Md Nasir Uddin, Kequan Yu, Ling, zhi Li, Junhong Ye, T. Tafsirojjaman, Wael Alhaddad

Developing machine learning model to estimate the shear capacity for RC beams with stirrups using standard building codes

Innovative Infrastructure Solutions, 2022; 7(3):227-1-227-20

Final publication at http://dx.doi.org/10.1007/s41062-022-00826-8

Developing Machine Learning Model to Estimate the Shear Capacity for RC Beams with Stirrups Using Standard Building Codes

Md Nasir Uddin [a*] Kequan Yu [a], Ling-zhi Li [a], Junhong Yu [a], T. Tafsirojjaman [b] and Wael Alhaddad [c]

[a] *Department of Disaster Mitigation for Structures, College of Civil Engineering, Tongji University, Siping Road 1239, Shanghai,200092, China.*

[b] *School of Civil Engineering and Surveying, University of Southern Queensland, Toowoomba, Australia.*

[c] Department of Structural Engineering, Tongji University, Shanghai 200092, China

Abstract

Shear failure in reinforced concrete (RC) beams with a brittle nature is a serious safety concern. Due to the inadequate description of the phenomenology of shear resistance (the shear behaviour of RC beams), several of the existing shear design equations for RC beams with stirrups have high uncertainty. Therefore, the predicted models with higher accuracy and lower variability are critical for the shear design of RC beams with stirrups. To predict the ultimate shear strength of RC beams with stirrups, machine learning (ML) based models are proposed in the present research. The models were created using a database of 201 experimental RC beams with stirrups gathered from earlier investigations for training and testing of the ML method, with 70% of the data being used for model training and the rest for testing. The performance of suggested models was evaluated using statistical comparisons between experimental results and state-of-the-art current shear design models (ACI 318-08, Canadian code, GB 510010-2010, NZS 3101, BNBC 2015). The suggested machine-learning-based models are consistent with experimentally observed shear strength and current

predictive models, but they are more accurate and impartial. To understand the model very well, sensitivity analysis is determining as input values for a specific variable affect the outcomes of a mathematical model. To compare the results with different machine learning models in training and testing R^2, RMSE and MSE are also established. Finally, proposed ML models such as gradient boost regressor and random forest give higher accuracy to evaluate the shear strength of the reinforcement concrete beam using stirrups

Keywords

RC Beam with Stirrups, Shear capacity, ML Algorithm, Standard building code, Sensitivity Analysis

1 Introduction

The shear behaviour of reinforced concrete (RC) structures has been intensively studied for decades. The intricacy of the shear transfer mechanism [1], [2], [3], [4], [5] the shear strength models for RC beams depends on various paraments and factors while considering different country codes. As a result, the precise value of shear strength remains a mystery. Consequently, empirical formulas and semiempirical formulas for design codes are proposed by regression analysis of experimental data against certain theoretical hypotheses. In structural design, the American concrete institute (ACI 318-08), Canadian building design code (CSA A23.3 04), Chinese design code (GB50010-2010), New Zealand building code (NZS3101), and Bangladesh building design code (BNBC 2015) have been used actual simulated in practice. In recent decades, numerous investigations have been done to enhance the ability of empirical formulas to predict the shear behaviour of concrete structures [6], [7]. Bazant and Kim [8] and Russo et al. [4] proposed the shear strength equations for normal-strength concrete and high strength concrete beam, respectively, and their size effect. Arslan et al. [5] equation are related to slender reinforced concrete (RC) beam without stirrups, where normal strength concrete and high strength concrete is considered for a shear design equation. Zsutty et al. [9] developed a database where statistical regression analysis was applied and established an equation for failure stress prediction for the sudden shear failure of the slender beam. However, the prediction accuracy of these equations and developed models is low; when multiple influence factors are incorporated, a model with high robustness and accuracy is required to predict the beam's shear capacity.

Back-propagating neural networks (RBPNNs) [10], artificial neural networks (ANNs) [11],[12],[13],[14][15],[2],[16],[17],[18],[19],[20], adaptive neuro-fuzzy inference systems (ANFIS) [21], [22],[23],[24], and have been utilized to forecast concrete qualities, damage detection, shear, and flexural strength and compressive strength prediction in recent decades. Other machine learning techniques such as random forest (RF) [25],[26], support vector machine (SVM) [27],[28],[29] are widely using for optimization and to evaluate the compressive strength of the concrete specimens. Concrete compressive strength and electrical resistivity are predicted using by gene expression programming (GEP) [30],[31],[32],[33]. The variable input factors were used in the GEP model as Portland cement, water, fine aggregates, coarse aggregate (CA), superplasticizer, fly ash (FA), blast furnace slag (BFS), and concrete age. The suggested literature, GEP, ANN, and SVM models used testing data for validation and training. The prediction results revealed that among the eight input parameters, water, cement, and age have the greatest impact on the compressive strength of specimens.

One of the major exploration areas in the use of ML techniques has been predicting the shear and flexural capacity of the beam, examples of which include studies by [34],[35],[34],[36]. The shear capacity of fibre-based-reinforced concrete slabs was investigated using regression model, ensemble tree (bagged and boosted), SVM, DT, GPR, ANNs with 148 experimental test results and eight input parameters in these

studies [36]. To predict the compressive and flexural strengths of steel fibre reinforced concrete, XGBoost and GBR showed the best results using seven input parameters [37]. Later, they valid the result using RMSE MSE and see their correlation among the variables. Solhmirzaei et al. [38] used the same technique to predict the failure mechanisms and shear capacity of ultra-high-performance fibre concrete, comparing the results to previous research and existing codes and predicting the best model for the problem. Similarly, a genetic algorithm (GA) was being extended to the shear capacity of the deep beam [39]. To compare the mathematical models and the different country code equations, a database of 381 deep beam experiments was used. For structural safety, the reliability analysis and the resistance factors for shear design were determined. The bond strength estimation of concrete-encased steel structures using ANN-PSO, GAPSO [40] showed more accurate results than ANN. To estimate the ultimate compressive strength of rectangular concrete-filled steel tubular (RCFST) columns [41], the hybrid model PANN was designed to compare the result with the design code. In the end, sensitivity analysis was also conducted to see the influence on input paraments in the design.

Cladera et al. [3] used the ANN method to develop a new method for predicting shear strength; a total of 123 test data was used, and compared to other methods, 0.23 standard deviation was gained. It is necessary to create a better database and use a novel ML algorithm design approach that may optimize the higher range of input qualities in the RC beam design to enhance the usability of RC beams with stirrups. The outcome of numerous elements of the RC members with stirrups was investigated using a computational framework based on machine learning. The results of multiple different testing procedures were gathered and fed into

the machine learning architecture. Hoang et al. [42] used pattern recognition algorithms to identify the critical parameters that affect the shear response of RC beams. In Fig. 1, stirrups prevent cracks from forming in beams and increase aggregate interlocking capacity, making the mechanical properties of stirrup beams more complicated. The suggested ML models' expressions are a function of beam geometric properties and loading configuration can be used to forecast the shear capacity of an RC beam with stirrups.

This study aimed to create ML models that could evaluate the shear capacity of the RC beam with stirrups in such a way that model predictions could be quantitatively connected to input data. First, a substantial experimental database is built, followed by comparing the current result with the projected model in the following parts. The data is then divided into training and testing sets at random using the standard proportions of 70%–30%. A robust machine learning approach called ANNs, GEP, RF, GBR, LR, and RR is utilized to develop a predictive shear strength model. The testing sets are then used to evaluate the ML model, and its predictions are compared to those of various existing classical mechanics-based models and existing codes. Finally, using a machine learning model, the elements that contribute to a certain shear strength estimation are investigated, and various conclusions are derived.

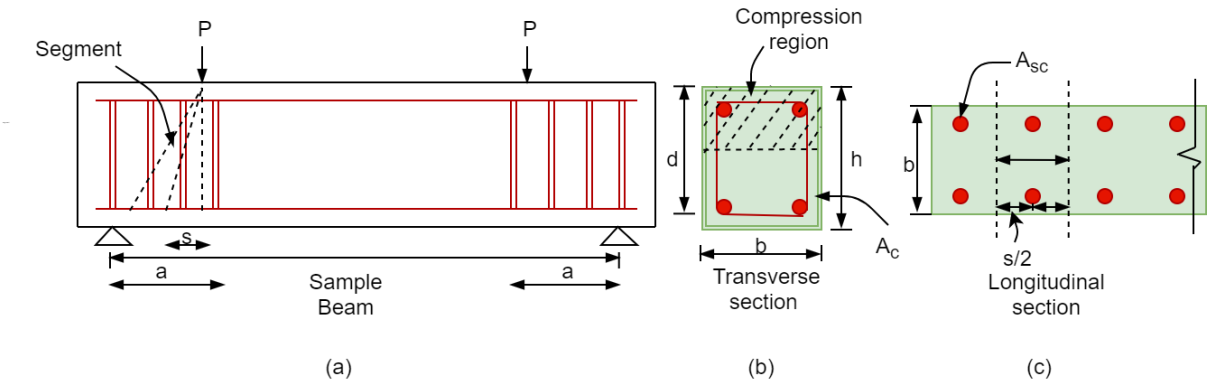


Fig. 1. Obtained various locations of the RC beams using stirrups in shear.

Table 1

Statistical parameters of the different variables used in the database

Parameters	Unit	Mean	Standard	Minimum	Average	Max	Operation
a	mm	1042.4	560.3	190.0	914.0	2553.0	Input
b	mm	208.2	83.0	76.0	203.0	457.0	Input
d	mm	357.4	127.7	95.0	383.0	851.0	Input
f_{co}	MPa	29.6	7.2	12.8	27.7	50.3	Input
ρ_{rt}	%	2.7	0.9	1.0	2.7	4.8	Input
f_{ry}	MPa	459.2	130.1	300.0	434.0	707.0	Input
ρ_{stp}	%	0.3	0.2	0.1	0.3	1.9	Input

f_{sy}	MPa	384.3	122.8	159.0	331.0	820.0	Input
V_{test}	KN	253.8	164.3	13.6	211.9	836.1	Output

2 Shear design of RC beam with stirrups

The following section summarizes the existing shear design models for RC beams with stirrups as defined in design guidelines.

2.1 ACI 318-08

The shear capacity of non-prestressed concrete members without shear reinforcement is calculated [43] using Eq. (1) and in SI units as:

$$v_c = \left(\frac{\lambda \sqrt{f_{co}}}{6} \right) bd \tag{1}$$

where λ is a factor that contributes to the lightweight of concrete.

The shear resistance can be calculated for a more detailed analysis by using Eq. (2) as follows

$$v_c = \left(0.16\sqrt{f_{co}} + 17_{\rho rt} \frac{V_u}{M_u} \right) bd \le 0.29\sqrt{f_{co}} bd \text{ V} \tag{2}$$

V_u and M_u are the calculated shear force and bending moment in the crucial section under consideration. In any instance, the ratio $\frac{V_u}{M_u}$ should not exceed 1.

According to the ACI, code V_n is considered as a normal shear strength

$$V_n = v_c + v_s \tag{3}$$

where from Eq. (4) v_c is taken and v_s expressed as a

$$v_s = \frac{A_v f_{sy} d}{s} = \rho_{rt} b \rho_{stp} d \tag{4}$$

where f_{sy} is the yielding strength of the transverse reinforcement, A_v is the area of vertical shear reinforcement. ρ_{rt} is the yielding of the longitudinal reinforcement ratio, s considered as a stirrup spacing, and ρ_{stp} is the transverse reinforcement ratio.

2.2 Building code CSA (Canadian Standards Association)

According to the Canadian code [44], the shear capacity of beams without web reinforcing is determined only by concrete compressive strength. Shear resistance in a beam without shear reinforcement is calculated Eq. (5) as follows:

$$v_c = 0.2\sqrt{f_{co}} bd \tag{5}$$

where b and d considered as a web width and effective depth, respectively. Concrete compressive strength f_{co} in MPa.

ACI building code v_s has the same value considering Eq. (4)

2.3 Chinese Code GB 50010-2010

To improve the safety in shear design, the coefficient of stirrups 1.50 in the formula is uniformly

changed to 1.0. For the shear capacity of with stirrups only [45].

$$V_n = 0.7 f_{to} bd + \rho_{rt} \frac{A_{sv}}{s} d \tag{6}$$

where f_{to} is the tensile strength of the concrete. ρ_{rt} is the yielding of longitudinal reinforcement ratio. The cross-sectional area of the bent steel bar in the bent plane A_{sv}, d is the effective depth.

$$V_n = 0.07 f_{co} bd + \rho_{rt} \frac{A_{sv}}{s} d \tag{7}$$

For converting the concrete tensile strength to compressive strength $f_{co} = 10 f_{to}$, have been used.

2.4 New Zealand concrete structure code

The nominal shear capacity of a reinforced concrete beam with transverse reinforcement [46] is given by

$$V_n = \left(0.07 + 10\rho_{rt}\right)\sqrt{f_{co}} bd + \rho_{stp}\rho_{rt} bd \tag{8}$$

2.5 BNBC 2015

Eq. (9) calculates the concrete shear strength (MPa) when subjected to the shear strength of unreinforced concrete parts [47] as follows:

$$v_c = 0.17\lambda\sqrt{f_{co}} bd \tag{9}$$

For more details, shear strength can be calculated by

$$v_c = \left(0.16\lambda\sqrt{f_{co}} + 17\rho_{rt}\frac{V_u d}{M_u} \right) bd \le 0.29\sqrt{f_{co}}bd \tag{10}$$

where, $\frac{V_u d}{M_u}$ shall not take greater than 1.0, where M_u occurs simultaneously with V_u at the section considered.

$$v_s = \frac{A_v f_{sy}}{bs} = \rho_{rt}\frac{A_{sv}}{s}d \tag{12}$$

where, the BNBC code follows the same as the ACI 318-08 code.

3 Database Creation

Total 201 sets experimental test results of shear strength test of reinforcement concrete beam with steel stirrups was obtained from [48], [49], [50], [51] [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65]. Each paper's experimental test results included a diverse set of parameters. To ensure the input variables, the majority of influence parameters were used to predict the shear capacity, which can also be used in the county code. Each literature contains eight input parameters. The parameters are shear span *a,* web width *b,* effective depth *d,* transverse reinforcement ratio ρ_{stp} , yielding strength of the transverse reinforcement f_{sy} , concrete compressive strength f_{co} , yielding of longitudinal reinforcement ratio ρ_{rt} , yielding strength of longitudinal reinforcement f_{ry} . Table 1 summarizes the model's main variable locations and datasets parameters explained as mean, standard deviation, minimum and maximum. The only output in this study's database was the experiment shear capacity V_{test} of the RC stirrups.

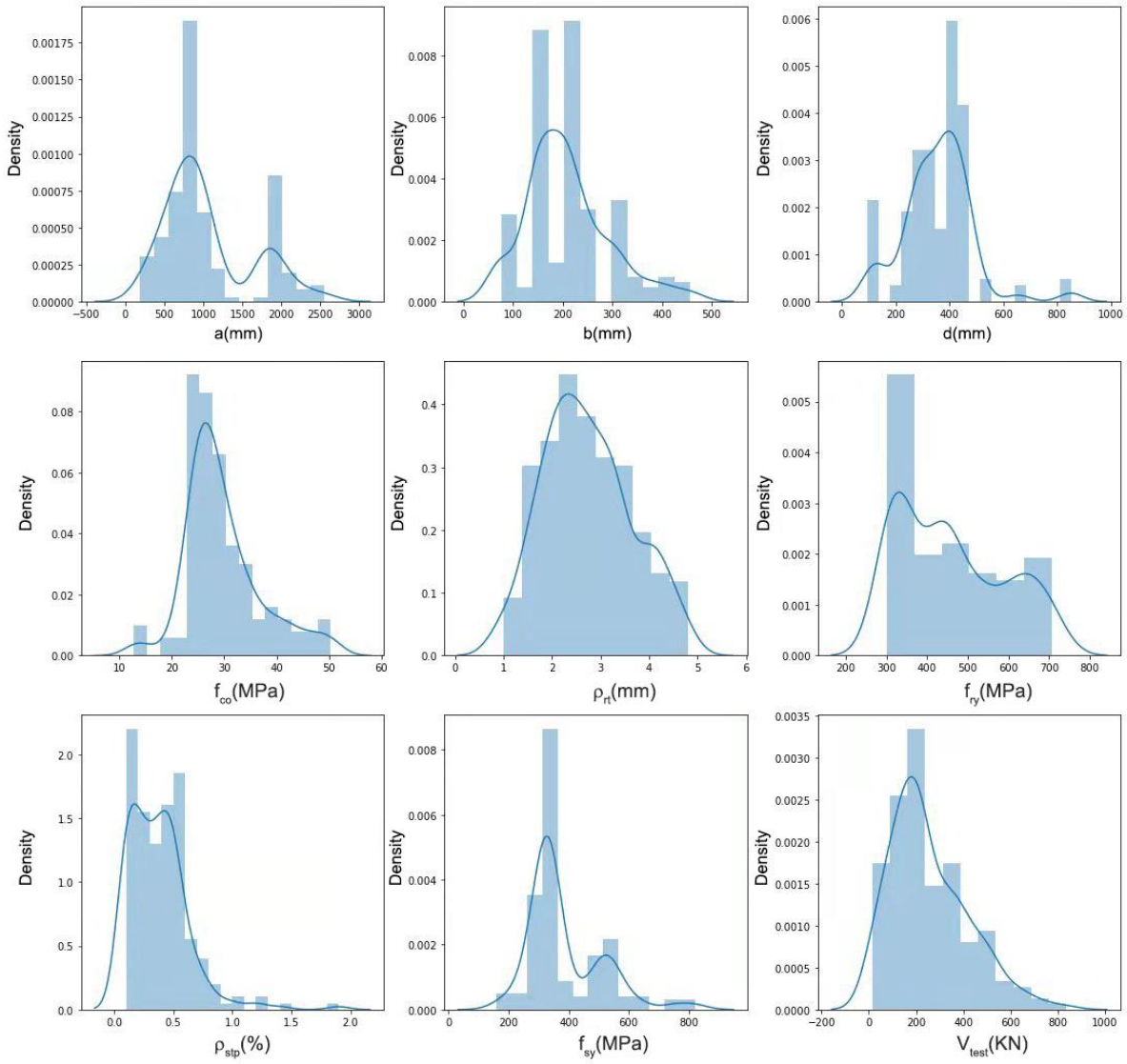


Fig. 2. Kernel Density estimation for input and output variables.

In **Fig. 2**, Kernel density estimation (KDE) estimated the probability density function (PDF) of the random variables in the dataset. It can easily show the distribution of the data in each variable. The high picks show more data on distribution at this stage. The KDE line shows the distribution line in the dataset. The bandwidth of the KDE is a parameter that has a high impact on the estimation. The higher bandwidth reveals smooth picks, and the low bandwidth shows more picks in the figures.

Table 1

Different machine learning models and their performance in predicting shear capacity for various problems.

Author	Year	Machine learning method	R^2
Olalusi et al .[66]	2020	GPR	0.93
		RF	0.95
		ANFIS	0.65
		SVM	0.64
Pham et al. [67]	2018	PSO	0.81
		ANNs	0.63
Olalusi et al .[68]	2020	GPR	0.99
		SVM	0.98
Gao et al. [69]	2021	XGBoost	0.91
Solhmirzaeia et al. [70]		SVM	0.81
	2020	ANNs	88.9
		k-NN	0.75

4 The Application of Machine learning

Table 2 describes the previous machine learning method that has been used for different structural

problems such as beam-column join, ultra-high performance concrete RC beam, soft soil, Steel fibre RC beam. The shear capacity has been predicted through GEP, SVM, ANNs, XGBoost, KNNs, PSO, ANFIS, and RF. Almost all the ML model shows better performance in predicting the shear capacity. Later, describe six machine learning models to evaluate our result and choose the best model to determine the RC beam's shear capacity with stirrups. The standard country code will evaluate the results.

4.1 Linear Regression

Statistical analysis using linear regression is a well-known technique for discovering relationships between two or more variables. [71]. After establishing a connection between the features (input), and the target (output), the learning process will be used to minimize the value of the loss function (like Mean Squared Error) [72]. The weights that minimize the loss function are also the ideal parameters for regression. This model has a low degree of accuracy due to its simplicity. The general structure of a multiple linear regression model is illustrated in **Eq. (13)** [73] follow:

$$\hat{y} = a_0 + \sum_{j=1}^{n} a_j X_j \tag{13}$$

where X_j are the features (input) of the dataset and $a_0, a_1, a_2, ..., a_n$ the parameters to the train set, the observed result showed $\hat{y}$.

4.2 Ridge Regression

When the independent variables are highly correlated, least-square estimates have a low bias but significant variance. When creating ridge regression models on a real dataset, the trade-off between bias and variance is often complex. The development of ridge regression is driven by the search for bias with a small mean square error (MSE).

Hoerl and Kennard (1970) hypothesized that the LS estimator might be unstable; the general structure of the model is illustrated in **Eq. (14)**.

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{14}$$

Before computing the inverse of the matrix $X'X$, it could be enhanced by adding a small constant λ value to the diagonal elements.

The ridge regression estimator is the outcome of this process.

$$\hat{\beta}_{ridge} = (X'X + \lambda K_p)^{-1}X'Y \tag{15}$$

The penalized sum of squares is minimized by using $\hat{\beta}_{ridge}$,

$$\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda\sum_{j=1}^{p}\beta_j^2 \tag{16}$$

where $\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2$, is equivalent to minimization of some $c > 0$, < 0 i.e., constraining the sum of the squared coefficients. The shrinkage parameters define as λ.

4.3 Developing GEP model

The optimization problem plays the role of individuals chosen for their fitness, and genetic variation is added by evolutionary algorithms utilizing one or more genetic operators. Near the 1950s, an artificial computational system was used based on an evolutionary algorithm [74].

GEP model is closely related to genetic algorithms and genetic programming (GP). It is classified as a complex tree structure known as an evolutionary algorithm used in computer models [75]. Fig. 3 demonstrates a modification to the GEP evaluation algorithm. It inherits linear chromosomes such as DNA strands of fixed length from GA and expressive parse trees of various sizes and shapes from the GP algorithm [76],[77]. The measure to evaluate in GEP is passing the genome on to the next population. Another one-of-a-kind function is developing things via chromosomes composed of genes that are subsequently expanded into the tail and head [78].In the GEP model, chromosome and pares trees work as genotype and phenotype. Those multiple pares tresses are called in GEP as expression trees. In Fig. 4, each gene is responsible for encoding a sub-expression tree (sub-ET). The sub-ETs can then engage in various ways, resulting in the cellular system's effective program. Each program in population, chromosome, execute a program and evaluate fitness are involved. The mutation cycle is repeated by adding new individuals for multiple generations until the optimal model is obtained. To revitalize the population, genetic operations, including reproduction, mutations, and crossings, are performed.

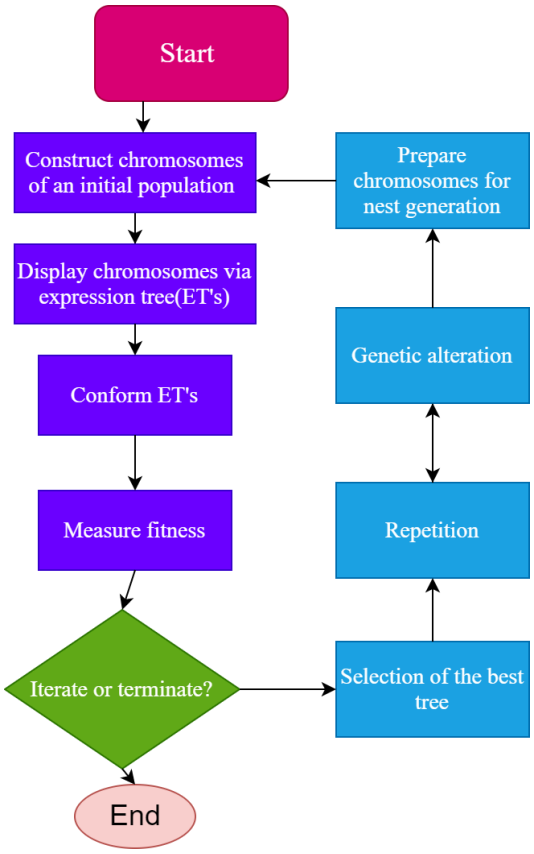


Fig. 3. The GEP algorithm's architectural flowchart.

The outcomes of the GEP model are influenced by the parameters chosen. Table 4 lists the function setups in the GeneXproTools v5.0 version that were used in the GEP model. The model's chromosome architecture, such as the number of genes, head size, and lining function, was determined based on literature [79],[80]. The final GEP models structure such as chromosomes, generals, training and testing records are shown in Table 3. These parameters are utilized to forecast the predicted value for the RC beam with stirrups. The number of trees and max complexity (Gene) was chosen from the previous experiences [81]. The model's performance is measured using the root mean square error (RMSE), and the GeneXproTools v5.0 version was used to create the model. The input parameters used for the final GEP structures are summarized in Table 1. As a result, V_{test} it could be expressed as

$$V_{test} = f\left(a, b, d, f_{co}, \rho_{rt}, f_{ry}, \rho_{stp}, f_{sy}\right) \tag{17}$$

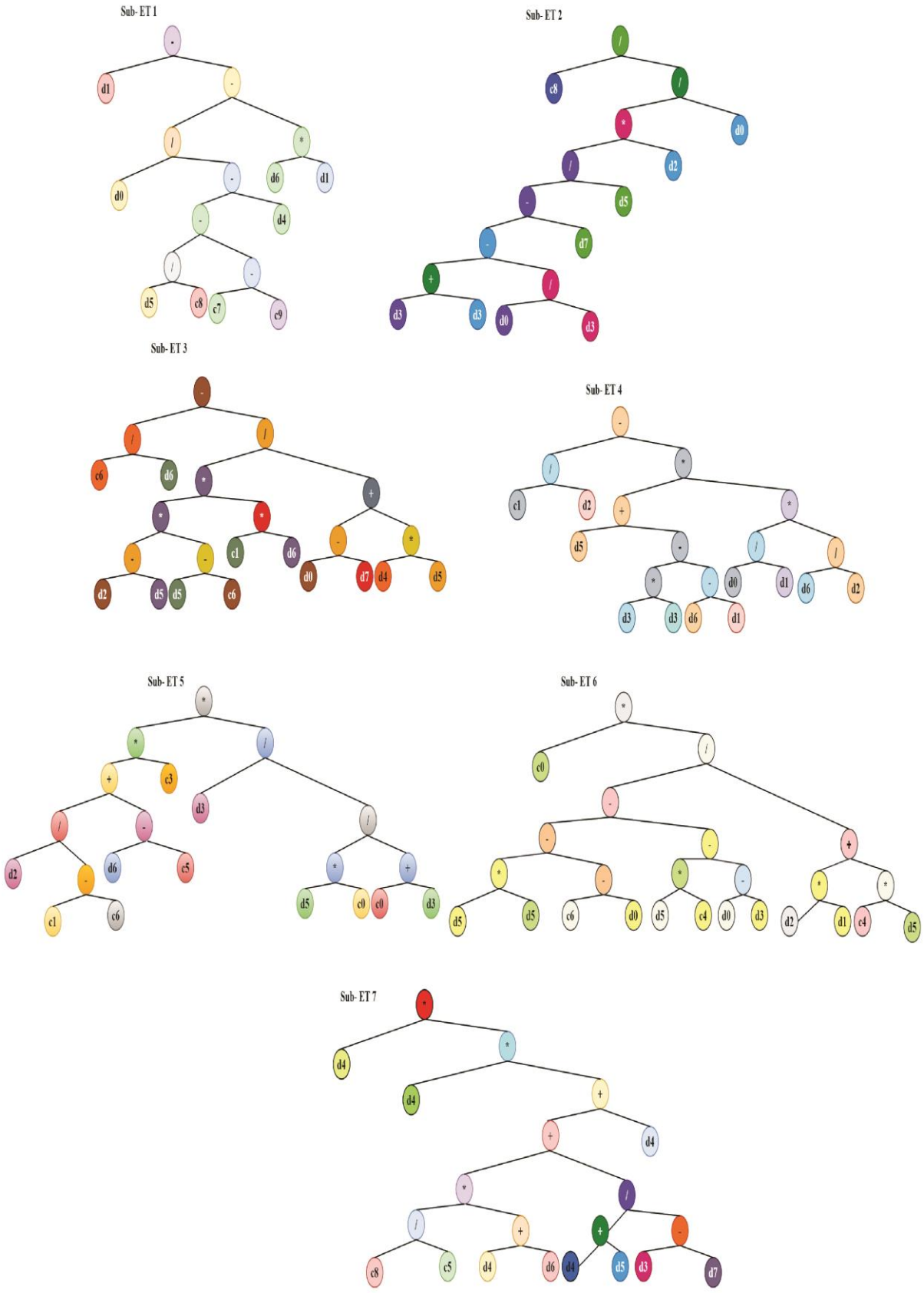


Fig. 4. Expression tree developed by gene expression tree (GEP) model.

Table 2

Parameter setting for the ultimate GEP structures

Parameter	Meaning	Setting
Chromosomes		50
	Number of Genes	7
General	Head size	13
	Tail size	13
	Dc size	14
	Gene size	14
	Linking function	Addition
	Error	Error Function
	Fitness function	RMSE
	Function set	$+, -, \times, \div$
Training records		200
Test records		52

Table 3

The statistic used in the test model for the result

Function Set			
Function	**Symbol**	**Weight**	**Arity**
Addition	+	2	2
Subtraction	-	2	2
Multiplication	×	2	2
Division	/	2	2
Numerical Constants	Model		
Constants per gene	10		
Data type	Floating- Point		
Lower Bound	-10		
Upper Bound	10		

4.4 Random Forest

Random forest (RF) [82] is another effective machine learning (ML) strategy that works by creating a large number of decision trees during training time to address classification, regression, and other tasks, which was proposed by Ho in 1995. And Leo Breiman developed an extension of the algorithm [83]. The primary aim of this study is to forecast the shear capacity of the RC beam with stirrups by focusing exclusively on the regression model. Before averaging the results, the Random Forest (RF) generates the

appropriate number of regression trees. Averaging the output of all separate trees yields the final predicted results. Finally, the following equation gives the RF regression predictor after l trees $\{T_l(x)\}$ have been produced [82].

$$f(x) = \frac{\sum_{l=1}^{L} T_l(x)}{L} \tag{18}$$

The RF regression method is a non-parametric regression approach comprised of a set of l trees $\{T_1(X), T_2(X), ..., T_l(X)\}$, where $X = \{x_1, x_2, ..., x_\beta\}$ is a β-dimension input vector that forms a forest—output k correspondent to each tress $Y_k = (k = 1, 2, ..., k)$.

For each RF regression tree building, a new training set is created, replaced by the previous training set. As a result, a randomly picked training sample is used each time a regression tree is made from the original experimental dataset. The out-of-bag (OOB) example is used to check for sample accuracy [84].

$$GI(t_{X(x_i)}) = 1 - \sum_{j=1}^{m} f(t_{X(x_i)}, j)^2 \tag{19}$$

When independent test data is used, the inherent validation qualities of random forests improve tree robustness. The result is obtained by averaging the tree estimations.

4.5 Gradient Boost Regressor

Leo Breiman observed that boosting can be viewed as an optimization procedure on an appropriate cost function, which gave rise to the concept of gradient boosting. This is defined by introducing and

minimizing a loss function $K(y, f(x))$, where the number of iterations M :

$$f_0(x) = arg\min_{\gamma} \sum_{l=1}^{L} K(y_i, \gamma) \tag{20}$$

Calculate a type of residual known as a pseudo-residual, where m = 1 to *M*, and $i = 1, 2, 3, ..., n$

$$r_{im} = -\left[\frac{\partial K(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{m-1}(x)} \tag{21}$$

Solve the one-dimensional optimization problem below to find a multiplier γ_m. The equation in the down,

$$\gamma_m = arg\min_{\gamma} \sum_{l=1}^{L} K(y_i, f_{m-1}(x) + \gamma h_m(x_i)) \tag{22}$$

For the training, set $\{(x_{i,} y_i)\}_{i=1}^{n}$ update the model, which emphasis as below

$$f_m(x) = f_{m-1}(x) + \gamma_m h_m(x) \tag{23}$$

For the output defined after the *M* iteration as below,

$$f_M(x) = \sum_{l=1}^{L} \gamma_m h_m(x) \tag{24}$$

4.6 Artificial Neuronal Network (ANNs)

By inventing a computer framework in 1943, Warren McCulloch and Walter Pitts lay the foundation for neural networks, which create a future computing system [85]. ANNs are a type of soft computing approach inspired by the behaviour of the human nervous system. It is widely employed in civil engineering research and technology [86],[87],[88],[89],[90],[91]. Neural network architecture comprises several major components, including inputs, weights, a sum function, an activation function, and outputs. The weights

modify the input signals and are added to the bias term as specified in Eq. (25) [92]. **Fig. 5** shows a schematic design of an artificial neuron model.

$$y_i = f\left(\sum_{i=0}^{L} x_i w_i - b\right) \tag{25}$$

where w_i represent the weight for input x_i and bias defined as b. L is the number of neurons and y is the output term.

The sigmoid function is employed as the activation function in multilayer feed-forward ANNs. The sigmoid function is a mathematical function expressed in Eq. (26) [93].

$$f(k) = \frac{1}{1+e^{-\alpha k}} \quad \text{where,} \quad k = \sum_{i=0}^{L} x_i w_i - b \tag{26}$$

where α is a constant, which is a shape parameter in Eq. (26) used to control the gradient of the semilinear structure.

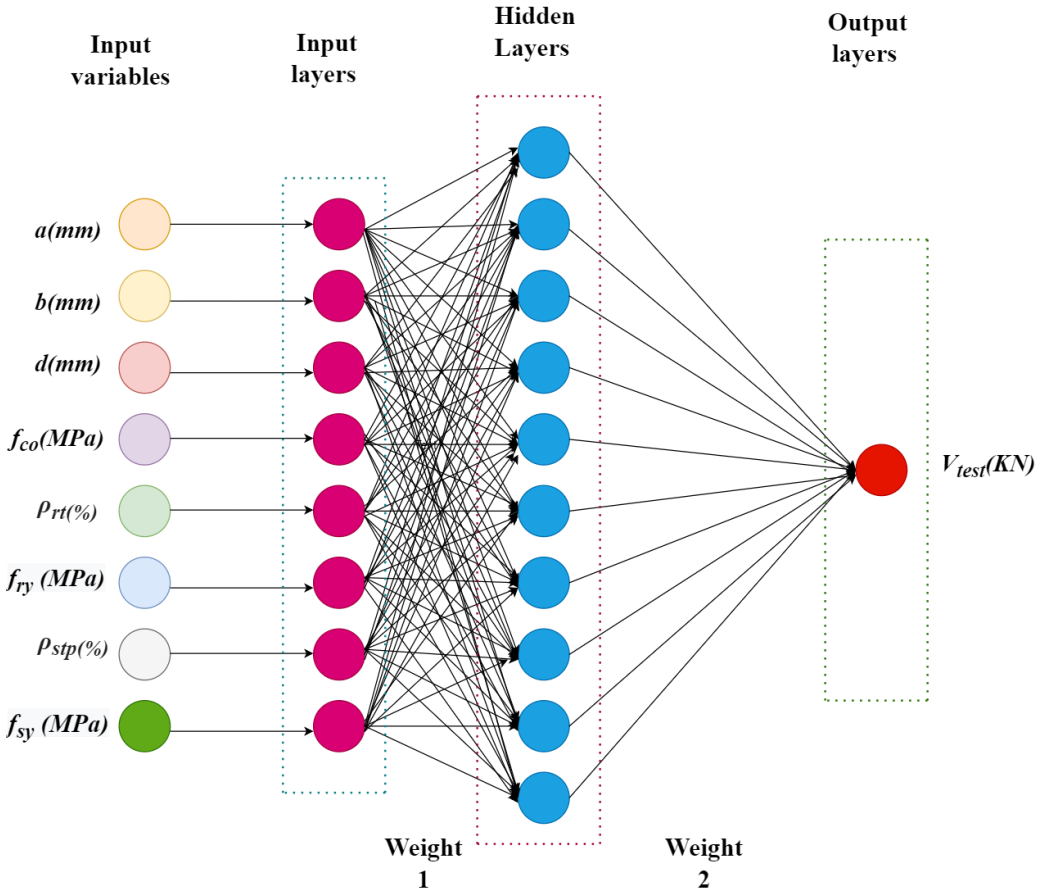


Fig. 5. The architecture of artificial neuronal networks (ANNs).

5 Results and Discussion

Table 4

Statistical indicators of the development of the machine learning model for the training and testing Phase

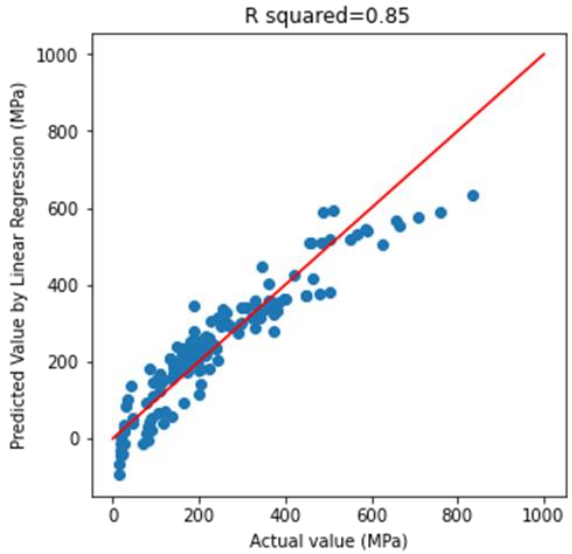
Training				Testing		
Statistical parameter	**R^2**	**MSE**	**RMSE**	**R^2**	**MSE**	**RMSE**
Linear Regression	0.85	4355.13	65.9	-1.43	52500.10	229.12

Ridge regression	0.87	3436.26	58.62	-0.93	50081.46	223.78
Random forest	0.98	359.06	18.94	0.94	1167.51	34.16
Gradient Boost Regressor	0.99	159.91	12.64	0.93	1293.72	35.96
ANNs	0.97	564.09	23.75	0.96	1150.31	34.52
GEP	0.92	344.1	44.54	0.95	388.2	44.98

5.1 Observation of the Result of Linear Regression and Ridge Regression

5.1.1 Developing the linear regression model and performance analysis

Multiple linear regression (MLR), often known as multiple regression, is a statistical technique for predicting the outcome of a predictor variable by combining many factors. MLR is two types which are including linear and nonlinear regression. Here we are focusing on the linear part of it.



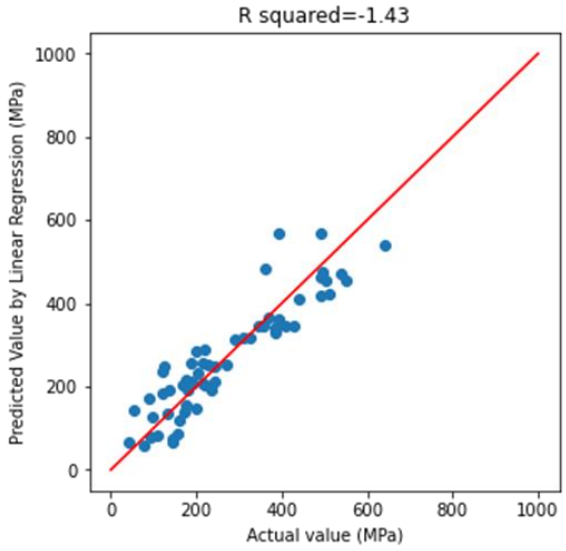


Fig. 6a. Training data fit in multiple linear regression model

Fig. 6b. Testing data fit in multiple linear regression model

As shown in Table 5, the multiple linear regression model is the model's performance of R^2, MSE, RMSE are calculated for training as 0.85, 4355.13, and 65.9, respectively. **Fig. 6a** shows that the multiple linear regression model is not efficient in estimating the shear capacity of the RC stirrups than GEP, and ANN, random forest, and gradient boost regressor model. **Fig. 6b** shows the negative result for R^2 -1.43, and MSE and RMSE are also calculated for testing 52500.10 and 229.12, respectively, which is much lower than another machine learning algorithm.

5.1.2 Developing the ridge regression model and performance analysis

Ridge regression is a model tuning technique that can be used to analyze data with multi-collinearity. Calculating the coefficients of multiple-regression models in cases with strongly correlated independent variables.

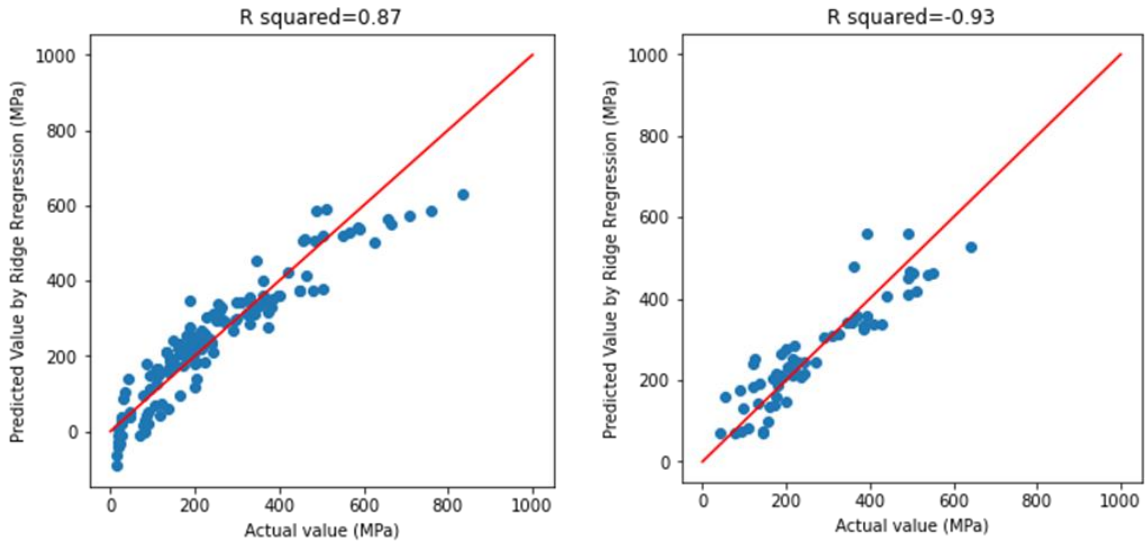


Fig. 7a. Training data fit in ridge regression model.

Fig. 7b. Testing data fit in ridge regression model.

The ridge regression model performance of R^2, MSE, RMSE are calculated for training as 0.87, 3436.26, and 58.62, respectively shown in Table 5. **Fig. 7a** shows that the ridge model is not efficient in estimating the shear capacity of the RC stirrups than GEP, and ANN, random forest, and gradient boost regressor model but it is showing better result compared to the multiple linear regression. **Fig. 7b** shows the negative result for R^2 -0.93, and MSE and RMSE are also calculated for testing 50081.46, and 223.78 respectively, which is much lower than another machine learning algorithm.

5.2 Observation of the Result between GEP and ANN model

5.2.1 Developing the GEP model and performance analysis

Eq. (28) shows a simplified relationship based on the GEP technique for estimating the RC beam with stirrups. After multiple attempts, the model was finalized with a minimum objective value of 0. 085. The best GEP model is shown below

$$V_{test} = d_1 - \left[\frac{d_0}{\left(\left(\frac{d_5}{c_8} \right) (c_7 - c_9) - d_4 \right)} - (d_6 \times d_1) \right] + \frac{c_8}{\left(\frac{2d_3 - \left(\frac{d_0}{d_3} \right) - d_7}{d_5} \right) \times d_2}{d_0}$$

$$+ \left(\frac{c_6}{d_6} \right) - \left(\frac{(d_2 - d_5) \times (d_5 - c_6) \times (c_1 \times d_6)}{(d_0 - d_7) + (d_4 \times d_5)} \right) \tag{27}$$

$$+\left(\frac{c_1}{d_6}\right)\left(d_5+\left(2d_3-\left(d_6-d_1\right)\right)\right)\times\left(\frac{d_0}{d_1}\right)\times\left(\frac{d_6}{d_2}\right)$$

$$+\left(\frac{d_2}{\left(c_1-c_6\right)}+\left(\frac{d_6}{c_5}\right)\times c_3\right)\times\left(\frac{d_3}{\frac{\left(d_5\times c_0\right)}{c_0+d_3}}\right)$$

$$+c_0\times\left(\frac{\left(2d_5\right)-\left(c_6-d_0\right)-\left(d_5\times c_4\right)-\left(d_0-d_3\right)}{\left(d_2\times d_1\right)+\left(c_4\times d_5\right)}\right)$$

$$+2d_4\left(\frac{c_8}{c_5}\right)\times\left(\frac{d_4}{d_6}\right)+\left(\frac{\left(d_4+d_5\right)}{d_3-d_7}\right)+d_4$$

The sub-ETs in Fig. 4 are used to create the preceding expression program, which comprises mathematical functions, constants, and variables in the GEP model. Seven sub-ETs (genes investigated in the model development) contain the complex solution's multiple components to the modelled problem and are linked by the addition mathematical operator. In Table 6, sub-ETs variables notation and constrain were reading in the direction of left-to-right and top-to-bottom. The parameters in **Fig. 4** are more complex and hierarchical structures, denoted by the letter "d" while the constants are indicated by the letter "c", a digit precedes both characters. Each sub-ET reflects a distinct aspect of the problem that must be addressed to produce a meaningful solution.

The process was repeated until no substantial change in R squared values or the fitness function was seen. After the method is stopped, other statistical metrics such as RMSE, MSE, and the trained model's OBJ function are calculated. Following iterations, the number of genes, head size, and arithmetic operations in the functions set increases while a linking function is established.

Table 3 summarizes the best parameters utilized to model the GEP-based model. After analyzing the model performance with different GEP models' variables in the training and testing sets, the final model was picked. The number of iterations in evolution also significantly affects a model's success. In the model, the best fitness is 21.95, and for the training and testing, the best fitness is 731.21 and 731.24. After 66304 generations, the model was terminated, and the method was completed when no significant changes were seen.

Table 5

Different variables notations as well as constants values in the model

Variable		**Constants**													
		Sub-ET1		**Sub-ET2**		**Sub-ET3**		**Sub-ET4**		**Sub-ET5**		**Sub-ET6**		**Sub-ET7**	
d_0	a	C_8	8.5	C_8	8.3	C_6	-6.5	C_1	-0.7	C_3	4.6	C_0	0.9	C_8	-0.9
d_1	b	C_7	9.1			C_1	-1.0			C_5	-10.6	C_4	-9.8	C_5	-5.9
d_2	d	C_9	-11							C_0	-4.4	C_6	-0.4		
d_3	f_{co}									C_1	-8.2				
d_4	ρ_{rt}									C_6	-0.9				
d_5	f_{ry}														
d_6	ρ_{stp}														

d_7	f_{sy}

Numerous statistical metrics, such as the root mean square error, the root mean square error, and r squared, are calculated further to emphasize the model's performance throughout the training stage. As shown in Table 5, the GEP model is determined the R^2, MSE, RMSE is calculated for training as 0.92, 344.1, 44.54, respectively and for the testing, the R^2, MSE, RMSE are calculated for testing as 0.95, 388.0, 44.98, respectively. **Fig. 8** shows that the GEP model is more efficient in estimating the RC stirrups' shear capacity than linear regression and ridge regression. And the ANN is showing much better results compared to the GEP model.

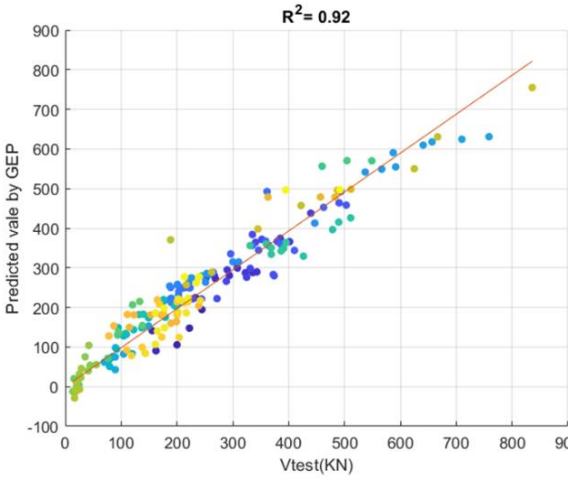


Fig. 8. The relationship between observed and expected shear capacity that the GEP model provides.

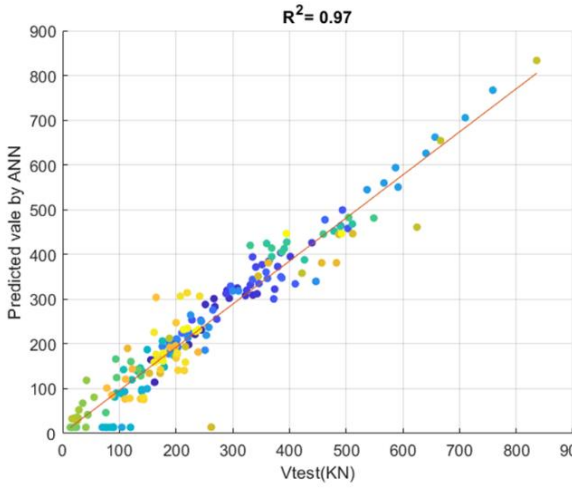


Fig. 9. The relationship between observed and expected shear capacity that the ANN model provides.

5.2.2 Developing the ANN model and performance analysis

As illustrated in **Fig. 4**, the ANN model comprises three layers: input, output, and hidden. The number

of neurons in the input and output layers equals the number of variables in the input and output layers. Thus, the input layer contains eight neurons, and the output layer has one. Whereas neurons in the buried layer are discovered using a hit-and-miss strategy. The ANN model predicts the shear capacity via training and testing. This model was built using 201 datasets, of which 70%, 30%(i.e., 141, 60) were utilized for training and testing, respectively. This model was chosen using the levenberg–marquardt (LM) technique. Ten hidden layers were used to test the ANN model.

As shown in Table 5, the ANN model is determined as an $R^2 = 0.97$. The MSE, RMSE are also calculated for training as 564.09, and 23.75, respectively. **Fig. 9** shows that the ANN model is more efficient in estimating the shear capacity of the RC stirrups than Linear regression and Ridge regression. Random forest and ANN are showing similar results, which is better than the GEP model

5.3 Observation of the Result Between Random Forest and Gradient Boost Regressor

5.3.1 Developing the random forest model and performance analysis

Random forest is a well-known machine learning method that works in multivariable problems. It gives outstanding accuracy across broad selection options of classification and regression predictive modelling problems [83]. Like bagging, random forest entails building many decision trees from bootstrap samples from the training dataset.

As shown in Table 5, the random forest model is the model's performance determined R^2 0.98. The MSE, RMSE are also calculated for training as 359.06, and 18.94 respectively. **Fig. 12a** shows that the random

forest model is more efficient in estimating the shear capacity of the RC stirrups than linear regression and ridge regression, GEP, and ANN models. Random forest and ANN are showing similar results in training, but they show less performance than the ANN model for the testing. **Fig. 12b** illustrates the R^2 value as 0.94 for the testing model, less than the training model in random forest.

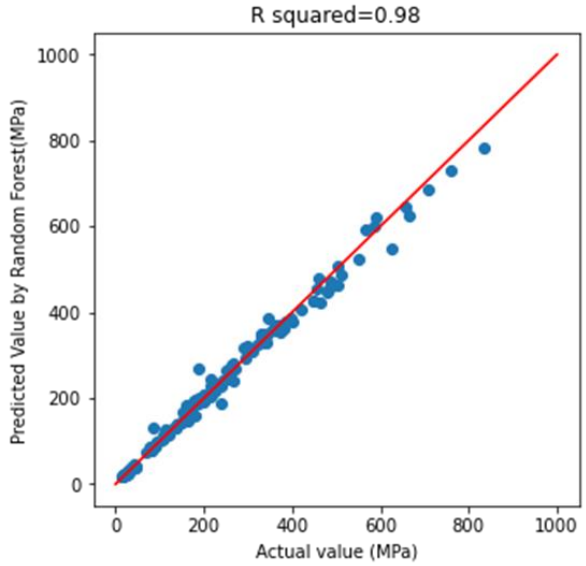


Fig. 12a. Training data fit in the random forest model.

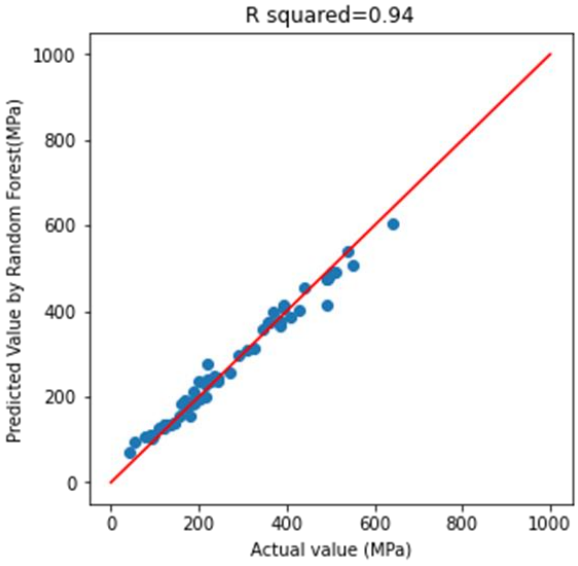


Fig. 12b. Testing data fit in the random forest model.

5.3.2 Developing the gradient boost regressor model and performance analysis

Gradient boosting is a highly effective technique for developing predictive models. Gradient boosting is effective when used with the loss function, weak learners, and the additive model. It enhances the performance of the underlying algorithm using several regularization approaches. The function [94] (weak hypothesis) demonstrates how a functional gradient approach of boosting has resulted in the creation of boosting algorithms in a wide variety of domains of machine learning and statistics, in addition to regression

and classification.

As shown in Table 5, the random forest model is the model's performance determined R^2 0.99. The MSE, RMSE are also calculated for training as 159.91 and 12.64, respectively. **Fig. 13a** shows that the gradient boost regressor is more efficient in estimating the RC stirrups' shear capacity than linear regression and ridge regression, GEP, and ANN model. Gradient boost regressor showing better results than for training and testing. Table 5 shows that the R^2, MSE, and RMSE are calculated for testing as 0.93, 1293.72, and 35.96, respectively. Fig. 13b illustrates the R^2 value as 0.93 for the testing model, less than the training model in gradient boost regressor.

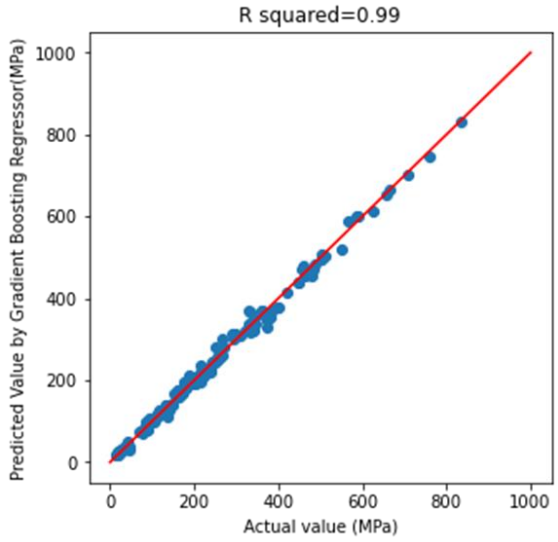


Fig. 13a. Training data fit in the gradient boost regressor model.

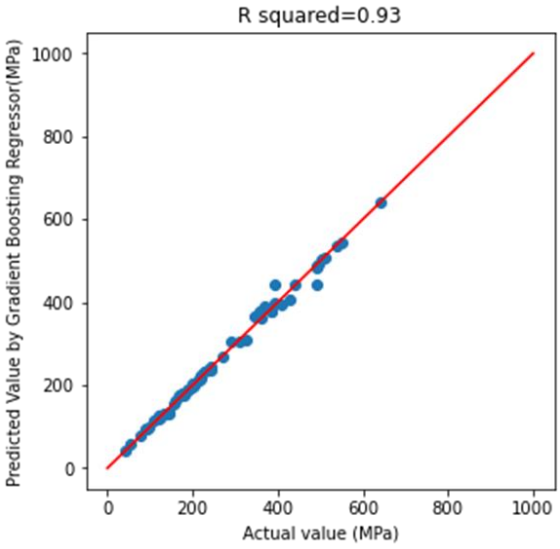


Fig. 13b. Testing data fit in the gradient boost regressor model.

5.4 Observation of the Result Between Actual Value and Predicted Value for Different Country Code

A total of 201 experimental test results were analyzed to determine the shear critical of an RC beam with stirrups only included in the model. To prevent the effect of arching, beams with a shear span-to-depth ratio $a/d < 2.5$ were omitted from consideration. **Fig. 15** illustrates the distribution of the effective depth, reinforcement ratio, and concrete compressive strength of the beams employed in this investigation. It can be observed from **Fig. 16** that the effective depth, beam weight, and shear span have a height influence on the predicted value.

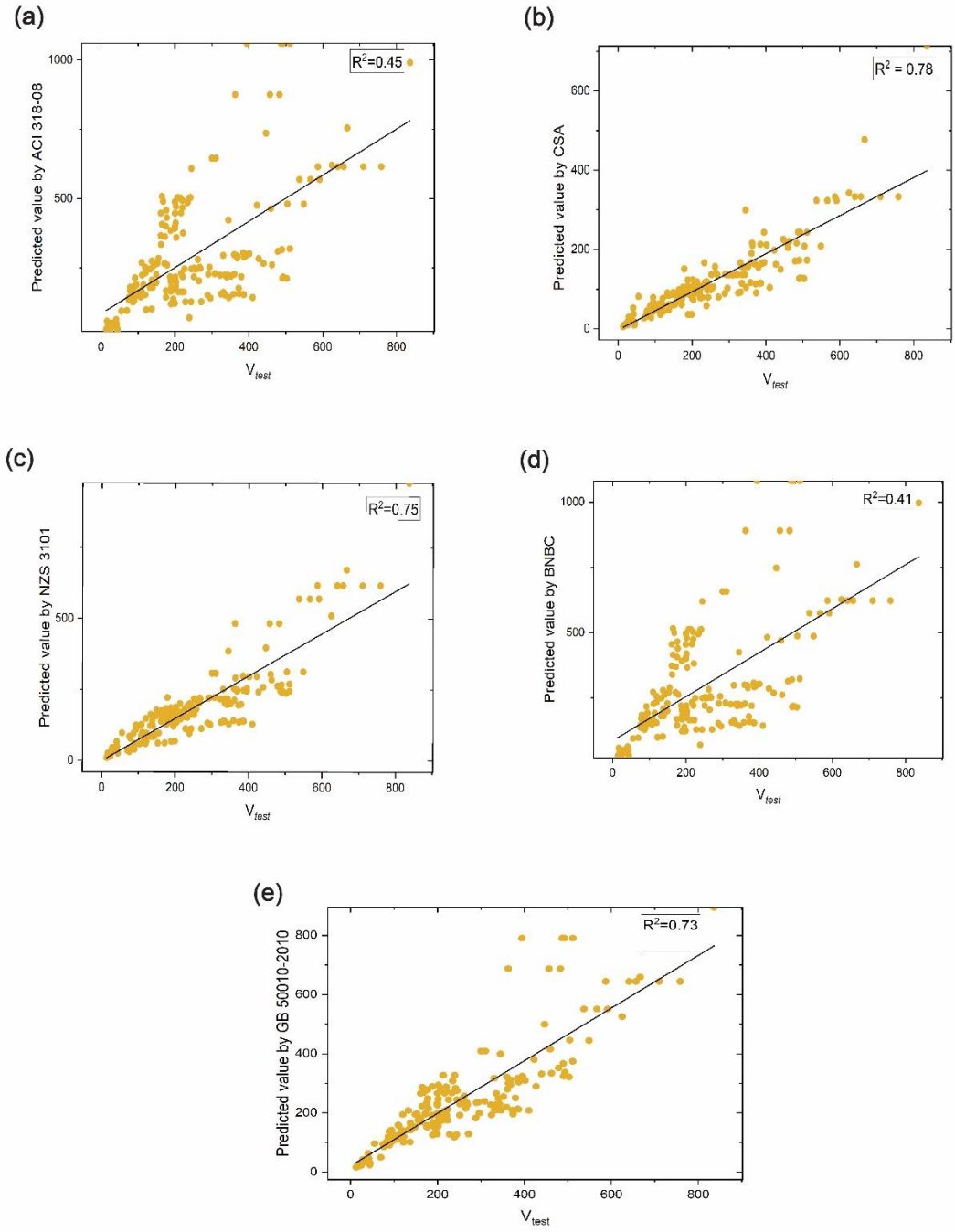


Fig. 14 Linear relation between actual value and predicted value: (a) ACI 318-08; (b) CSA A23.3-04 ;(c) NZS 3101; (d) BNBC 2015; (e) GB 50010-2010.

Fig. 14 illustrates that the relationship between experimental shear capacity and predicted shear force

using various codes. When these values are compared to the R squared value of 0.78, CSA A23.3-04 has the best-fit data points. NZS 3101 and GB 50010-2010 both evaluate the shear capacity of beams with 0.75 and 0.73 shear strengths, respectively. ACI 318-08 and BNBC 2015 have the same shear capacity results of expected values 0.45 and 0.41, respectively.

Table 7 presents a detailed examination of the expected shear resistance from various codes using regression analysis. As can be seen from this table, GB 50010-2010 has the lowest standard deviation of 0.32 and the lowest mean of 1.06. The ACI 318-08 and CSA A23.3-04 standard deviation is 0.66 and 0.69, but CSA A23.3-04 is showing higher result in mean 2.23 while ACI 318-08 is 1.33.

The machine learning model shows the relationship between experimental shear capacity and predicted shear force using various models in Table 7. Random forest and gradient boost regressor show the lowest mean of 0.98 and 0.985 and standard deviation of 0.078 and 0.075, respectively. It shows a significant correlation between the predicted value and the proposed ML model in the paper. ANN and GEP model is showing for mean 0.993 and 1.01 and standard deviation 0.147 and 1.02 respectively. ANN model is better than the GEP model to identifying shear capacity in this database. Multilinear regression and ridge regressor showing not good results compared to the other ML model. To satisfy the powerful machine learning model RMSE, MSE is also calculated in Table 5. The r squared value (Table 4) shows the best correlation between the actual and predicted value for the training and testing stage, which describes powerful ML algorithms with high prediction capacity for large-scale data. It is also suggested that more databases will give more

accurate results to implement those powerful algorithms in the future.

Fig.15 displays a/d the shear capacity for the sample RC beam with the stirrups database. Where the residual assumption the variability is approximately about the same on various data in the x-axis. For random forest, gradient boost regressor, ANN, and GEP models show almost the same sequence. This model shows the mean is normally distributed, with no systematic curvature and non-normality, indicating no real obvious problem with the assumption model. Table 7 illustrates that mean and standard deviation are compared to lower, whereas the existing code has a higher result—gradient boost regressor showing the best result among all leering machine models.

Fig. 15 also shows that ACI 318 -08 and BNBC 2015 have lower results than the proposed robust machine learning algorithm. CSA A23.3-04, NZS 3101, and GB 50010-2010 show the variable distribution is almost the same in the x-axis and no systematic curvature generated. It means the data is normally distributed, and predicted values are valid.

Table 6

Variation of proposed shear strength for RC beam with stirrups.

Compare results	**Statistical approach**	
	Mean	**Standard Deviation**
$\frac{V_{test}}{ANN}$	0.993	0.147
$\frac{V_{test}}{RF}$	0.980	0.078

$\frac{V_{test}}{GBR}$	0.985	0.075
$\frac{V_{test}}{GEP}$	1.01	1.02
$\frac{V_{test}}{LR}$	1.04	2.609
$\frac{V_{test}}{RidgeR}$	1.175	3.321
$\frac{V_{test}}{ACI\ 318-08}$	1.33	0.66
$\frac{V_{test}}{CSA\ \text{A23.3}-04}$	2.23	0.69
$\frac{V_{test}}{\text{GB 50010-2010}}$	1.06	0.32
$\frac{V_{test}}{NZS\ 3101}$	1.42	0.51
$\frac{V_{test}}{BNBC\ 2015}$	1.35	0.68

It can be observed that the machine learning model (gradient boost regressor) which is the best model to predict the RC beam with stirrups. The same accuracy is adopted and satisfies the predicted best model for training and testing using different civil engineering problems to identify shear capacity.

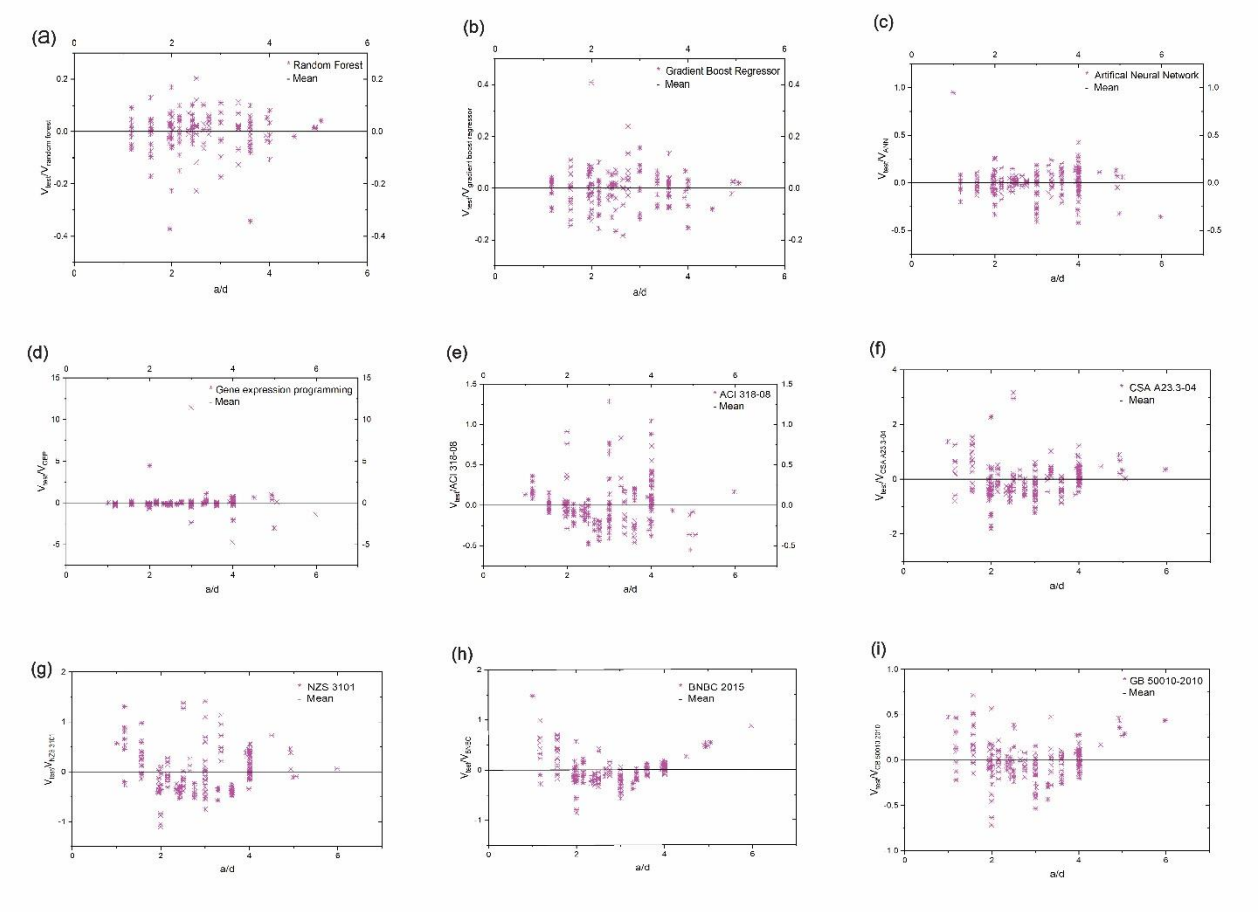


Fig. 15. Shear capacity result predicted by diffecnt mechanics-based model : (a) Random Forest ;(b) Gradent Boost Regressor ;(c) Artifical Neural Network; (d) Gene Expression Programming;(e) ACI -318-08; (f) CSA A23.3-04;(g) NZS 3101; (h) BNBC 2015;(i) GB 50010-2010.

5.5 Sensitivity analysis

The relative strength of the effect (RSE) of an input "i" unit on an output unit "j" is a quantity that can formerly quantify the relative influence of input elements on output. The greater the absolute value of RSE, the greater the effect on the output unit of the corresponding input unit. Sensitivity analysis was used to verify the relative importance of each parameter in the machine learning model [95]. To use this technique, all data pairs were used to construct a data array X in the following manner:

$$X = \{x_1, x_2, x_3, ...x_i, ..., x_n\} \tag{28}$$

Here, x_i the variable in the array X is the length vector of *t* as

$$x_i = \{x_{i1}, x_{i2}, x_{i3}, ..., x_{it}\} \tag{29}$$

The following equation expresses the strength of the relationship (r_{ij}) between datasets x_i and x_j .

$$r_{ij} = \frac{\sum_{t=1}^{k} x_{it} x_{jt}}{\sqrt{\sum_{t=1}^{k} x_{it}^2} \sqrt{\sum_{t=1}^{k} x_{jt}^2}} \tag{30}$$

The strengths of the relationships (r_{ij} values) between the input and observed (V_{test}) parameters are depicted in Fig. 16. x_{it} and x_{jt} is the *t* input variables. The data indicate that $a(mm)$, $b(mm)$, $d(mm)$ and $f_{sy}(MPa)$ have the most significant influence on V_{test} .To calculate the sensitivity the gradient boost regressor has been used to see the influence of the variables.

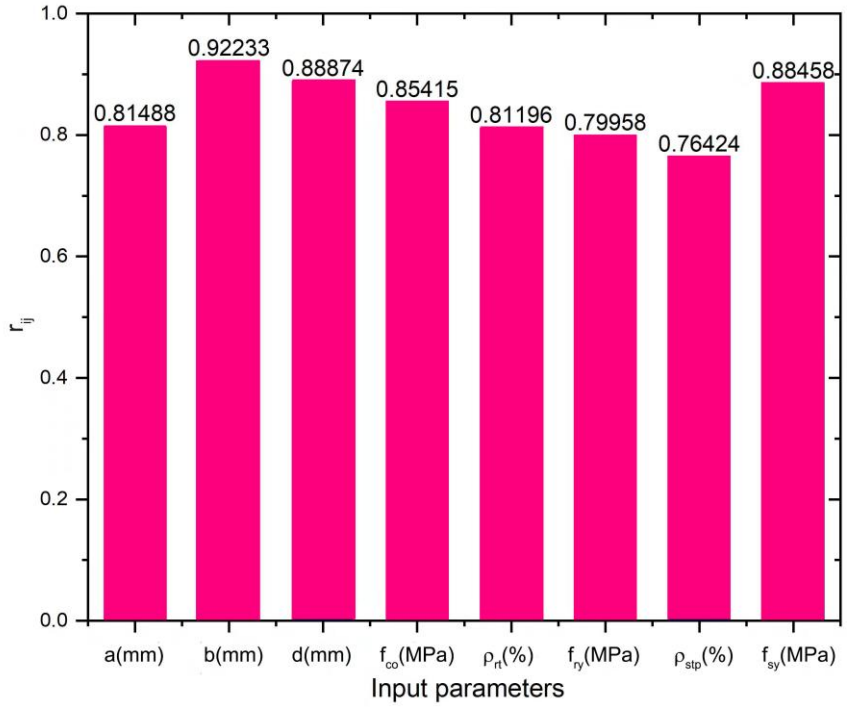


Fig. 16. Relationship strength between input and output parameters.

6 Conclusion

A machine learning-based algorithm is proposed for estimating the shear capacity for RC beam with stirrups. The objective of this research was to assess the predictive performance of six machine learning algorithms for estimating the shear strength of an RC beam with stirrups: multiple linear regression, ridge regressor, GEP, ANN, random forest, and gradient boost regressor. The result shows that the robust machine learning algorithm strongly influences calculating the shear strength of the RC beam. Among the six models, gradient boost regressor, random forest, and ANN have the highest prediction performance; as a result, we conclude that gradient boost regressors, random forests, and artificial neural networks are all suitable for determining the shear strength of an RC beam with stirrups. The following conclusions are taken from the

findings of this study:

1. The analysis of data from various experiments demonstrates that the critical parameters governing the shear capacity of RC beams with stirrups are compressive strength, shear span, web width, effective depth, transverse reinforcement ratio, transverse reinforcement yielding strength, longitudinal reinforcement yielding strength, longitudinal reinforcement yielding ratio, and longitudinal reinforcement yielding strength. Sensitivity analysis shows shear span, web width, and effective depth are mostly influenceable on the dataset.

2. Machine learning algorithms such as gradient boost regressor, random forest, and ANN effectively predict the shear strength of RC beam with stirrups. In this study, gradient boost regressor is the best among all machine learning models, which have 0.99 % accuracy for training.

3. Prediction results reveal that the gradient boost regressor and the random forest model are more effective at predicting shear capacity. At the same time, the ANN algorithm shows satisfactory performance in estimating all datasets and training and testing data 0.97 and 0.96, respectively.

4. The proposed expressions using gradient boost regressor and random forest yield good prediction of shear capacity of RC beam with stirrups beams with R^2 of 99% and 0.98%. The developed machine learning model can be used for practical design purposes as they are derived

based on data from different experimental programs.

5. The suggested machine learning-based models were evaluated for accuracy against predictive models integrated with international design standards and literature (ACI 318-08, CSA A23.3-04, GB 510010-2010, NZS 3101, BNBC 2015). The best model has the closest R^2 value to one and the smallest MAE and RMSE values. The overall precision of the machine learning-based strength models beats that of existing models when these parameters are used.

6. In **Fig. 15**, the proposed design formula (random forest (RF), gradient boost regression, gene expression programming (GEP), and artificial neural network (ANNs)) is compared to the present design formula (ACI 318-08, Canadian code, GB 510010-2010, NZS 3101, BNBC 2015) on 201 tested beams, along with the AVG and standard deviation, and the number of unsafe predictions for the beam. The proposed design formula appears to be superior to the existing design code. It delivers the most consistent predictions (lowest mean value) and the lowest number of inaccurate estimates.

7. Future work will focus on developing multiple gene expression models that apply to a wide variety of attributes by incorporating additional test data and combining machine learning-based models with computational frameworks such as hybrid PSO-ANN and then optimizing with gene expression modeling—creating a novel design formula for predicting the shear capacity of RC beam.

7 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.

8 Author Contribution

Md Nasir Uddin conceived the research, designed research model and writing the manuscript. Ling-zhi Li supervised and editing the paper. Kequan Yu provided the research background, interpreted the results, Junhong Yu complied all the figures. T. Tafsirojjaman and Wael Alhaddad revised the code. All authors read and approved the manuscript.

9 Acknowledgments

Gladly thank Tongji University for providing digital library support through research papers cited in this work.

References

[1] Bazant ZP, Sun Hsu-Huei. Size Effect in Diagonal Shear Failure: Influence of Aggregate Size and Stirrups. ACI Mater J 1987;84:259–72. https://doi.org/10.14359/1614.

[2] Amani J, Moeini R. Prediction of shear strength of reinforced concrete beams using adaptive neuro-fuzzy inference system and artificial neural network. Sci Iran 2012;19:242–8. https://doi.org/10.1016/j.scient.2012.02.009.

[3] Cladera A, Marí AR. Shear design procedure for reinforced normal and high-strength concrete beams using artificial neural networks. Part II: Beams with stirrups. Eng Struct 2004;26:927–36. https://doi.org/10.1016/j.engstruct.2004.02.011.

[4] Russo G, Somma G, Angeli P. Design shear strength formula for high strength concrete beams. Mater Struct Constr 2004;37:680–8. https://doi.org/10.1617/14016.

[5] Arslan G. Cracking shear strength of RC slender beams without stirrups. J Civ Eng Manag

2008;14:177–82. https://doi.org/10.3846/1392-3730.2008.14.14.

[6] Bentz EC, Vecchio FJ, Collins MP. Simplified modified compression field theory for calculating shear strength of reinforced concrete elements. ACI Struct J 2006;103:614–24. https://doi.org/10.14359/16438.

[7] Von Ramin M, Matamoros AB. Shear strength of reinforced concrete members subjected to monotonic loads. ACI Struct J 2006;103:83–92. https://doi.org/10.14359/15089.

[8] Bazant ZP, Kim JK. Size Effect in Shear Failure of Longitudinally Reinforced Beams. J Am Concr Inst 1984;81:456–68. https://doi.org/10.14359/10696.

[9] Zsutty TC. Beam Shear Strength Prediction by Analysis of Existing Data. ACI J Proc 1968;65. https://doi.org/10.14359/7526.

[10] Abuodeh OR, Abdalla JA, Hawileh RA. Prediction of shear strength and behavior of RC beams strengthened with externally bonded FRP sheets using machine learning techniques. Compos Struct 2020;234:111698. https://doi.org/10.1016/j.compstruct.2019.111698.

[11] Xu J, Zhou L, He G, Ji X, Dai Y, Dang Y. Comprehensive machine learning-based model for predicting compressive strength of ready-mix concrete. Materials (Basel) 2021;14:1–18. https://doi.org/10.3390/ma14051068.

[12] Salehi H, Das S, Chakrabartty S, Biswas S, Burgueño R. Structural damage identification using image-based pattern recognition on event-based binary data generated from self-powered sensor networks. Struct Control Heal Monit 2018;25:1–21. https://doi.org/10.1002/stc.2135.

[13] Datteo A, Lucà F, Busca G. Statistical pattern recognition approach for long-time monitoring of the G.Meazza stadium by means of AR models and PCA. Eng Struct 2017;153:317–33. https://doi.org/10.1016/j.engstruct.2017.10.022.

[14] Hooda Y, Kuhar P, Sharma K, Verma NK. Emerging Applications of Artificial Intelligence in Structural Engineering and Construction Industry. J Phys Conf Ser 2021;1950.

https://doi.org/10.1088/1742-6596/1950/1/012062.

[15] Chou JS, Tsai CF, Pham AD, Lu YH. Machine learning in concrete strength simulations: Multi-nation data analytics. Constr Build Mater 2014;73:771–80. https://doi.org/10.1016/j.conbuildmat.2014.09.054.

[16] Bandara RP, Chan TH, Thambiratnam DP. Structural damage detection method using frequency response functions. Struct Heal Monit 2014;13:418–29. https://doi.org/10.1177/1475921714522847.

[17] Leong HY, Ong DEL, Sanjayan JG, Nazari A, Kueh SM. Effects of Significant Variables on Compressive Strength of Soil-Fly Ash Geopolymer: Variable Analytical Approach Based on Neural Networks and Genetic Programming. J Mater Civ Eng 2018;30:04018129. https://doi.org/10.1061/(asce)mt.1943-5533.0002246.

[18] Oreta AWC, Kawashima K. Neural Network Modeling of Confined Compressive Strength and Strain of Circular Concrete Columns. J Struct Eng 2003;129:554–61. https://doi.org/10.1061/(asce)0733-9445(2003)129:4(554).

[19] Toghroli A, Suhatril M, Ibrahim Z, Safa M, Shariati M, Shamshirband S. Retraction Note to: Potential of soft computing approach for evaluating the factors affecting the capacity of steel–concrete composite beam. J Intell Manuf 2020;31:1311. https://doi.org/10.1007/s10845-019-01528-2.

[20] Uddin MN, Li LZ, Khan RKM, Shahriar F, Sob LWT. Axial Capacity Prediction of Concrete-Filled Steel Tubular Short Members Using Multiple Linear Regression and Artificial Neural Network. Mater Sci Forum 2021;1047:220–6. https://doi.org/10.4028/www.scientific.net/msf.1047.220.

[21] Sonebi M, Cevik A. Prediction of Fresh and Hardened Properties of Self-Consolidating Concrete Using Neurofuzzy Approach. J Mater Civ Eng 2009;21:672–9.

https://doi.org/10.1061/(asce)0899-1561(2009)21:11(672).

[22] Naderpour H, Mirrashid M. Shear Failure Capacity Prediction of Concrete Beam–Column Joints in Terms of ANFIS and GMDH. Pract Period Struct Des Constr 2019;24:04019006. https://doi.org/10.1061/(asce)sc.1943-5576.0000417.

[23] Santamaria JL, Morales L, Lima P. Neuro Fuzzy Inference Systems for Estimating Normal Concrete Mixture Proportions. Comput Civ Eng 2019 2019:43–50.

[24] Nazari A, Sanjayan JG. Modeling of Compressive Strength of Geopolymers by a Hybrid ANFIS-ICA Approach. J Mater Civ Eng 2015;27:04014167. https://doi.org/10.1061/(asce)mt.1943-5533.0001126.

[25] Zhang J, Ma G, Huang Y, sun J, Aslani F, Nener B. Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression. Constr Build Mater 2019;210:713–9. https://doi.org/10.1016/j.conbuildmat.2019.03.189.

[26] Matin SS, Farahzadi L, Makaremi S, Chelgani SC, Sattari G. Variable selection and prediction of uniaxial compressive strength and modulus of elasticity by random forest. Appl Soft Comput J 2018;70:980–7. https://doi.org/10.1016/j.asoc.2017.06.030.

[27] Omran BA, Chen Q, Jin R. Comparison of Data Mining Techniques for Predicting Compressive Strength of Environmentally Friendly Concrete. J Comput Civ Eng 2016;30:04016029. https://doi.org/10.1061/(asce)cp.1943-5487.0000596.

[28] Cook R, Lapeyre J, Ma H, Kumar A. Prediction of Compressive Strength of Concrete: Critical Comparison of Performance of a Hybrid Machine Learning Model with Standalone Models. J Mater Civ Eng 2019;31:04019255. https://doi.org/10.1061/(asce)mt.1943-5533.0002902.

[29] Pham A-D, Hoang N-D, Nguyen Q-T. Predicting Compressive Strength of High-Performance Concrete Using Metaheuristic-Optimized Least Squares Support Vector Regression. J Comput Civ Eng 2016;30:06015002. https://doi.org/10.1061/(asce)cp.1943-5487.0000506.

[30] Algaifi HA, Alqarni AS, Alyousef R, Bakar SA, Ibrahim MHW, Shahidan S, et al. Mathematical prediction of the compressive strength of bacterial concrete using gene expression programming. Ain Shams Eng J 2021. https://doi.org/10.1016/j.asej.2021.04.008.

[31] Shahmansouri AA, Akbarzadeh Bengar H, Jahani E. Predicting compressive strength and electrical resistivity of eco-friendly concrete containing natural zeolite via GEP algorithm. Constr Build Mater 2019;229:116883. https://doi.org/10.1016/j.conbuildmat.2019.116883.

[32] Chu HH, Khan MA, Javed MF, Zafar A, Ijaz Khan M, Alabduljabbar H, 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 2021. https://doi.org/10.1016/j.asej.2021.03.018.

[33] Mahdinia S, Eskandari-Naddaf H, Shadnia R. Effect of cement strength class on the prediction of compressive strength of cement mortar using GEP method. Constr Build Mater 2019;198:27–41. https://doi.org/10.1016/j.conbuildmat.2018.11.265.

[34] Kang MC, Yoo DY, Gupta R. Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. Constr Build Mater 2021;266:121117. https://doi.org/10.1016/j.conbuildmat.2020.121117.

[35] Sun J, Wang Y, Yao X, Ren Z, Zhang G, Zhang C, et al. Machine-learning-aided prediction of flexural strength and asr expansion for waste glass cementitious composite. Appl Sci 2021;11. https://doi.org/10.3390/app11156686.

[36] Alotaibi E, Mostafa O, Nassif N, Omar M, Arab MG. Prediction of Punching Shear Capacity for Fiber-Reinforced Concrete Slabs Using Neuro-Nomographs Constructed by Machine Learning. J Struct Eng 2021;147. https://doi.org/10.1061/(asce)st.1943-541x.0003041.

[37] Fu B, Feng DC. A machine learning-based time-dependent shear strength model for corroded reinforced concrete beams. J Build Eng 2021;36. https://doi.org/10.1016/j.jobe.2020.102118.

[38] Solhmirzaei R, Salehi H, Kodur V, Naser MZ. Machine learning framework for predicting failure mode and shear capacity of ultra high performance concrete beams. Eng Struct 2020;224:111221. https://doi.org/10.1016/j.engstruct.2020.111221.

[39] Shahnewaz M, Rteil A, Alam MS. Shear strength of reinforced concrete deep beams – A review with improved model by genetic algorithm and reliability analysis. Structures 2020;23:494–508. https://doi.org/10.1016/j.istruc.2019.09.006.

[40] Wang X, Liu Y, Xin H. Bond strength prediction of concrete-encased steel structures using hybrid machine learning method. Structures 2021;32:2279–92. https://doi.org/10.1016/j.istruc.2021.04.018.

[41] Nguyen MST, Kim SE. A hybrid machine learning approach in prediction and uncertainty quantification of ultimate compressive strength of RCFST columns. Constr Build Mater 2021;302:124208. https://doi.org/10.1016/j.conbuildmat.2021.124208.

[42] Hoang ND. Image Processing-Based Recognition of Wall Defects Using Machine Learning Approaches and Steerable Filters. Comput Intell Neurosci 2018;2018. https://doi.org/10.1155/2018/7913952.

[43] ACI Commettee 318. Building code requirements for structural plain concrete (ACI 318.08) and commentary. Farmington Hills, MI: American Concrete Institute; 2008.

[44] Canadian Standards Association. CAN/CSA A23.3-04 Design of Concrete Structures. CSA, Rexdale, Ontario, Canada: 2004.

[45] Chinese Concrete Code. Code for design of concrete structure(GB 50010–2010). Beijing: 2011.

[46] New Zealand Standard. Concrete structures standard (NZS 3101.1). 2006.

[47] Institute H and BR. The Bangladesh National Building Code (BNBC) 2015 (Final Draft). vol. 1. 2015.

[48] Bresler B, Scordelis AC. Shear strength of reinforced concrete beams. ACI J Proc 1963;60:51–

74.

[49] Clark AP. Diagonal tension in reinforced concrete beams. ACI J Proc 1951;48:145–56.

[50] Placas A, Regan PE. Shear failure of reinforced concrete beams. ACI J Proc 1971;68:763–73.

[51] Lee J, Choi I, Kim S. Shear behavior of reinforced concrete beams with high-strength stirrups. ACI Struct J 2011;108:620–9.

[52] Mattock AH, Wang Z. Shear strength of reinforced concrete members subject to high axial compressive stress. ACI J Proc 1984;81:287–98

[53] Anderson NS, Ramirez JA. Detailing of stirrup reinforcement. ACI Struct J 1989;86:507–15.

[54] Mphond AG, Frantz GC. Shear Tests of High- and Low-Strength Concrete Beams with Stirrups. Am Concr Inst 1985;87:179–96.

[55] Swamy RN, Andriopoulos AD. Contribution of aggregate interlock and dowel forces to the shear resistance of reinforced beams with web reinforcement. Am Concr Institute, Farmingt Hills, MI 1974:129–66.

[56] Lee J-Y, Hwang H-B. Maximum Shear Reinforcement of Reinforced Concrete Beams. ACI Struct J 2010;107:580–8.

[57] Tompos EJ, Frosch RJ. Influence of beam size, longitudinal reinforcement, and stirrup effectiveness on concrete shear strength. ACI Struct J 2002;99:559–67.

[58] Yoon Y-S, Cook WD, Mitchell D. Minimum shear reinforcement in normal, medium, and high-strength concrete beams. ACI Struct J 1996;93:576–584.

[59] Sarsam KF, Al-Musawi JMS. Shear design of high- and normal strength concrete beams with web reinforcement. ACI Struct J 1992;86:658–64.

[60] Xie Y, Ahmad SH, Yu T, Hino S, Chung W. Shear ductility of reinforced concrete beams of normal and high-strength concrete. ACI Struct J 1994;91:140–149.

[61] Elzanaty AH, Nilson AH, Slate FO. Shear capacity of reinforced concrete beams using high-

strength concrete. ACI J Proc 1986;83:290–6.

[62] Bresler B, Scordelis AC. Shear strength of reinforced concrete beams—Series II. 1964. https://doi.org/10.1128/AEM.69.8.4901-4909.2003.

[63] Bresler B, Scordelis AC. Shear strength of reinforced concrete beams—Series III. vol. Rep. No. 6. 1966.

[64] Krefeld WJ, Thurston CW. Studies of the shear and diagonal tension strength of simply supported reinforced concrete beams. ACI J Proc 1966;63:451–76.

[65] Frosch RJ. Behavior of large-scale reinforced concrete beams with minimum shear reinforcement. ACI Struct J 2000;97:814–20.

[66] Olalusi OB, Awoyera PO. Shear capacity prediction of slender reinforced concrete structures with steel fibers using machine learning. Eng Struct 2021;227:111470. https://doi.org/10.1016/j.engstruct.2020.111470.

[67] Pham BT, Son LH, Hoang TA, Nguyen DM, Tien Bui D. Prediction of shear strength of soft soil using machine learning methods. Catena 2018;166:181–91. https://doi.org/10.1016/j.catena.2018.04.004.

[68] Olalusi OB, Spyridis P. Machine learning-based models for the concrete breakout capacity prediction of single anchors in shear. Adv Eng Softw 2020;147:102832. https://doi.org/10.1016/j.advengsoft.2020.102832.

[69] Gao X, Lin C. Prediction model of the failure mode of beam-column joints using machine learning methods. Eng Fail Anal 2021;120:105072. https://doi.org/10.1016/j.engfailanal.2020.105072.

[70] Solhmirzaei R, Salehi H, Kodur V, Naser MZ. Machine learning framework for predicting failure mode and shear capacity of ultra high performance concrete beams. Eng Struct 2020;224:111221. https://doi.org/10.1016/j.engstruct.2020.111221.

[71] Seltman HJ. Experimental Design and Analysis.
2008. http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf.

[72] Weisberg S. Applied Linear Regression Models. John Wiley & Sons, Inc. A; 2005.

[73] Khademi F, Akbari M, Mohammadmehdi S, Nikoo M. Multiple linear regression , arti fi cial neural network , and fuzzy logic prediction of 28 days compressive strength of concrete. Front Struct Civ Eng 2017. https://doi.org/10.1007/s11709-016-0363-9.

[74] Box GEP. Evolutionary operation: A method for increasing industrial productivity. Appl Stat 1957;VI:121–41. https://doi.org/10.1109/9780470544600.ch4.

[75] Ferreira C. Gene Expression Programming: a New Adaptive Algorithm for Solving Problems. Complex Syst 2001;13:87–129.

[76] Ferreira C. Automatically defined functions in gene expression programming. Stud Comput Intell 2006;13:21–56. https://doi.org/10.1007/11521433_2.

[77] Xie Z, Li X, Di Eugenio B, Nelson PC, Xiao W, Tirpak TM. Using gene expression programming to construct sentence ranking functions for text summarization 2004:1381-es. https://doi.org/10.3115/1220355.1220557.

[78] Saridemir M. Genetic programming approach for prediction of compressive strength of concretes containing rice husk ash. Constr Build Mater 2010;24:1911–9. https://doi.org/10.1016/j.conbuildmat.2010.04.011.

[79] Gandomi AH, Alavi AH, Mousavi M, Tabatabaei SM. A hybrid computational approach to derive new ground-motion prediction equations. Eng Appl Artif Intell 2011;24:717–32. https://doi.org/10.1016/j.engappai.2011.01.005.

[80] Gandomi AH, Alavi AH, Mirzahosseini MR, Nejad FM. Nonlinear Genetic-Based Models for Prediction of Flow Number of Asphalt Mixtures. J Mater Civ Eng 2011;23:248–63. https://doi.org/10.1061/(asce)mt.1943-5533.0000154.

[81] Azim I, Yang J, Iqbal MF, Mahmood Z, Javed MF, Wang F, et al. Prediction of Catenary Action Capacity of RC Beam-Column Substructures under a Missing Column Scenario Using Evolutionary Algorithm. KSCE J Civ Eng 2021;25:891–905. https://doi.org/10.1007/s12205-021-0431-0.

[82] Ho TK. Random decision forests. Proc Int Conf Doc Anal Recognition, ICDAR 1995;1:278–82. https://doi.org/10.1109/ICDAR.1995.598994.

[83] Breiman L. Random Forests. Mach Learn 2001:5–32.https://doi.org/10.1023/A:1010933404324.

[84] Zhou J, Li E, Wei H, Li C, Qiao Q, Armaghani DJ. Random forests and cubist algorithms for predicting shear strengths of rockfill materials. Appl Sci 2019;9:1–16. https://doi.org/10.3390/app9081621.

[85] McCulloch WS, Pitts W. A Logical Calculus of Ideas Immanent in Nervous Activity. Bull Math Biophys 1943;5:115–33. https://doi.org/10.1007/BF02478259.

[86] Meesaraganda LVP, Saha P, Tarafder N. Artificial neural network for strength prediction of fibers' self-compacting concrete. vol. 816. Springer Singapore; 2019. https://doi.org/10.1007/978-981-13-1592-3_2.

[87] Nagajothi S, Elavenil S. Influence of Aluminosilicate for the Prediction of Mechanical Properties of Geopolymer Concrete – Artificial Neural Network. Silicon 2020;12:1011–21. https://doi.org/10.1007/s12633-019-00203-8.

[88] Jang HS, Shuli X, Lee M, Lee YK, So SY. Use of artificial neural network for the simulation of radon emission concentration of granulated blast furnace slag mortar. J Nanosci Nanotechnol 2016;16:5268–73. https://doi.org/10.1166/jnn.2016.12268.

[89] Özcan F, Atiş CD, Karahan O, Uncuoğlu E, Tanyildizi H. Comparison of artificial neural network and fuzzy logic models for prediction of long-term compressive strength of silica fume

concrete. Adv Eng Softw 2009;40:856–63. https://doi.org/10.1016/j.advengsoft.2009.01.005.

[90] Uddin MN, Khan RKM, Chhattal M, Sagar A. Managing IoT to Establish Smart Cities for Sustainable Development. Proc. Int. Conf. energy, Resour. Environ. Sustain. Dev., 2019.

[91] Rajeshwari R, Mandal S. Prediction of compressive strength of high-volume fly ash concrete using artificial neural network. vol. 25. Springer Singapore; 2019. https://doi.org/10.1007/978-981-13-3317-0_42

[92] Kong X, Khambadkone AM. Modeling of a PEM fuel-cell stack for dynamic and steady-state operation using ANN-based submodels. IEEE Trans Ind Electron 2009;56:4903–14. https://doi.org/10.1109/TIE.2009.2026768.

[93] Anderson JA. Cognitive and Psychological Computation with Neural Models. IEEE Trans Syst Man Cybern 1983;SMC-13:799–815. https://doi.org/10.1109/TSMC.1983.6313074.

[94] Llew M, Jonathan B, Peter B, Marcus F. Boosting Algorithms as Gradient Descent. Adv Neural Inf Process Syst 12 MIT Press 1999:512–518.

[95] Jalali SH, Heidari M, Mohseni H. Comparison of models for estimating uniaxial compressive strength of some sedimentary rocks from Qom Formation. Environ Earth Sci 2017;76:1–15. https://doi.org/10.1007/s12665-017-7090-y.