87)} + M + \frac{N_S}{N_O}} \tag{6}$$

Table 6
GEP model details.

Data	Model parameters	Performance				Ref
Database size: 298, **Data source**: Literature, **Data division**: Train: 70%, Validation: 30% Test: 0%, **Number of inputs**: 10, **Output**: Compressive strength (8.2–63 MPa)	**Number of chromosomes** 150 **Number of genes** 4 **Head size**: 10 **Linking function** Multiplication **Functions set** +, −, /, ×,$\sqrt[3]{}$ **Constants per gene** 10 **Type of data** Floating data **Upper bound value** 10 **Lower bound value** – 10 **Mutation rate** 0.001380 **Inversion rate** 0.005460 **IS transportation rate** 0.005460 **RIS transportation rate** 0.005460 **Gene recombination rate** 0.007550 **Gene transportation rate** 0.002770 Shape normalization done	Coefficient of correlation TR 0.858 VAL 0.964 RMSE TR 5.971 VAL 2.643 MAE TR 5.823 VAL 2.057 RSE TR 0.325 VAL 0.0675 RRMSE TR 16.949 VAL 4.949 Performance index TR 0.091 VAL 0.025				[59]
Database size: 311, **Data source**: Literature, **Data division**: Train: 70%, Validation: 30% Test: 0%, **Number of inputs**: 8, **Output**: Compressive strength (10.5–63 MPa)	**Number of chromosomes** 150 **Number of genes** 3 **Head size**: 10 **Linking function** Multiplication **Functions set** +, −, /, ×, **Constants per gene** 10 **Type of data** Floating data **Upper bound value** 10 **Lower bound value** – 10 **Mutation rate** 0.001380 **Inversion rate** 0.005460 **Gene transportation rate** 0.002770 **One-point and two-point recombination rate** 0.00277	Coefficient of correlation TR 0.864 VAL 0.984 RMSE TR 5.901 VAL 1.521 MAE TR 5.624 VAL 1.319 RSE TR 0.253 VAL 0.0315 RRMSE TR 15.495 VAL 3.895 Performance index TR 0.083 VAL 0.019				[62]

Note: TR: Train, VAL: Validation, TS: Test

$$C = \frac{F}{A_G} - \left(E_W\% \times M - \frac{0.0003}{\frac{N_S}{N_O} - E_W\%} \right) - 0.0003 \tag{7}$$

$$D = \sqrt[3]{\frac{\left(P\% - \frac{S\%}{W\%}\right)1.16}{T}} + \sqrt[3]{\frac{0.17}{\frac{F}{A_G}}} + 0.77 \tag{8}$$

According to the 3 genes employed by Chu, et al. [62] the equation for compressive strength prediction consists of three equations Eq.9 to Eq.12 where T, t, A, %A_g, M, %S, %P, %E_w denote initial temperature required in the curing regime, time for curing in hours, age of samples in days, percentage of total aggregate by volume, molarity of sodium hydroxide (NaOH) solution, SiO_2 solids percentage in sodium silicate (Na_2SiO_3) solution, superplasticizer and extra water as percent FA respectively.

$$f_c^{'}(MPa) = E \times F \times G \tag{9}$$

$$E = 4.1 - \frac{18.46}{8.94 + P\%} \times \frac{S\% + E_w\%}{A_g - S\%} \tag{10}$$

$$F = 1 + \frac{(2 \times A) - 131.23}{T \times t} \tag{11}$$

$$G = \frac{T - (2 \times M)}{9} - \frac{2}{M - 9} + 11.91 \tag{12}$$

Compared to other methods such as liner regression, nonlinear regression and multi expression programming (MEP) that provide correlations to predict the properties of concrete, GEP has better prediction accuracy both in training and validation, Table 7. Based on the observations of Ali Khan, et al. [12], coefficient of correlation (R) for the training of the GEP model is 0.858 and it is 6% higher than the linear model and 3% higher than the non-linear model, Table 7. Furthermore, GEP model recorded validation performance of 0.963 and 7% and 4% higher accuracy than the linear and nonlinear models. Similar results were observed by Hong-Hu Chu a et al. [62] where it achieved 0.864 and 0.984 for training and validation performance, which is 4% and 9% higher than the MEP model respectively.

Based on Table 7 it can be observed that both models resulted higher training error compared to validation error of the corresponding model which is the opposite of the general behavior of the machine learning models.

4.2. Random Forest (RF)

Random forest (RF) has also been adopted in two studies as a forecasting tool. The RF technique is comprised of three main steps that included the assembling of trained regression trees via a training dataset, the calculation of the mean value of a single regression tree outcome, and the validation of predicted results via a validation dataset [59]. Optimization algorithms used in the RF models ensure that inputs-output correlations are properly established, outliers are

Table 7
Performance comparison.

Author	Model	Train			Validation		
		R	MAE	RMSE	R	MAE	RMSE
Ali Khan, et al. [12]	GEP	0.858	5.823	5.971	0.964	2.057	2.643
	Linear	0.807	6.543	6.986	0.897	4.967	5.546
	Non-Linear	0.835	6.053	6.593	0.924	4.875	5.054
Hong-Hu Chu a et al. [62]	GEP	0.864	5.624	5.901	0.984	1.319	1.521
	MEP	0.831	5.632	6.605	0.938	2.330	3.013

accounted for, and variance and bias among model trees are kept as low as possible. During the study of Gomaa, et al. [64] a database of 180 unique records with 20 input variables was randomly split in to 75% and 25% as training and testing. Furthermore, grid search optimization method was adopted and identified 500 no. of trees and 9 no. of leaves per tree as the optimum values. Further the study identified, when the number of splits was fewer than nine, the logical divisions in the databases were numerically insufficient and too basic to completely incorporate the complicated connections between inputs and outputs. When the number of splits was greater than nine, the complexity of the trees (CARTs) increased the probability of bias, which led to overfitting. Similarly, when the number of trees was fewer than 500, the RF model lacked sufficient bootstraps to make reliable predictions (for new alkali activated concrete compositions). However, when an excessive number of trees (more than 500) were utilized, the computational complexity of the model unquestionably grew.

However, during the study of Khan, et al. [59], original trained set was used to calculate a new trained dataset comprised of boot-strap data. In this step, some of the data points were removed and swapped with the present data points. The removed data points were assembled into other datasets and were called out-of-bag data points. Then, the regression function was estimated using 2/3 of the data points, and the out-of-bag data points were used in validating the model. This process was continued till the required accuracy was achieved. Use of optimization algorithms enables the RF regressor to achieve higher model performance [1]. Coefficient of correlation values of 0.9826 and 0.9943 were observed for training and validation for the model developed by Khan, et al. [59] while Gomaa, et al. [64] achieved 0.96 for the training and 0.97 for the testing. Furthermore, studies revealed that RF regressor outperformed some of the prediction models such as GEP, Linear regression and non-linear regression with improved results for training giving a 14%, 22% and 18% increase and 3%, 11% and 8% for validation.

5. Parameter sensitivity analysis

Sensitivity analysis is used to evaluate how input variables affect output variation derived by machine learning models. Four approaches have been used to detect the parameter sensitivity. The first approach calculated the sensitivity percentage using Eq.13 and Eq.14, where $f_{min}(y_j)$ and $f_{max}(y_j)$ are the j^{th} minimum and maximum predictive model output, respectively, while input values are kept constant at the mean value. N_j gives the range of j^{th} input variable by taking the difference between $f_{max}(y_j)$ and $f_{min}(y_j)$ [4,59,62].

$$N_j = f_{\max}(y_j) - f_{\min}(y_j) \tag{13}$$

$$Sensitivity\% = \frac{N_j}{\sum_{i=1}^{n} N_j} \tag{14}$$

The second approach was permutation feature importance (PFI). 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. Based on the definition, the calculation expression of the PFI value for a specific variable obtained by using mean absolute error (MSE) as error measurement function is shown in Eq.15, where MAE_{perm} and MAE_{orig} are mean absolute error before and after randomly adjusting *X*i sequence, respectively [5].

$$PFI = MAE_{perm} - MAE_{orig} \tag{15}$$

The third approach simulated the model by choosing the number of input variables for the prediction as m-1, where the m denotes the number of input variables, and the value of the excluded parameter is set to 0. Simulation performance was measured in terms of coefficient of determination (R^2) and the results were compared [58]. The fourth approach used a decision-tree structure of the RF model, after it being trained, and fully validated against the test set, which was analyzed to estimate and rank each attribute's importance in accordance with the magnitude of influence it exerts on compressive strength and is presented as "importance" which is a dimensionless parameter [64].

Parameter sensitivity studies on geopolymer concrete, Table 8, identified that fly ash/ aggregates ratio is more sensitive towards the compressive strength prediction as it determines the internal void structure of geopolymer concrete. Huynh, et al. [4] observed the highest sensitivity percentage, 35.55% for fly ash / aggregates ratio, which was 119.17% higher than the second highest sensitivity parameter in the same study, Table 8. However, the individual effect of coarse aggregates, fine aggregates and fly ash found to be less sensitive towards the output predictions. Peng and Unluer [5] noticed least PFI values of 0.03 and 0.02 for coarse aggregates and fine aggregates out of 14 input variables used, while Gomaa, et al. [64] observed third and second lowest "importance" values 58 and 59 for coarse aggregates and fine aggregates compared to the 15 input variables used. Furthermore, studies of Dao, et al. [58] and Peng and Unluer [5] noticed a lower sensitivity of the fly ash content to the output results with a R^2 value of 0.839 and PFI of 0.03 which are 1.14% and 95.7% lower than the highest values.

R^2 value of 0.839 and PFI of 0.03 which are 1.14% and 95.7% lower

Table 8
Input parameter sensitivity analysis results.

Input parameter	Sensitivity indices					
	Sensitivity %			PFI	R^2	Importance
	[62]	[4]	[59]	[5]	[58]	[64]
Curing Temperature (^{0}C)	26.25	16.22	25.30	0.07	-	483
Curing Time (hrs.)	11.58	14.93	-	0.09	-	301
Age of specimen (days)	9.00	-	5.88	-	-	370
NaOH molarity (mol/L)	9.75	16.18	2.32	0.25	-	-
Total aggregate by volume (%)	10.61	-	-	-	-	-
Superplasticizer (% fly ash)	12.21	-	6.71	0.26	-	-
SiO_2 in Na_2SiO_3 (%)	8.90	-	-	-	-	-
Water content (% or kg)	11.70	-	18.85	0.09	0.835	176
NaOH/ Na_2SiO_3	-	12.90	0.85	0.12	-	-
Fly ash / Aggregates	-	35.55	-	-	-	-
Alkali liquid / Fly ash	-	4.22	6.85	0.06	-	-
SiO_2/ H_2O in Na_2SiO_3	-	-	22.94	-	-	-
Total aggregate percentage by volume	-	-	7.71	-	-	-
Fine to total aggregates	-	-	2.61		-	-
Fly ash (Kg)	-	-	-	0.03	0.839	-
SiO_2 in fly ash (%)	-	-	-	0.71	-	51
Al_2O_3 in fly ash	-	-	-	0.21	-	-
Coarse aggregates (kg)	-	-	-	0.03	-	58
Fine aggregates (kg)	-	-	-	0.02	-	59
NaOH solution weight (kg)	-	-	-	0.08	0.822	-
Na_2SiO_3 solution weight (kg)	-	-	-	0.30	0.833	-
Curing regime						754
Mixing procedure	-	-	-	-	-	644
Loss on ignition (%)	-	-	-	-	-	607
Specific surface area of fly ash (m^2/kg)	-	-	-	-	-	155
Na_2O in fly ash (%)	-	-	-	-	-	137
K_2O in fly ash (%)	-	-	-	-	-	107
CaO in fly ash (%)	-	-	-	-	-	99
Al_2O_3 in fly ash	-	-	-	-	-	64

than the highest values. Sensitivity of SiO_2% in fly ash was reported to have two contradictory results where Peng and Unluer [5] observed it as the highest sensitive parameter while Gomaa, et al. [64] as the least sensitive. This observation in the sensitivity study could be due to the notable differences in the ranges of SiO_2 percentage values used in the databases. Peng and Unluer [5] utilized a broad range of SiO_2 percentages ranging from 31.32% to 71.5%, whereas Gomaa, et al. [64] used a more narrow range of SiO_2 percentages, spanning from 36.9% to 43.9%. Properties of fly ash such as Al_2O_3%, Na_2O%, CaO%, K_2O% and specific surface area were found to have less sensitivity while loss in ignition was ranked as third highest sensitive parameter [64]. Curing conditions, which control the dissolution kinetics of fly ash and precipitation kinetics of the reaction products, is identified as a sensitive parameter in the model outputs. Gomaa, et al. [64] observed the highest "importance" value of 754 for the curing regime (oven curing, Ambient curing, and moist curing) while the importance value of curing temperature, with the parameter range of 27–70 ^{0}C, and time, with a parameter range of 1–28 days, were 483 and 301 respectively. Similar observation on curing temperature was made by Chu, et al. [62] and Khan, et al. [59] being the most sensitives parameter of their study with a value of 26.25% and 25.3% respectively while Huynh, et al. [4] identified curing temperature as the second most sensitive with a value of 16.22%. On the other hand, Peng and Unluer [5] identify the sensitivity of curing temperature and time as minimal with values of 0.07 and 0.09 for the PFI which are 87.3% and 90.1% less than the most sensitive in the study.

Sensitivity of Na_2SiO_3 in the model output is prominent as it controls the SiO_2 availability for polymerization and hence the features of resultant sialates network. Dao, et al. [58] and Peng and Unluer [5] identified Na_2SiO_3 has having the second highest sensitivity in their studies, 45.62% and 42.25% lower compared to the highest sensitivity of the studies respectively. However, Chu, et al. [62] identify SiO_2% in Na_2SiO_3 has least sensitivity towards the output predictions, while Khan, et al. [59] identify the ratio between SiO_2% in Na_2SiO_3 and water as the second highest sensitivity for the model predictions. Sensitivity of NaOH content and concentration has been identified to have a moderate contribution towards model outputs with PFI values of 0.08 and 0.25 respectively. Gomaa, et al. [64] used 8 mixing procedures and identified as the second highest sensitive parameter with an importance value of 644 in their study. Mixing procedures are designed to represent different mixing sequences of materials and mixing durations. Further details on mixing procedures can be found by refereeing to the paper.

6. Discussion

Machine learning techniques have been adopted by several researchers as a novel approach to forecast the compressive strength of geopolymer concrete. Most of the studies focused on predicting compressive strength as the model output by using properties of fly ash, activators, curing conditions, aggregates, and extra water as the model inputs. Fly ash, either as weight or a ratio with another parameter, has been used in all the machine learning models, while SiO_2 and Al_2O_3 content of fly ash have been used in only two studies. However, total SiO_2 and Al_2O_3 in the fly ash found to be ineffective in the geopolymer synthesis as more contribution comes from the amorphous SiO_2 and Al_2O_3 [66]. Furthermore, other chemical, physical, and mineralogical properties of fly ash, which have been found to be crucial input variables during the experimental studies, have not been included in any of the machine leaning models [66]. Thus, the highly varying nature and source dependency of fly ash have been not incorporated into machine learning models, which has been suggested as resulting in poor performance for test data in some of the machine learning models [59,64].

The effect of alkaline activator has been considered in most of the models. Properties such as concentration of alkali hydroxide controls the rate of dissolution of precursor material by alkali hydrolysis and formation of concentrated gel phase which eventually rearrange and reorganize in to geopolymers. Moreover, silicate ions released from the alkali silicate solution determine the type of end product which in turn influences the properties of geopolymer. Furthermore, sensitivity analysis of the model parameter highlighted curing time, curing temperature and water content as the three principal input variables which control the reaction process of geopolymer. Higher temperatures result high reactivity leaving less unreacted particles. Furthermore, it governs the setting time of concrete and has significant influence on initial and final setting time of geopolymer concrete. Water content controls the workability of geopolymer concrete, where high quantities cause segregation and reduction in compressive strength.

From our point of view incorporation of physical, chemical, and mineralogical properties of fly ash may lead to standard predictive model which can be used estimate compressive strength of geopolymer concrete irrespective of the origin of the fly ash. Furthermore, it enables mixing of fly ash with different sources having distinct properties to optimize the resultant concrete. NaOH concentration, NaOH content and Na_2SiO_3 content seems to be important input variables for the machine leaning models as it represents the activator in the geopolymerization. However, NaOH solid content used for NaOH solution preparation, Na_2O and SiO_2 solid content in Na_2SiO_3 along with the total water content seems to be more effective as the input variables for the models. Furthermore, curing time and temperature are required input variables for machine learning models as it controls the setting time and reaction kinetics of the geopolymer concrete. All input variables sensitivity analysis were conducted using the machine learning models developed for the corresponding study. As a result, the size and quality of the data set and level of hyper parameter tuning may have significant impacts on the sensitivity of each input variables. This explains the reason for some input variables to have higher sensitivity in some models while the same input variables may have the least in other models [5,64].

Size of the database used for the machine learning model development depends on the relevancy and reliability of the dataset. Furthermore, it depends on the model complexity, learning algorithm, required error margin and input diversity. Since papers discussed in the review have used distinct models and sets of hyper parameters, making conclusions on the size of the data set on the performance of the machine leaning model is ineffective. However, small dataset with high reliability and relevancy could perform better than a large data set which is less reliable and relevance. Data splitting between training, validation and testing is essential for model training, reducing model overfitting and assessing the generalization capability respectively. The most utilized and accepted data splitting ratio in the literature was 70:15:15 percent for training, validation, and testing. However, models with K-fold cross validation method divide data into k number of folds and used k-1 no. of folds for training while remaining set is used for validation or testing. This ensure that the entire instances are utilized for training and validation/testing of the model, and every instance is once utilized for the validation/testing purpose.

Data normalization has been used in several machine learning models while details were missing about other types of preprocessing techniques such as missing data identification and removal of duplicated data which may cause an incorrect view of the overall statistics of data and removal of outliers which often tend to affect the model's overall learning, and prediction performance. Furthermore, compressive strength of geopolymer concrete highly depend on the size and the shape of the mould used for sample casting. Thus, normalization factors for mould type need to be applied to get the better performance.

Model hyperparameters are crucial factors that determine the behavior and the performance of the machine learning model. Tuning of hyperparameters and feature selection enables ANN models to achieve similar performance compared to complex models such as DNN, ResNet and ANFIS. Proper selection of hidden layers and hidden neurons in ANN, based on the complexity of the problem and availability of data, improves the model predictions, and avoid both underfitting and overfitting of the model. As a result, performance of the model will increase

while demand for the computational requirement will reduce. However, even a higher number of hidden layers and hidden neurons combined with proper dropout probability resulted in better predictions [4] at the cost of higher computational power. Randomly initialized weights and biases of a neural network were tuned to optimize network performance as defined by the network performance function using the training function. There are two different ways in which training can be implemented: incremental mode and batch mode. All the ANN models reviewed here used batch mode training functions such as Levenberg-Marquardt and Bayesian regularization. Bayesian regularization functions slightly outperformed Levenberg-Marquardt training function while other training functions resulted poor performance during the model development [13].

DNN, Resnet and ANFIS are extended versions of ANN. DNN consists of multiple hidden layers with high number of hidden neurons at each layer [4,63]. When dealing with small number of datasets DNN models are easily prone to be overfitted and can be overcome by layer dropout and layer normalization [4]. However, dropout easily ignore the important features when there are less influential input variables. DNN models developed by Nguyen, et al. [63] used a large drop out ratio (0.7) which ignores 70% of the nodes. As a result, the model without dropout outperformed the model with dop out which can be due to ignoring of important features due to large drop out ratio. However Huynh, et al. [4] developed a DNN model with 0.2 drop out ratio and observed, model with drop out, outperform the model without drop out. Resnet was introduced to overcome the vanishing/exploding gradient of neural network by using skip connections. However, Resent models also suffer from model overfitting which can be avoided by careful implementation of layer dropout and special attention should be given to not avoid important features when dealing with small number of input variables. Similar to DNN models Resnet models resulted poor performance with large dropout ratio [63] and improved performance with small dropout ratio [4] compared to the models without applying dropout. An ANFIS is a kind of ANN that is based on Takagi–Sugeno fuzzy inference system. Khan, et al. [14] and Dao, et al. [58] developed ANN and ANFIS models and results indicate the superiority of the ANFIS model compared to the ANN models. This is because ANFIS integrates both neural networks and fuzzy logic principles and it has a potential to capture the benefits of both in a single framework. Its inference system corresponds to a set of fuzzy IF–THEN rules that have learning capability to approximate nonlinear functions.

Unlike all other machine learning methods, GEP generates a set of equations to predict the properties of geopolymer concrete based on the user-defined number of genes and function sets. The number of genes represent the number of variables involved, while the function set defines how each parameter is combined in the final equation to predict the compressive strength. However, proper selection of number of genes and function sets are an iterative process, and the final equations may not necessarily capture the complex behavior of geopolymer concrete as it would be captured by a neural network-based model. Tuning of hyperparameters for any machine learning model is iterative process and depends on the size of dataset and number of input variables used to train the model. Compared to complex neural networks, simplified models such as support vector machine and RF have proven their ability to predict the properties of geopolymer concrete with higher prediction accuracy. However, the generalization capability of the model in practical applications and tuning of model hyperparameters need further investigation.

Only 5 models used a test dataset to assess the generalization capability while the rest of the 16 models used only training and validation datasets. Assessing the model performance using the validation dataset is not a reliable approach because the validation data is already exposed to the model during the hyperparameter tuning and hence the model can predict reasonable results for the validation dataset. However, based on the training and validation results of the model, DNN and Resnet outperform all other models [4,63]. As discussed earlier due to large number of hidden neurons DNN and Resent can easily overfit the model. As such in the absence of separate test dataset to assess the generalization capability, the results given are insufficient to determine whether these model are superior to the other models. Based on the performance indices on the test sets SVM and ANN are identified as the best methods to predict the compressive strength of geopolymer concrete. However, it should also be noted that the model performances may also have been affected by the choice of model inputs and the level of hyperparameter optimization.

Performance analysis of the machine learning models were conducted using several statistical parameters. Coefficient of determination was used by 50.0% of the studies while 41.6% used coefficient of correlation to assess the model performance. The Coefficient of correlation and coefficient of determination evaluate overall performance of the models and are used for the model comparison purposes. However, they do not provide an idea about the magnitude of the error. Root mean squared error (RMSE) was utilized in 83.33% of the studies while, mean absolute error (MAE) and mean squared error (MSE) were used in 66.6% and 25.0% of the studies respectively. Root mean squared error (RMSE), mean squared error (MSE), and mean absolute error (MAE) represent the magnitude of error, and the model comparisons using them are meaningful when models are built using same set of data.

From our point of view by using several independent statistical error measures enables to verify the results of model performance. However, even with the overall performance measures, a fair comparison of various machine learning models reviewed here is impossible, as these models have been developed using distinct data sets, and any data preprocessing performed (e.g. missing data treatment, outlier removal etc.) is unknown. In the current review all the studies have used either coefficient of correlation or coefficient of determination along with another error measures which give an indication of overall error magnitude, such as RMSE, MSE or MAE except for a single study [61] which have not adopted any measures evaluate the model performance.

Evaluation of model performance using laboratory trials for verification is an important step to identify the potential practical implications such as the models' behavior for new sources of fly ash outside the database, workability, and initial setting time. This will enable to assessment of the performance of the model and further optimization of the model while ensuring the practicability of the resultant concrete. Fig. 4 depicts the model predictions and experimental results of five studies while the other studies do not perform experimental tests to validate their model performance.

The coefficient of determination calculated using the experimental

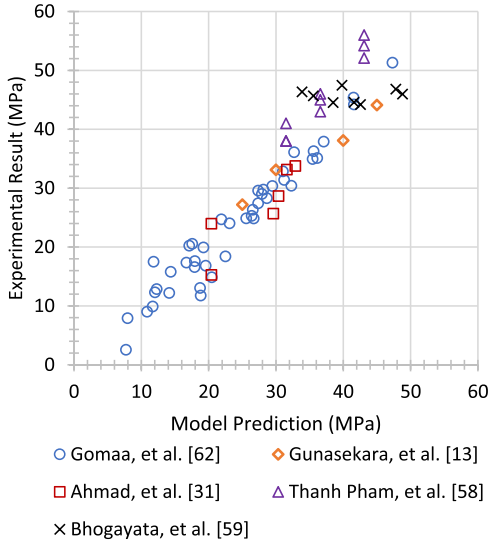


Fig. 4. Model predictions vs experimental results comparison.

results and model predictions for the study conducted by Gunasekara, et al. [13] is 96.6, while the models developed by Gomaa, et al. [64] and Thanh Pham [60] predict the compressive strength with coefficients of determination of 94.3 and 94.5, respectively. Models developed by Ahmad, et al. [33] and Bhogayata, et al. [61] showed coefficients of determination of 76.0 and 0.45, respectively. To ensure the models generalization capability it is recommended that all the developed models should be validated with an experimental data employing multiple sources of fly ash to identify model performance and improve accuracy.

As a summary the review paper discuss several machine learning methods used to determine the compressive strength of geopolymer concrete. Studies have used different databases in terms of both size and the input variables to develop the models. Furthermore, types of hyperparameters and the level of optimization were not mentioned in most of the studies. As a result, a proper comparison between the studies is hard to perform. In general, some models [5,13,63,64] have used data normalization before training the models and it has been identified that it as a crucial step in machine learning model development. Only a single attempt [12] has been made to take in to account the shape and size factor of the specimens used to evaluate the compressive strength in different studies. This strength normalization is crucial as the data had been gathered from different studies which may had used different types of molds. Studies have not mentioned about outlier detection methods, any missing data treatment or any other data pre-processing used, which could have affected the model performance.

The models under consideration in this paper have employed a set of input variables to forecast the compressive strength of geopolymer concrete. However, the approach used to derive the values for these input variables raises a concern. Because the models can generate predictions for any given input variable values, irrespective of their practicality in real-world applications. A potential remedy to this issue is to utilize the input variable values those documented in the literature. However, even with this reference, the process of identifying the most fitting values for these input variables can be a time-intensive endeavor. Another approach to identify the values for the input variables is to use graphical or statistical methods developed in the literature. However, the accuracy of the values of the input variables would depend on the accuracy of the graphical or statistical model.

7. Conclusions

This paper has reviewed the machine learning techniques used to predict compressive strength of geopolymer concrete. The following conclusions can be summarized from the overall review within this study:

1. The reported models were accurate for the limited ranges for the input variables such as chemical composition, particle size distribution, surface area and amorphous content of fly ash. However, the models may not perform well for a completely new source of fly ash with properties lying outside the established limited ranges.
2. ANN has been used in most of the studied due to its robust nature. Both sigmoid activation function and rectified linear unit activation functions appeared effective for the hidden layer activation while Bayesian regularization function and Levenberg-Marquardt training function were effective in model training.
3. Simple models like SVM and RF achieved similar performance to ANN models, while generalization capabilities of RF should be evaluated using test dataset and further understanding on hyperparameters is required to fine-tune the models.
4. It has been observed that complex nonlinear models such as DNN, ResNet and ANFIS provide accurate prediction in conjunction with dropout and layer normalization. However, further research is required to identify whether the model loose some of the high sensitivity nodes due to drop out.
5. GEP model provides a set of equations to predict the properties of geopolymer concrete while other black box models such as ANN, SVM etc. do not provide such direct expressions to predict the properties of geopolymer concrete.
6. Only five models were validated using laboratory experiments while others have used a part of the database to either as test data or validation data to assess the model. This approach will not give an insight into the model performance for practical applications when the model enconters a completely new source of fly ash with completely distinct characteristics from those available in the database.
7. Based on the Sensitivity analysis [4,5,59,62], fly ash to aggregate ratio, curing conditions, Na_2SiO_3 content have been identified as the most sensitive input variables, while fly ash content, aggregate content, Al_2O_3%, Na_2O%, CaO%, K_2O % and specific surface area have been found to be less sensitive towards the model outputs.

8. Research gaps and recommendations

This review identifies several gaps in knowledge as follows. All the studies have selected input variables only referring to past literature, and no justification was given for the selection criteria. By incorporating a reasonable feature selection method used in the machine learning model development (such as: information gain, fisher's score, correlation coefficient, and neighborhood component analysis), this can be addressed. Furthermore, details on the data preprocessing are hardly found on the studies except for a few that motioned about data normalization. Unprocessed data may contain outliers, duplicated records, and missing data, which can affect the quality of models developed and their generalization capability. Use of proper data preprocessing and reporting them along with the data used and the evaluation performance enable readers to compare their models with those of others. This will help identifying most appropriate preprocessing techniques for the data set in concern. Moreover, details on the fine-tuning of model hyperparameters hardly found in the literature, which could have been used to enhance the model performance. Some studies have used error measure that are directly related one another. For the model evaluation, several independent statistical error measures should be used, with at least one relative error measure (such as coefficient of determination, correlation coefficient) and at least one absolute error measure (such as RMSE, MAE). Further, in addition to such summary error measures, appropriate graphical representation of model performances (such as Fig. 4 in this manuscript) are comprehensible to any reader. Assessing the generalization capability of the model using a test dataset completely unexposed to model fitting stage should incorporated into the model development and validation process in all the machine learning models. Another ideal addition, wherever possible, is further validation of such model using laboratory experiments. On another side, when such experiments are conducted with multiple samples (e.g. triplicate tests), making all such measurements available can help understand the model errors in compared to the magnitudes of the inherent uncertainties in the production process itself.

Furthermore, all the models were developed to predict the compressive strength of the geopolymer concrete for a given combination of input variables. However, the problem of practical interest is to know the values of such input variables that can provide a particular compressive strength. Currently, how ANN models can be useful in such scenario is not clear. Past data records can provide some suitable prior estimates for the input variables. How ANN models can be incorporated to further refine such estimates has to be investigated. A systematic way to extract such prior estimates from past data also requires attention in the future models. Implementing such a systematic approach to directly predict the mix proportions required to achieve a desired compressive strength, in the form of a user-friendly interface that does not require understanding machine learning would certainly benefit industry.

In proposing future recommendations, the development of a

comprehensive machine learning model that integrates the chemical, physical, and mineralogical properties of fly ash as inputs is recommended. In addition, expanding the model to predict other important mechanical properties of geopolymer concrete, like flexural strength, split tensile strength, and elastic modulus is recommended. These properties are crucial for structural engineering and including them in the model would make it more comprehensive. This way, the model can be a useful tool for engineers and researchers, helping them better understand how geopolymer concrete behaves mechanically. This expansion will make the model more versatile and contribute to the wider application of geopolymer concrete in different structural situations. Such a user-friendly model has the potential to revolutionize the industry's ability to employ geopolymer concrete in practical applications, offering a seamless solution for both practitioners and researchers.

CRediT authorship contribution statement

David Law: Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Madushan Rathnayaka:** Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Chamila Gunasekara:** Writing – review & editing, Supervision, Project administration, Methodology, Formal analysis, Conceptualization. **Dulakshi Karunasinghe:** Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Weena Lokuge:** Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Kushan Wijesundara:** Writing – review & editing, Supervision, Formal analysis, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships

that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

The scholarship provided by the School of Engineering in RMIT University and Faculty of Engineering in University of Peradeniya to the first author is acknowledged. The financial support provided by Australian Research Council Discovery Early Career Researcher Grant (DE230101221) is also acknowledged.

References

[1] C. Shi, A.F. Jiménez, A. Palomo, New cements for the 21st century: the pursuit of an alternative to Portland cement, Cem. Concr. Res. 41 (7) (2011) 750–763.
[2] L. Imtiaz, S.K.U. Rehman, S. Ali Memon, M. Khizar Khan, M. Faisal Javed, A review of recent developments and advances in eco-friendly geopolymer concrete, Appl. Sci. 10 (21) (2020) 7838.
[3] I. Tekin, Properties of NaOH activated geopolymer with marble, travertine and volcanic tuff wastes, Constr. Build. Mater. 127 (2016) 607–617.
[4] A.T. Huynh, et al., A machine learning-assisted numerical predictor for compressive strength of geopolymer concrete based on experimental data and sensitivity analysis, Appl. Sci. 10 (21) (2020), https://doi.org/10.3390/app10217726.
[5] Y. Peng, C. Unluer, Analyzing the mechanical performance of fly ash-based geopolymer concrete with different machine learning techniques, Constr. Build. Mater. 316 (2022), https://doi.org/10.1016/j.conbuildmat.2021.125785.
[6] D. Hardjito, S.E. Wallah, D.M. Sumajouw, B.V. Rangan, On the development of fly ash-based geopolymer concrete, Mater. J. 101 (6) (2004) 467–472.
[7] A.A. Aliabdo, M. Abd Elmoaty, M.A. Emam, Factors affecting the mechanical properties of alkali activated ground granulated blast furnace slag concrete, Constr. Build. Mater. 197 (2019) 339–355.
[8] M.J. Castro-Alonso, L.E. Montañez-Hernandez, M.A. Sanchez-Muñoz, M.R. Macias Franco, R. Narayanasamy, N. Balagurusamy, Microbially induced calcium carbonate precipitation (MICP) and its potential in bioconcrete: microbiological and molecular concepts, Front. Mater. 6 (2019) 126.
[9] G. Habert, et al., Environmental impacts and decarbonization strategies in the cement and concrete industries, Nat. Rev. Earth Environ. 1 (11) (2020) 559–573.
[10] G.G. Mejeoumov, Improved Cement Quality and Grinding Efficiency by Means of Closed Mill Circuit Modeling, Texas A&M University, 2007.
[11] H.U. Ahmed, A.S. Mohammed, A.A. Mohammed, R.H. Faraj, Systematic multiscale models to predict the compressive strength of fly ash-based geopolymer concrete at various mixture proportions and curing regimes, PLoS One 16 (6) (2021) e0253006, https://doi.org/10.1371/journal.pone.0253006.
[12] M. Ali Khan, A. Zafar, A. Akbar, M.F. Javed, A. Mosavi, Application of gene expression programming (GEP) for the prediction of compressive strength of geopolymer concrete, Materials 14 (5) (2021), https://doi.org/10.3390/ma14051106.
[13] C. Gunasekara, P. Atzarakis, W. Lokuge, D.W. Law, S. Setunge, Novel analytical method for mix design and performance prediction of high calcium fly ash geopolymer concrete, Polymers 13 (6) (2021), https://doi.org/10.3390/polym13060900.
[14] M.A. Khan, et al., Geopolymer concrete compressive strength via artificial neural network, adaptive neuro fuzzy interface system, and gene expression programming with K-fold cross validation, Front. Mater. 8 (2021), https://doi.org/10.3389/fmats.2021.621163.
[15] X. Lu, et al., Recycling of coal fly ash in building materials: a review, Minerals 13 (1) (2023) 25.
[16] X.C. Yu, Y. Chen, Y. Chang, I.-S. Wu, J, The drivers of collaborative innovation of the comprehensive utilization technologies of coal fly ash in China: a network analysis, Environ. Sci. Pollut. Res. Int. (29) (2022) 56291–56308.
[17] N.W.A.Z. Khairuddin, A.K. Phillip, E. Sujan, M.F, Coal power plant fly ash characterization assessment for geopolymerization process, Key Eng. Mater. 908 (2022) 678–684.
[18] Y.J. Shi, K.-X. Jiang, T.A. Zhang, Cleaner extraction of alumina from coal fly ash: Baking-electrolysis method, Fuel 273 (2020) 117697.
[19] S. Nazar, et al., Machine learning interpretable-prediction models to evaluate the slump and strength of fly ash-based geopolymer, J. Mater. Res. Technol. 24 (2023) 100–124.
[20] P.K. Dash, S.K. Parhi, S.K. Patro, R. Panigrahi, Efficient machine learning algorithm with enhanced cat swarm optimization for prediction of compressive strength of GGBS-based geopolymer concrete at elevated temperature, Constr. Build. Mater. 400 (2023) 132814.
[21] M.S. Reddy, P. Dinakar, B.H. Rao, A review of the influence of source material's oxide composition on the compressive strength of geopolymer concrete, Microporous Mesoporous Mater. 234 (2016) 12–23.
[22] Y. Zhang, R. Xiao, X. Jiang, W. Li, X. Zhu, B. Huang, Effect of particle size and curing temperature on mechanical and microstructural properties of waste glass-slag-based and waste glass-fly ash-based geopolymers, J. Clean. Prod. 273 (2020) 122970.
[23] J. Van Jaarsveld, J. Van Deventer, G. Lukey, The characterisation of source materials in fly ash-based geopolymers, Mater. Lett. 57 (7) (2003) 1272–1280.
[24] L.N. Assi, E.E. Deaver, P. Ziehl, Effect of source and particle size distribution on the mechanical and microstructural properties of fly Ash-Based geopolymer concrete, Constr. Build. Mater. 167 (2018) 372–380.
[25] J.L. Provis, P. Duxson, J.S. van Deventer, The role of particle technology in developing sustainable construction materials, Adv. Powder Technol. 21 (1) (2010) 2–7.
[26] B. Singh, G. Ishwarya, M. Gupta, S.K. Bhattacharyya, Geopolymer concrete: a review of some recent developments, Constr. Build. Mater. 85 (2015) 78–90, https://doi.org/10.1016/j.conbuildmat.2015.03.036.
[27] N. Li, C. Shi, Z. Zhang, H. Wang, Y. Liu, A review on mixture design methods for geopolymer concrete, Compos. Part B Eng. 178 (2019), https://doi.org/10.1016/j.compositesb.2019.107490.
[28] H. Nguyen, T. Vu, T.P. Vo, H.-T. Thai, Efficient machine learning models for prediction of concrete strengths, Constr. Build. Mater. 266 (2021), https://doi.org/10.1016/j.conbuildmat.2020.120950.
[29] M. Azimi-Pour, H. Eskandari-Naddaf, A. Pakzad, Linear and non-linear SVM prediction for fresh properties and compressive strength of high volume fly ash self-compacting concrete, Constr. Build. Mater. 230 (2020), https://doi.org/10.1016/j.conbuildmat.2019.117021.
[30] L.V. Zhang, A. Marani, M.L. Nehdi, Chemistry-informed machine learning prediction of compressive strength for alkali-activated materials, Constr. Build. Mater. 316 (2022), https://doi.org/10.1016/j.conbuildmat.2021.126103.
[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] S. Aneja, A. Sharma, R. Gupta, D.-Y. Yoo, Bayesian regularized artificial neural network model to predict strength characteristics of fly-ash and bottom-ash based geopolymer concrete, Materials 14 (7) (2021) 1729.
[33] M. Ahmad, K. Rashid, Z. Tariq, M. Ju, Utilization of a novel artificial intelligence technique (ANFIS) to predict the compressive strength of fly ash-based geopolymer, Constr. Build. Mater. 301 (2021), https://doi.org/10.1016/j.conbuildmat.2021.124251.
[34] RSN 336-84, Republican Building Norms for Production and Use of Slag Alkaline Binders, 1984.
[35] S. C, A mixture proportioning method for alkali-activated slag concrete, Build Energy Effic., 1989.
[36] P. 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.

[37] Y.K. Yang C, Q. Pan, K. Chen, X. Zhu, Parameter and design method for alkaliactivated slag concrete mix, J. Civ., Arch. Environ. Eng. 39 (122) (2017).
[38] M.S. Reddy, P. Dinakar, B.H. Rao, Mix design development of fly ash and ground granulated blast furnace slag based geopolymer concrete, J. Build. Eng. 20 (2018) 712–722.
[39] 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).
[40] R. Anuradha, V. Sreevidya, R. Venkatasubramani, B.V. Rangan, Modified guidelines for geopolymer concrete mix design using Indian standard, 2012.
[41] M.W. Ferdous, O. Kayali, and A. Khennane, A detailed procedure of mix design for fly ash based geopolymer concrete, Presented at the Fourth Asia-Pacific Conference on FRP in Structures, 2013.
[42] W. Ferdous, A. Manalo, A. Khennane, O. Kayali, Geopolymer concrete-filled pultruded composite beams–concrete mix design and application, Cem. Concr. Compos. 58 (2015) 1–13.
[43] S.V. Patankar, Y.M. Ghugal, S.S. Jamkar, Mix design of fly ash based geopolymer concrete. Advances in Structural Engineering, Springer, 2015, pp. 1619–1634.
[44] M.T. Junaid, O. Kayali, A. Khennane, J. Black, A mix design procedure for low calcium alkali activated fly ash-based concretes, Constr. Build. Mater. 79 (2015) 301–310.
[45] B.V. Rangan, Fly Ash-based Geopolymer Concrete, 2008.
[46] T.S. Ng, S.J. Foster, Development of a mix design methodology for high-performance geopolymer mortars, Struct. Concr. 14 (2) (2013) 148–156.
[47] A. Rafeet, R. Vinai, M. Soutsos, W. Sha, Guidelines for mix proportioning of fly ash/GGBS based alkali activated concretes, Constr. Build. Mater. 147 (2017) 130–142.
[48] D. Bondar, Q. Ma, M. Soutsos, M. Basheer, J.L. Provis, S. Nanukuttan, Alkali activated slag concretes designed for a desired slump, strength and chloride diffusivity, Constr. Build. Mater. 190 (2018) 191–199.
[49] D. Bondar, S. Nanukuttan, J.L. Provis, M. Soutsos, Efficient mix design of alkali activated slag concretes based on packing fraction of ingredients and paste thickness, J. Clean. Prod. 218 (2019) 438–449.
[50] M. Olivia, H. Nikraz, Properties of fly ash geopolymer concrete designed by Taguchi method, Mater. Des. (1980-2015) 36 (2012) 191–198.
[51] S. Riahi, A. Nazari, D. Zaarei, G. Khalaj, H. Bohlooli, M.M. Kaykha, Compressive strength of ash-based geopolymers at early ages designed by Taguchi method, Mater. Des. 37 (2012) 443–449.
[52] G. Khalaj, S.E.S. Hassani, A. Khezrloo, Split tensile strength of OPC-based geopolymers: application of DOE method in evaluating the effect of production parameters and their optimum condition, Ceram. Int. 40 (7) (2014) 10945–10952.
[53] A. Mehta, R. Siddique, B.P. Singh, S. Aggoun, G. Łagód, D. Barnat-Hunek, Influence of various parameters on strength and absorption properties of fly ash based geopolymer concrete designed by Taguchi method, Constr. Build. Mater. 150 (2017) 817–824.
[54] L.Z. Yang Y, J. Cai, J. Zhang, X. Chen, Mixture proportioning of inorganic polymer concrete, Guangdong Water Resour. Hydr. 27 (34) (2017).
[55] M.N. Hadi, N.A. Farhan, M.N. Sheikh, Design of geopolymer concrete with GGBFS at ambient curing condition using Taguchi method, Constr. Build. Mater. 140 (2017) 424–431.
[56] N. Li, et al., A mixture proportioning method for the development of performance-based alkali-activated slag-based concrete, Cem. Concr. Compos. 93 (2018) 163–174.
[57] Y.S.N. Kishore, S.G.D. Nadimpalli, A.K. Potnuru, J. Vemuri, M.A. Khan, Statistical analysis of sustainable geopolymer concrete, Mater. Today Proc. (2021), https://doi.org/10.1016/j.matpr.2021.08.129.
[58] D.V. Dao, H.B. Ly, S.H. Trinh, T.T. Le, B.T. Pham, Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete, Materials 12 (6) (2019), https://doi.org/10.3390/ma12060983.
[59] M.A. Khan, et al., Compressive strength of fly-ash-based geopolymer concrete by gene expression programming and random forest, Adv. Civ. Eng. 2021 (2021) 1–17, https://doi.org/10.1155/2021/6618407.
[60] T. Thanh Pham, A neural network approach for predicting hardened property of geopolymer concrete, Int. J. Geomate 19 (74) (2020) 176–184, https://doi.org/10.21660/2020.74.72565.
[61] A. Bhogayata, S. Kakadiya, R. Makwana, Neural network for mixture design optimization of geopolymer concrete, Acids Mater. J. 118 (4) (2021), https://doi.org/10.14359/51732711.
[62] H.-H. Chu, et al., Sustainable use of fly-ash: use of gene-expression programming (GEP) and multi-expression programming (MEP) for forecasting the compressive strength geopolymer concrete, Ain Shams Eng. J. 12 (4) (2021) 3603–3617.
[63] K.T. Nguyen, Q.D. Nguyen, T.A. Le, J. Shin, K. Lee, Analyzing the compressive strength of green fly ash based geopolymer concrete using experiment and machine learning approaches, Constr. Build. Mater. 247 (2020) 118581.
[64] E. Gomaa, T. Han, M. ElGawady, J. Huang, A. Kumar, Machine learning to predict properties of fresh and hardened alkali-activated concrete, Cem. Concr. Compos. 115 (2021), https://doi.org/10.1016/j.cemconcomp.2020.103863.
[65] G. Dong, H. Liu, Feature Engineering for Machine Learning and Data Analytics, CRC Press, 2018.
[66] C. Gunasekara, Influence of Properties of Fly Ash from Different Sources on the Mix Design and Performance of Geopolymer Concrete, RMIT University, 2022.
[67] H.W. Nugteren, Coal fly ash: from waste to industrial product, in: Particle & Particle Systems Characterization, 24, 2007, pp. 49–55.
[68] N. Li, C. Shi, Q. Wang, Z. Zhang, Z. Ou, Composition design and performance of alkali-activated cements, Mater. Struct. 50 (2017) 1–11.
[69] S. Haykin, Neural Networks and Learning Machines, 3/E, Pearson Education India,, 2009.
[70] S. Sharma, S. Sharma, A. Athaiya, Activation functions in neural networks, Towards Data Sci. 6 (12) (2017) 310–316.
[71] S. Agatonovic-Kustrin, R. Beresford, Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research, J. Pharm. Biomed. Anal. 22 (5) (2000) 717–727.
[72] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 770-778.
[73] D. Karaboga, E. Kaya, Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey, Artif. Intell. Rev. 52 (2019) 2263–2293.
[74] C. Ferreira, Gene expression programming: a new adaptive algorithm for solving problems, arXiv preprint cs/0102027, 2001.
[75] C. Ferreira, Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence, Springer, 2006.