frontiers | Frontiers in Materials

TYPE Original Research
PUBLISHED 30 March 2023
DOI 10.3389/fmats.2023.1159079

OPEN ACCESS

EDITED BY
Di Wang,
Aalto University, Finland

REVIEWED BY
Jiaolong Ren,
Shandong University of Technology, China
Mahdi Hasanipanah,
University of Technology Malaysia, Malaysia
Jinshun Xue,
Hubei University of Arts and Science, China

*CORRESPONDENCE
Mengmeng Zhou,
a7378576701@163.com

SPECIALTY SECTION
This article was submitted to Structural Materials, a section of the journal Frontiers in Materials

RECEIVED 05 February 2023
ACCEPTED 16 March 2023
PUBLISHED 30 March 2023

CITATION
Wang Q and Zhou M (2023), Development of hybrid SVM-FA, DT-FA and MLR-FA models to predict the flexural strength (FS) of recycled concrete.
Front. Mater. 10:1159079.
doi: 10.3389/fmats.2023.1159079

Development of hybrid SVM-FA, DT-FA and MLR-FA models to predict the flexural strength (FS) of recycled concrete

Qiang Wang and Mengmeng Zhou*

School of Mines, China University of Mining and Technology, Xuzhou, China

Recycled concrete from construction waste used as road material is a current sustainable approach. To provide feasible suggestions for civil engineers to prepare recycled concrete with high flexural strength (FS) for the road pavement, the present study proposed three hybrid machine learning models by combining support vector machine (SVM), decision tree (DT) and multiple linear regression (MLR) with the firefly algorithm (FA) for the computational optimization, named as SVM-FA, DT-FA, and MLR-FA, respectively. Effective water-cement ratio (WC), aggregate-cement ratio (AC), recycled concrete aggregate replacement ratio (RCA), nominal maximum recycled concrete aggregate size (NMR), nominal maximum normal aggregate size (NMN), bulk density of recycled concrete aggregate (BDR), bulk density of normal aggregate (BDN), water absorption of RCA (WAR) and water absorption of NA (WAN) were employed as the input variables. To determine the predicting results of varying hybrid models, root mean square error (RMSE) and correlation coefficient (R) were used as performance indexes. The results showed that the SVM-FA demonstrated the highest R values and the lowest RMSE values, and the fitting effect of the predicted values and the actual values of the FS of recycled concrete is the best. All the above analysis proving that the SVM optimized by FA hyperparameters has the highest prediction accuracy and SVM-FA can provide engineers a more accurate and convenient tool to evaluate the FS of recycled concrete. The results of sensitivity analysis showed that WC has the most significant influence on the FS of recycled concrete, while RCA has the weakest influence on the FS, which should be noticed when engineers apply recycled concrete to road design in the future.

KEYWORDS

machine learning, flexural strength, hyperparameter, firefly algorithm, root mean square error

1 Introduction

- As the urbanization process enters the era of accelerated development, the construction industry has achieved new heights in the process of development (Arumugam et al., 2018; Corches and Popa, 2018; Albuquerque et al., 2022). In order to adapt to the needs of urban development, residential buildings, infrastructure, roads and bridges and other facilities are constantly updated, a large number of old infrastructures have been dismantled, resulting in a huge amount of construction waste (Dong et al., 2018; Faruqi and Siddiqui, 2020). At present, the main way to deal with construction waste is centralized stacking or landfill, but due to the characteristics of concrete is difficult to degrade, the above treatment will cause environmental pollution and land waste, not in line with the theme of sustainable development (Hassan et al., 2018; Luangcharoenrat et al., 2019; Kim, 2021). At the same time, the construction of new infrastructure has a large demand for concrete, and the preparation of concrete will use natural sand and other non-renewable resources, the exploitation of these resources has caused serious damage to the ecological environment (Vilventhan et al., 2019; Nunes and Mahler, 2020; Yin, 2021; Xiao et al., 2022). Therefore, the recycling and utilizing of construction waste can not only avoid the environmental pollution and ecological damage brought by it, but also reduce the mining and utilizing of natural sand and stone, which is of great significance in practical engineering application and social and economic benefits (Wang et al., 2020a; Wang et al., 2020b). In recent years, in order to promote the recovery and utilization of construction waste, civil engineers put forward the method of making recycled concrete from construction waste into recycled aggregate (Chi and Pei, 2018; Liang et al., 2019; Li et al., 2020; Hasanipanah et al., 2021). Recycled concrete refers to a new type of concrete formed by recycled aggregate which is partially or completely replaced by natural aggregate after the waste concrete is processed by a specific technology (Liu et al., 2019; Lin et al., 2020; Pacheco et al., 2021; Shmlls et al., 2022). Recycled concrete is not only beneficial to reduce the pollution caused by construction waste to the environment, but also to relieve the burden of resource exploitation to the environment. The cost of recycled concrete with the same strength is lower than that of ordinary concrete (Qiao et al., 2018; Weglorz et al., 2018; Xiong et al., 2018; Pani et al., 2020).

- The road surface is in direct contact with the external environment and bears the driving load, meanwhile, it is also subjected to the rain erosion and the temperature change of freezing-thawing alternation, which puts forward higher requirements for the performance of concrete (Liu et al., 2019; Hoai-Bao and Quoc-Bao, 2021). FS is an important index to measure the quality of recycled aggregate concrete, and many researchers have carried out research on it. Xiao et al. (2013) used the laboratory test method to study the relationship between the compressive strength as well as the residual FS of recycled aggregate concrete with different recycled aggregates (0%, 30%, 50%, 70%, and 100%) and temperature, and the results showed that, both the residual compressive strength and the residual FS of recycled aggregate concrete increase with the increase of temperature, and they are affected by the replacement rate of recycled aggregate. In order to evaluate the feasibility of recycled aggregate concrete made of crushed clay brick and concrete block as recycled aggregate for filling wall materials, Chen et al. (2018) used orthogonal experiment method to study the influence of recycled brick particle content, water-cement ratio, lime content, aggregate cement ratio, aggregate replacement rate on the FS and compressive strength of recycled aggregate concrete, and used multiple regression analysis to analyze the relationship between recycled aggregate concrete and various factors. Kachouh et al. (2021) analyzed the effects of the replacement rate of recycled aggregate and steel fiber content on the FS of recycled aggregate concrete and desert dune sand by using three- and four-point bending tests, and proved that the replacement rate of recycled aggregate has a greater effect on the compression performance and steel fiber has a greater effect on the FS. Guan et al. (2022) evaluated the effects of steel fiber volume fraction, recycled aggregate replacement rate and concrete strength on the flexural properties of steel fiber reinforced recycled aggregate concrete beams using a combination of test and finite element method. The results showed that all specimens were broken in the compression zone, and the FS of steel fiber reinforced recycled aggregate concrete beams increase with the increase of the volume fraction of steel fiber and the strength of concrete, while the increase of the replacement rate of recycled aggregate decreases the FS of steel fiber reinforced recycled aggregate concrete beam. Yuan et al. (2022) proposed using gradient boosting and random forest to evaluate the compressive strength and FS of recycled aggregate concrete, and confirmed that the random forest model the predicted performance of random forest on recycled aggregate concrete better than gradient boosting by evaluating the error parameters and fitting parameters of the two machine learning models.
- However, the above methods used to evaluate the flexural strength of recycled aggregate concrete has disadvantages as follow:
- Laboratory testing involves casting, curing and testing samples, which requires a large amount of cost, substantial effort and time (Jamei et al., 2021), especially considering the influence of multiple factors on the flexural strength of recycled aggregate concrete, the amount of work will increase exponentially.
- Some researchers have realized that the prediction effect of hybrid machine learning models is better than that of single machine learning models (Guo and Wang, 2017; Wang et al., 2019; Zhu et al., 2021; Hasanipanah et al., 2022), and studied the evaluation effect of hybrid machine learning models, but it is necessary to compare the evaluation effect of different hybrid machine learning models and select the model with higher prediction effect.
- In order to compare the prediction effect of different hybrid machine learning models on the FS of recycled concrete, and select the model with high prediction accuracy for engineers as an environmentally friendly tool to evaluate the FS of recycled concrete, this study proposed to use the SVM-FA, DT-FA and MLR-FA to predict the FS of recycled concrete. There are many factors affecting the FS of concrete, and the influencing mechanism is also relatively complex (Shariati et al., 2020; Zhang et al., 2022). Research and analysis showed that the strength of concrete is mainly affected by the raw material quality, mix proportion of concrete and other factors (Jiang et al., 2019; Wang et al., 2021; Xie et al., 2022). In order to analyze the effect of the mix proportion of concrete and aggregate quality on the FS of recycled concrete, this study selected WC, AC, RCA, NMR, NMN, BDR, BDN, WAR and WAN as the input variables and collected 51 data sets from the previous published literatures. RMSE values, R values and the fitting effect between predicted values and actual values were

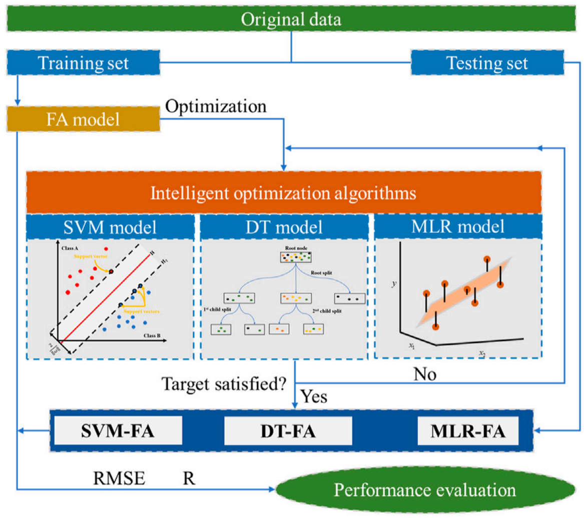

FIGURE 1
Frame flow chart.

used to evaluate the prediction accuracy of the models, and analyzed the sensitivity of recycled concrete to different input variables. The framework of this study is shown in Figure 1.

2 Materials and methods

2.1 Data collection

In this study, relevant data on the FS of recycled concrete were collected from the previous published literatures, and a reliable database was formed to verify the prediction effect of different machine learning models on the FS of recycled concrete. WC, AC, RCA, NMR, NMN, BDR, BDN, WAR, WAN are the input variables, FS is the output variable in this database. The distribution and statistical analysis of data sets are shown in Figure 2; Table 1. It can be seen from Figure 2; Table 1 that the coverage of data sets is wide, the mean and median are close, and the standard deviation and variance are small, proving the distribution of data sets is reasonable.

2.2 Correlation analysis

In order to ensure that the prediction effect of the models on the FS of recycled concrete is not affected by the high correlation between the input variables, the correlation between the nine input variables is analyzed before the models training, and the results are shown in Figure 3. It can be seen from the figure that the correlation coefficient of the same input variable is 1. Between different variables, the correlation coefficients between WAN and NMR (0.833), WAR and BDR (0.699) are the largest, but the correlation coefficients between other input variables are lower, less than 0.6. Overall, the correlation between input variables is low. Therefore, the prediction effect of the models will not be affected by the multicollinearity of the input variables, using them as input variables to verify the prediction effect of the models on the FS of recycled concrete.

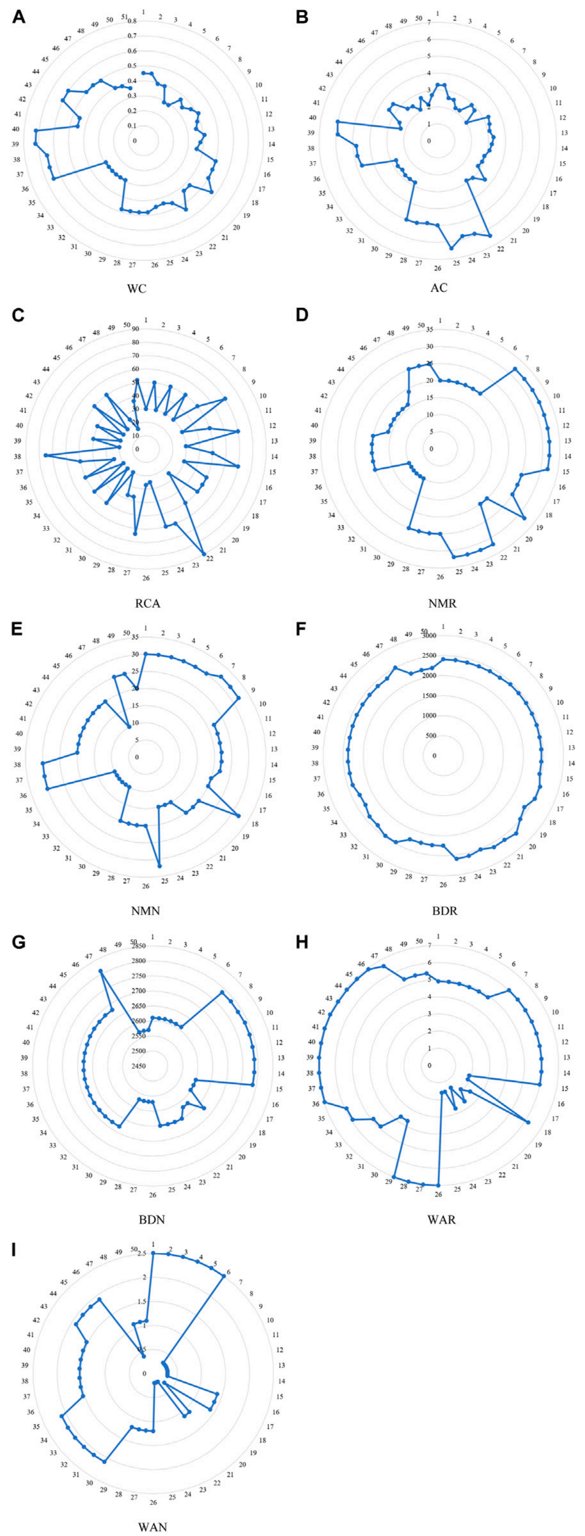

FIGURE 2
The data distribution of input variables [**(A)**: WC, **(B)**: AC, **(C)**: RCA, **(D)**: NMR, **(E)**: NMN, **(F)**: BDR, **(G)**: BDN, **(H)**: WAR, **(I)**: WAN].

TABLE 1 Statistical analysis of input variables data.

Variables	Minimum	Maximum	Range	Median	Average	Std.	Variance
WC	0.290	0.720	0.430	0.445	0.445	0.112	0.013
AC	2.000	6.400	4.400	2.950	3.496	1.236	1.529
RCA (%)	16.000	90.000	74.000	45.000	41.900	17.050	290.730
NMR	10.000	32.000	22.000	20.000	22.660	7.220	52.180
NMN	10.000	32.000	22.000	20.000	21.740	6.810	46.350
BDR	2200.000	2661.000	461.000	2400.000	2410.740	104.700	10961.470
BDN	2570.000	2810.000	130.000	2680.000	2668.340	71.830	5159.500
WAR	1.500	7.000	5.500	6.000	5.270	1.750	3.070
WAN	0.200	2.500	2.300	1.400	1.290	0.780	0.600

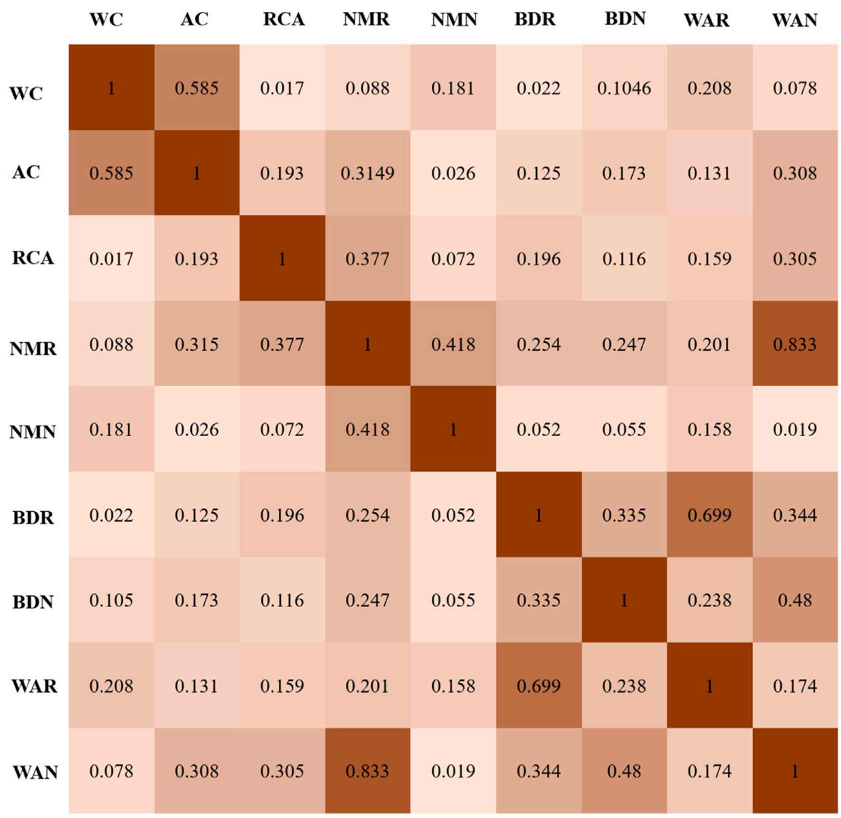

FIGURE 3
The correlation analysis of input variables.

2.3 Algorithm

2.3.1 Firefly algorithm (FA)

Firefly algorithm can not only optimize single-peak function and multi-peak function, but also has strong local search ability to find the optimal solution in a small area. Compared with other optimization algorithms, firefly algorithm has the advantages of convenient operation, simple implementation, fewer parameters, and less influence of parameters on the algorithm (Arora and Kaur, 2022). In nature, fireflies communicate with each other through their own fluorescence and will be attracted by fireflies with higher brightness than themselves and move towards them, thus completing the position update. FA needs to comply with the following three conditions: firstly, all fireflies are gender-neutral, and each firefly is attracted to the other; Secondly, the attraction between fireflies is only related to brightness and distance, and is

TABLE 2 The code of FA.

Firefly algorithm
Initialize a population of fireflies x_i $(i = 1, 2 \cdots n)$
Objective function $f(x_i)$, $x_i = (x_{i1}, x_{i2}, \cdots x_{im})^T$
Brightness I_i at x_i is determined by $f(x_i)$
Define light absorption coefficient γ
While $t < Maxiteration$
for $i = 1; n$
for $j = 1; n$
if $I_i < I_j$
Move firefly i towards j in m
end if
end for j
end for i
Rank the fireflies and find the current best
$t = t + 1$
end
Post process results and visualization

directly proportional to brightness and inversely proportional to distance (Rezaei and Rezaei, 2022). The fireflies with weaker brightness move to the fireflies with stronger brightness, while the fireflies with highest brightness move randomly (Huang et al., 2020); Finally, the brightness of fireflies is affected by the objective function. For the maximization problem, the brightness of fireflies is proportional to the objective function, while for the minimization problem, the brightness of fireflies is inversely proportional to the objective function (Huang et al., 2021). The code for the FA is shown in Table 2.

2.3.2 Support vector machine (SVM)

SVM is a common classification method suitable for solving small sample and non-linear problems (Huang et al., 2021; Li and Lv, 2021; Gupta et al., 2022). Based on the principle of structural risk minimization, SVM compresses the original data set into the support vector set, gives the rules of support vector determination in the process of learning new knowledge with subsets, and obtains the upper bound of learning error probability (Ahmad et al., 2021). The linear classification plane is defined as:

$$G(D) = \alpha \cdot D + a = 0 \tag{1}$$

In the formula, α is the weight vector of the classification plane, a is the classification threshold, and can be obtained by any support vector or by any pair of support vectors of the two classes. Through normalization of the discriminant function, all samples meet $|G(D)| = 1$, namely:

$$y_i\left[(\alpha \cdot D_i)^i + a\right] - 1 \geq 0 \tag{2}$$

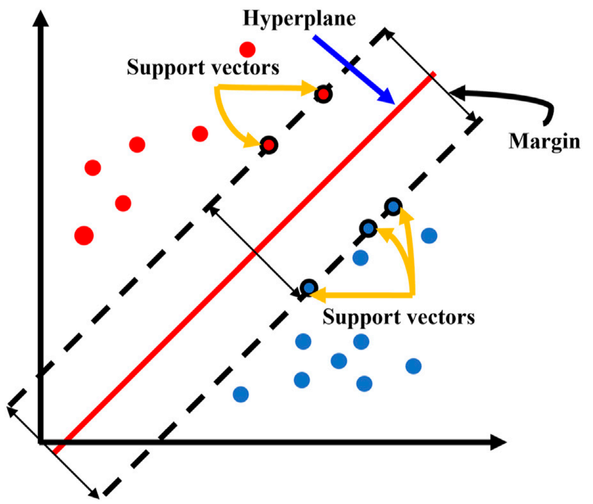

FIGURE 4
The schematic diagram of SVM.

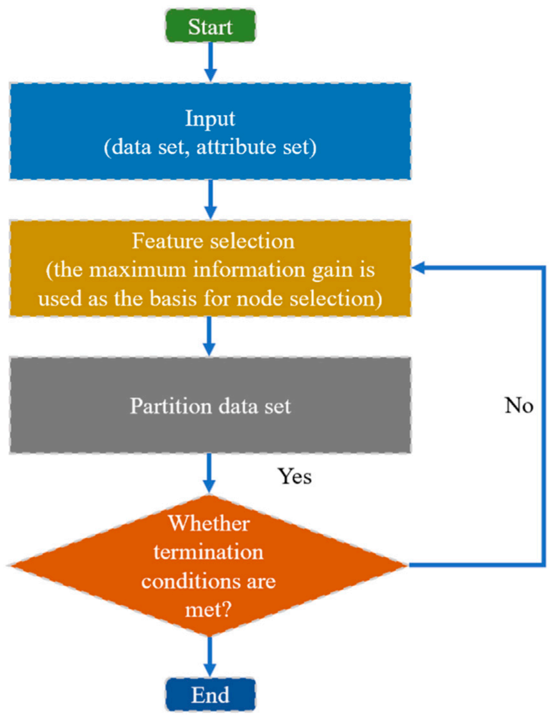

FIGURE 5
The flow chart of DT.

In the formula $i = 1, 2, \cdots N$, y_i is the category mark of the sample. If the sample belongs to class C, then $y_i = 1$, otherwise $y_i = -1$, D_i is the corresponding sample. The design goal is to minimize the interval value. According to the above analysis, Lagrange function is defined as:

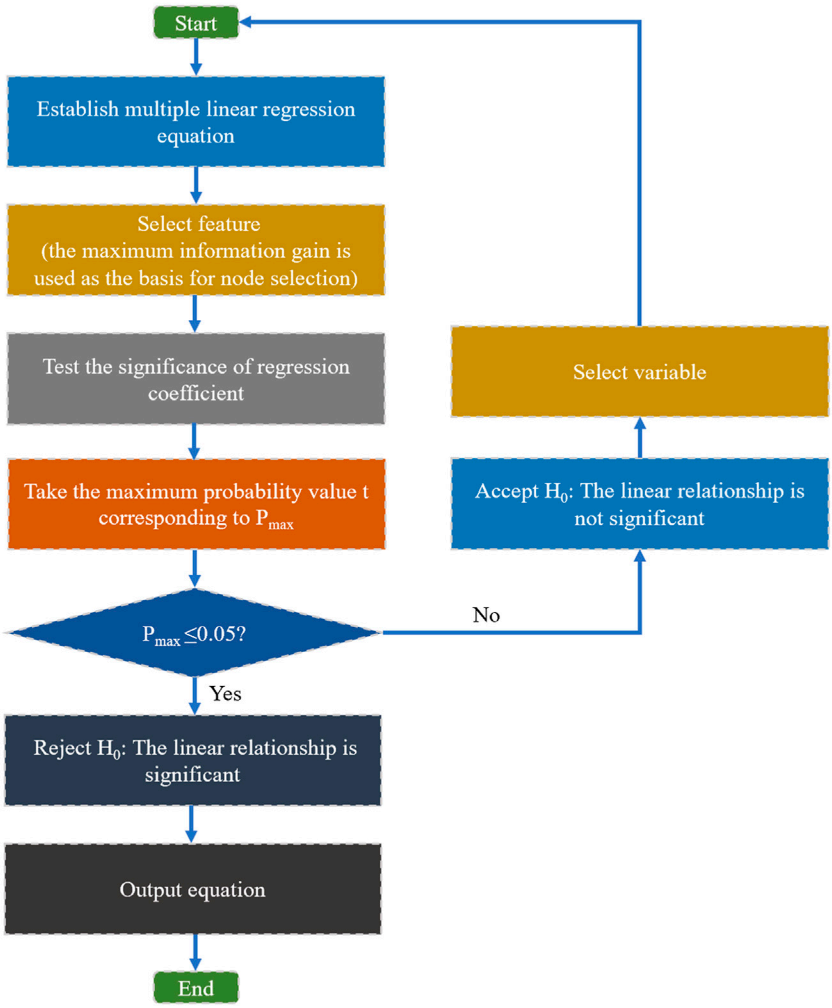

FIGURE 6
The flow chart of MLR.

TABLE 3 Hyperparameters of the machine learning models tuned by FA.

Machine learning Models	Hyperparameters Tuned by FA	Range values (or Requirement) of the Hyperparameters
SVM	C_penalty	0.1–10
	kernel	Linear
	tol	1×10^{-4}-1×10^{-2}
DT	criterion	Gini, Entropy
	max_deep	1–100
	min_samples_split	2–10
	min_samples_leaf	1–10
MLR	tol	1×10^{-5}-1×10^{-3}
	C_inverse	0.1–10

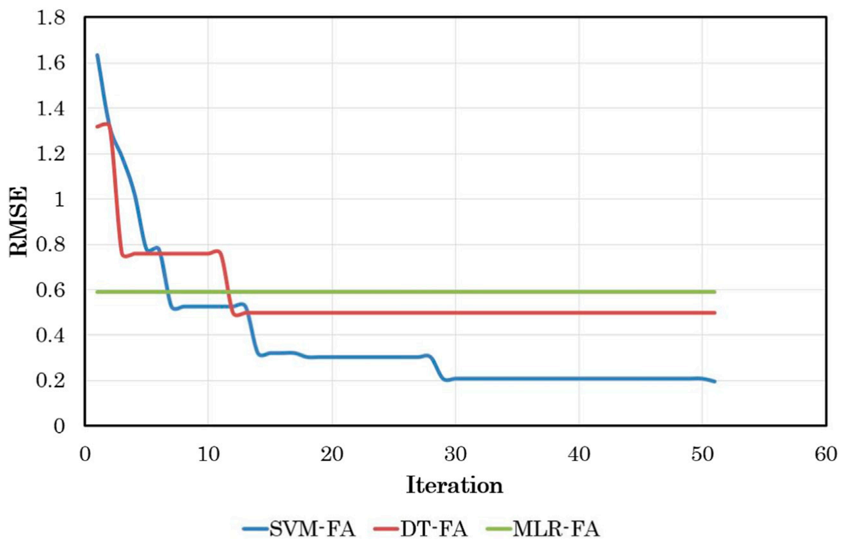

FIGURE 7
The relationship between the number of iterations and RMSE values.

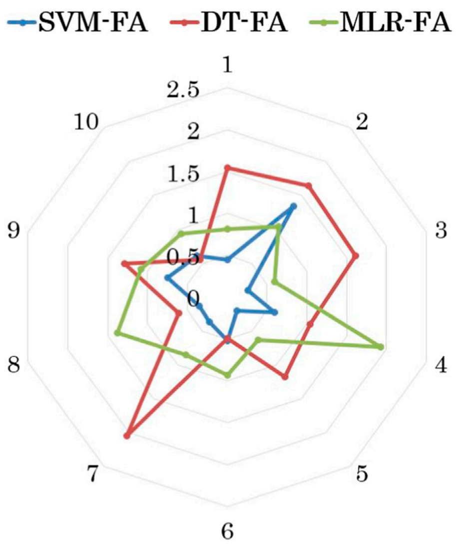

FIGURE 8
RMSE values of different folds.

$$L(\alpha, a, b) = \frac{1}{2}(\alpha \cdot \alpha) - \sum_{i=1}^{n} \beta_i \{y_i [\alpha \cdot D_i + a] - 1\} \tag{3}$$

In the formula, $y_i = 1, \beta_i > 0$ is Lagrange multiplier, solve partial differential equations of α and a , and set them to 0. The original problem is transformed into a duality problem, with constraints:

$$\sum_{i=1}^{n} y_i \beta_i = 0, \quad \beta_i > 0 \tag{4}$$

Solve for β_i that maximizes the value of the following function:

$$Q(\beta) = \sum_{i=1}^{n} \beta_i - \frac{1}{2} \sum_{i,j=1}^{n} \beta_i \beta_j y_i y_j \left(D_i D_j\right) \tag{5}$$

If β_i^* is the optimal solution of the above formula, then:

$$\alpha^* = \sum \beta_i^* \times y_i D \tag{6}$$

The weight coefficient vector of the optimal classification surface is obtained. Calculate the following classification functions:

$$f(D) = sign\{(\alpha^* \cdot D) + a^*\} = sign\left\{\sum \beta_i{}^* y_i (D_i \cdot D) + a^*\right\} \tag{7}$$

If $(D) = 1$, D belongs to this class, otherwise, D does not belong to this class.

The schematic diagram of SVM is shown as Figure 4.

2.3.3 Decision tree (DT)

DT is a common classification method consisting of a set of branching decision test groups (Bonsignori et al., 2021; Azad and Moshkov, 2022). When DT is used to predict the category of the new sample, the corresponding attributes are tested starting from the root node, and the prediction results are obtained when the leaf node is reached. The process of learning of DT is usually a recursive equation (Keikes et al., 2021). In each step, a partition attribute is selected for the formulated data set, and the subset of the data set is determined according to the partition of the attribute, each subset is regarded as the next data set, and the partition of the attribute is selected until the end condition bit is satisfied (Wei et al., 2021; Panhalkar and Doye, 2022). The flow chart of DT is shown in Figure 5.

2.3.4 Multiple linear regression (MLR)

MLR is a common quantitative analysis method, which is widely used in the field of economics due to its superior performance, and is usually used to explain the relationship between variables affected by

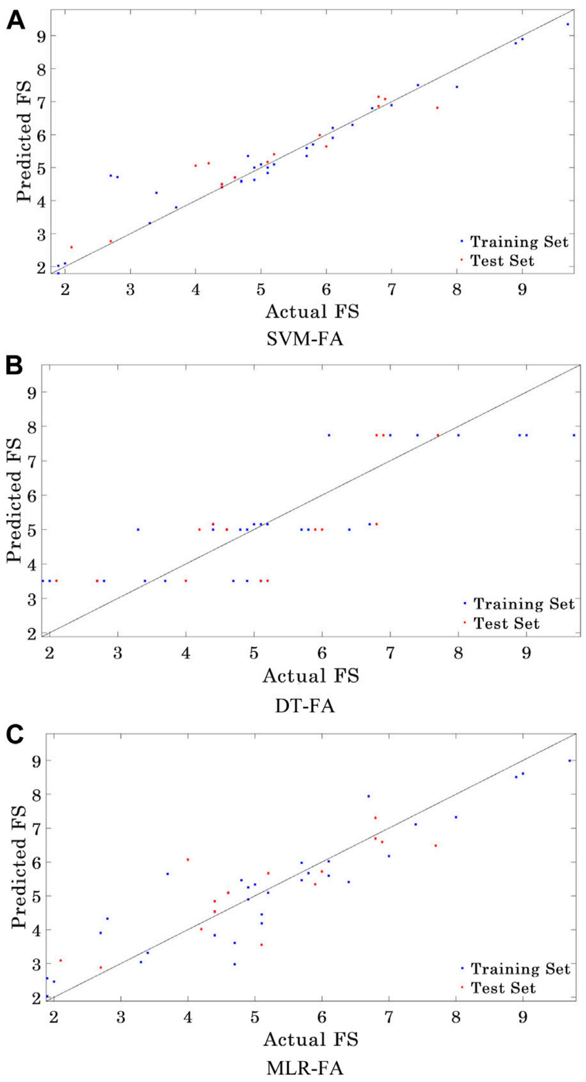

FIGURE 9
The fitting effect of the predicted values and the actual values of the FS of recycled concrete by the machine learning models [**(A)**: SVM-FA, **(B)**: DT-FA, **(C)**: MLR-FA].

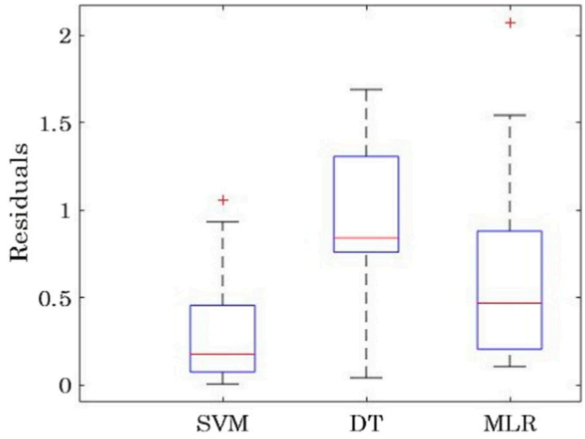

FIGURE 10
Box plot of prediction error of different models.

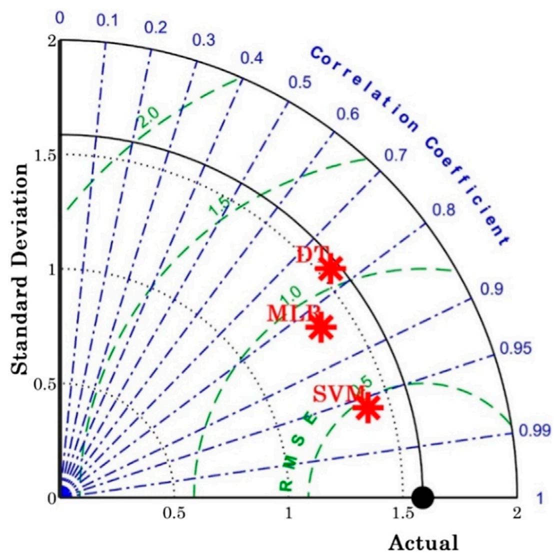

FIGURE 11
Radar plots of R values and RMSE values of different hybrid machine learning models.

multiple variables (Korkmaz, 2021; Rossi Salamanca-Neto et al., 2021; Sibyan et al., 2022), and its general form is:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots \beta_i X_{ii} + \cdots + \beta_k X_{ki} + m_i \quad (8)$$

Where Y is the dependent variable, X are the independent variables, k is the number of explanatory variables, β is the regression coefficient, and m is the random interference term. As can be seen from the above equation, dependent variables are affected by multiple independent variables. During model construction, independent variables with greater influence on dependent variable should be selected and corresponding coefficients of each independent variables should be determined. Then, after judging the fitting effect of the model on actual data, whether the model is suitable for prediction should be further confirmed (Pallardy, 2022; Peng et al., 2022). MLR needs to satisfy the following four assumptions:

- The independent variables are fixed and there is no high correlation.
- The mean and homoscedasticity of m are zero, and no sequential correlation, that is, the expected value of m is zero.
- X is independent of m.
- The random interference term m satisfies normal distribution.

The flow chart of MLR is shown in Figure 6.

Among the four machine learning models, FA is used to tune the hyperparameters of SVM, DT and MLR, while SVM, DT and MLR are used to predict the FS of recycled concrete. The hyperparameters of the three algorithms tuned by FA and the required ranges or values corresponding to the hyperparameters are shown in Table 3.

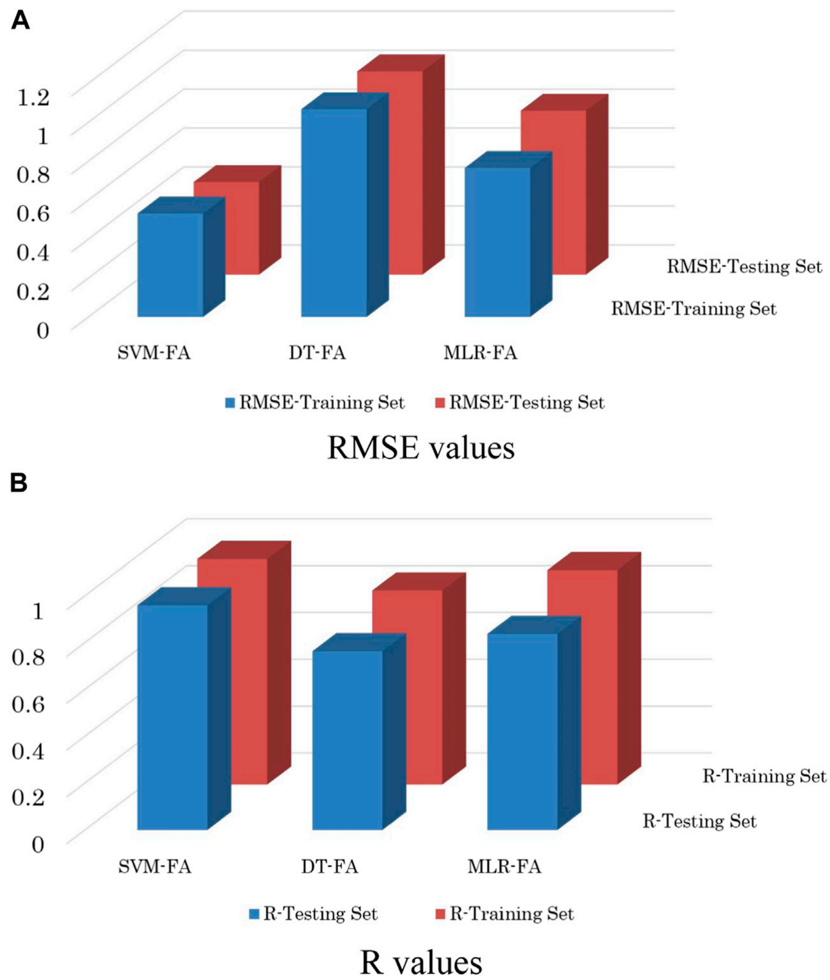

FIGURE 12
The R values and RMSE values of different hybrid machine learning models [**(A)**: R values, **(B)**: RMSE values].

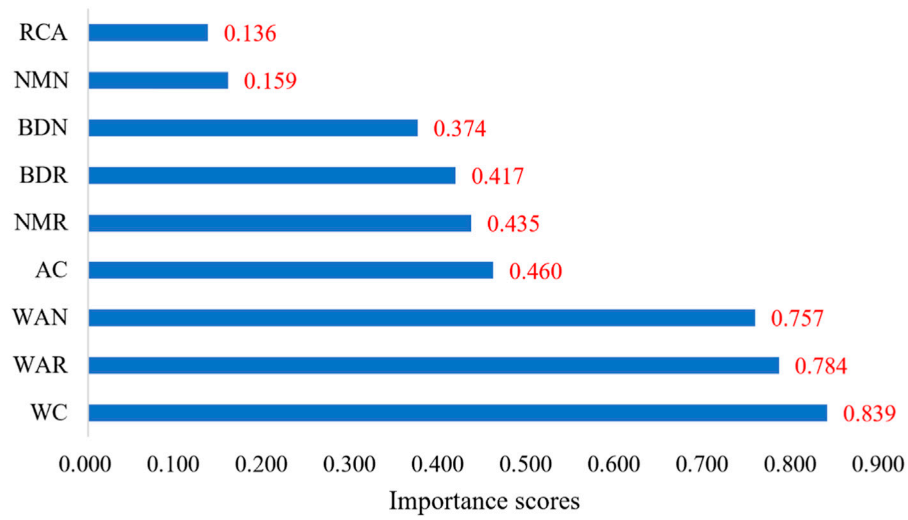

FIGURE 13
The importance scores of the input variable.

3 Result and analysis

3.1 Hyperparameters tuning

In order to improve the accuracy of the model, it is necessary to optimize the hyperparameters of the models before the model training. In this study, FA was used to optimize the hyperparameters of SVM, DT and MLR and the relationship between the number of iterations and RMSE values is shown in Figure 7. It is obviously that with the increase of the number of iterations, the RMSE values of both SVM and DT decrease firstly and then tend to be stable, while the RMSE values of MLR remain basically unchanged. In other words, FA has a higher hyperparameter tuning effect on SVM and DT (Qi et al., 2018), but a poorer hyperparameter tuning effect on MLR.

In order to compare the hyperparameters tuning effects of FA on SVM, DT and MLR, and then select the optimal hyperparameters of the three models, this study selected 10-fold cross-validation method to optimize the hyperparameters of the three models. The relationship between the number of iterations and RMSE values is shown in Figure 8. The figure shows that the RMSE values of SVM is the smallest on the whole, that is, the hyperparameters tuning effect of SVM is the best. The minimum RMSE values of SVM, DT and MLR are obtained at the fourth, sixth and third iterations respectively, that is, the optimal hyperparameters of SVM, DT and MLR are obtained at the fourth, sixth and third iterations respectively.

3.2 Evaluation of the model

Model evaluation is an important step to verify the accuracy of model prediction. The fitting effect of SVM-FA, DT-FA and MLR-FA on the predicted values and actual values of the FS of recycled concrete in the training set and test set is shown in Figure 9. It is clearly that among the SVM-FA, DT-FA and MLR-FA, the SVM-FA has the best fitting effect on the FS of recycled concrete of the predicted values and actual values, proving that SVM-FA is the best model for the prediction on the FS of recycled concrete among the three hybrid machine learning models.

In order to further verify that among the three hybrid machine learning models of SVM-FA, DT-FA and MLR-FA, SVM-FA has the best prediction effect on the FS of recycled concrete, this study further analyzed the prediction effect of the three models. The box diagram of residual results of SVM-FA, DT-FA and MLR-FA is shown in Figure 10 it can be seen from the figure that the residual error of SVM is the smallest among the above three machine learning models, that is, SVM-FA is the model with the best prediction effect.

Figure 11 shows the radar diagram of R values and RMSE values in the testing set of the three hybrid machine learning models. As can be seen in the figure, The RMSE value of SVM-FA is about 0.5 and the R value is about 0.95. On the whole, SVM-FA has the higher R value and the lower RMSE value, which proves that SVM-FA has the best prediction effect, among the three hybrid machine learning models.

The analysis of R values and RMSE values of SVM-FA, DT-FA and MLR-FA is shown in Figure 12. As shown in the figure the SVM-FA has the lowest RMSE value and the highest R value in both the training set and the test set, which again verifies that the SVM in the training set and the test set has the best prediction effect on the FS of recycled concrete among the above three models.

3.3 Importance of variables

Sensitivity analysis is helpful to judge the influence of input variables on the FS of recycled concrete. According to the evaluation of effect of models, SVM-FA is the model with the best prediction effect on the FS of recycled concrete among the three intelligent optimized models in this study. Therefore, the importance score of input variables were extracted from SVM-FA, shown as Figure 13. It is obvious that WC and WAR are the most sensitive variables of the FS of recycled concrete (Xu et al., 2020; Yuan et al., 2022), and the importance score are 0.839 and 0.784 successively, while the importance score of RCA is 0.136 ranking the lowest. Although RCA has the lowest score for the importance of the FS of recycled concrete, which does not mean that RCA can be excluded from the influencing factors when researchers study the FS of recycled concrete and civil engineers prepare the recycled concrete with high FS. It only means that according to the database of this study (Li et al., 2022), RCA has the lowest score for the importance of the FS of recycled concrete.

4 Conclusion

Three intelligent optimization models (SVM-FA, DT-FA, MLR-FA) were proposed to predict the FS of recycled concrete. In order to compose a reliable database for the training and validation of the models, 51 datasets were collected, with nine variables (WC, AC, RCA, NMR, NMN, BDR, BDN, WAR, and WAN) as input variables. Before the training of models, the distribution of input data sets and the correlation of input variables were analyzed, and the results showed that the distribution range of data sets of input variables were wide, and there was no high correlation between different input variables, which proved that using the database to evaluate the prediction accuracy of the models, used to predict the FS of recycled concrete was reasonable.

In this study, the prediction accuracy of the models for the FS of recycled concrete was verified by analyzing the R values, RMSE values and the fitting effect of the predicted values and the actual values. The results showed that compared with DT-FA and MLR-FA, the predicted values and the actual values of SVM-FA have better fitting effect, and have higher R value and lower RMSE value, proving that SVM-FA is the most suitable model for predicting the FS of recycled concrete among the three intelligent optimization models. The sensitivity of each input variable was analyzed, and the ranking of importance of the influenced variables was WC (0.839), WAR (0.784), WAN (0.757), AC (0.460), NMR (0.435), the BDR (0.417), BDN (0.374), NMN (0.159), and RCA (0.136). The above ranking of importance is based on the data set of this study, so the lowest ranking of importance of RCA does not mean that civil engineers can ignore RCA when designing the recycled concrete with high FS.

The research results formed in this study can be used in practical engineering. Engineers can use the model (SVM-FA) obtained in this study to reliably predict the FS of recycled concrete, which can be used as the basis for road design, but the developed model is only suitable for the assessment of FS of recycled concrete. Hence, in future research, a multi-objective optimization design should be proposed, considering the objectives of mechanical properties, economical and carbon emissions.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

The authors confirm contribution to the paper as follows: Study conception and design: QW and MZ; Data collection: MZ; analysis and interpretation of results: QW and MZ; Draft manuscript preparation: QW and MZ. All authors reviewed the results and approved the final version of the manuscript.

Funding

This research was supported by the Fundamental Research Funds for the Central Universities (Grant No. 2021QN1006). The writers are grateful for this support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ahmad, M., Ahmad, F., Huang, J., Iqbal, M. J., Safdar, M., and Pirhadi, N. (2021). Probabilistic evaluation of CPT-based seismic soil liquefaction potential: Towards the integration of interpretive structural modeling and bayesian belief network. *Math. Biosci. Eng.* 18 (6), 9233–9252. doi:10.3934/mbe.2021454

Albuquerque, A., Pacheco, J., and de Brito, J. (2022). Reliability-based recommendations for EN1992 carbonation cover design of concrete with coarse recycled concrete aggregates. *Struct. Concr.* 23 (3), 1873–1889. doi:10.1002/suco.202100406

Arora, S., and Kaur, R. (2022). An escalated convergent firefly algorithm. *J. King Saud University-Computer Inf. Sci.* 34 (2), 308–315. doi:10.1016/j.jksuci.2018.10.007

Arumugam, S. K., Muhamad, R., and Yahya, K. (2018). "Iop, Implementation of eco-costs per value ratio (EVR) on construction waste management in Shah Alam, Malaysia," in 10th Asia Pacific Structural Engineering and Construction Conference, November 2018, Langkawi, Malaysia.

Azad, M., and Moshkov, M. (2022). A Bi-criteria optimization model for adjusting the decision tree parameters. *Kuwait J. Sci.* 49 (2).

Bonsignori, V., Guidotti, R., and Monreale, A. (2021). Deriving a single interpretable model by merging tree-based classifiers, 24th International Conference on Discovery Science (DS). October 2021, Halifax, NS, Canada, Springer, 347–357.

Chen, Z. P., Zhang, Y. Q., Chen, J. J., and Fan, J. (2018). Sensitivity factors analysis on the compressive strength and flexural strength of recycled aggregate infill wall materials. *Appl. Sciences-Basel* 8 (7), 1090. doi:10.3390/app8071090

Chi, C.-P., and Pei, C.-C. (2018). "Iop, finite element simulation analysis on crack resistance of recycled concrete beams of hybrid fiber with waste polypropylene fiber," in 4th International Conference on Applied Materials and Manufacturing Technology (ICAMMT), May 2018, Nanchang, China, Nanchang Hangkong Univ, Natl Def Key Discipline Lab Light Alloy Proc Sci.

Corches, M. T., and Popa, D. V. (2018). Environmental management of construction and demolition waste in Romania. *J. Environ. Prot. Ecol.* 19 (1), 146–151.

Dong, A., You, W., and Zhang, H. (2018). "Research on construction waste treatment methods based on system dynamics," in 6th international symposium on project management (ISPM), July 2018, Chongqing, China, Chongqing Univ Posts & Telecommunicat, 874–882.

Faruqi, M. H. Z., and Siddiqui, F. Z. (2020). A mini review of construction and demolition waste management in India. *Waste Manag. Res.* 38 (7), 708–716. doi:10.1177/0734242x20916828

Guan, Q. Y., Yang, M. Y., Shi, K., and Zhang, T. (2022). Experimental study and finite element analysis on the flexural behavior of steel fiber reinforced recycled aggregate concrete beams. *Materials* 15 (22), 8210. doi:10.3390/ma15228210

Guo, Y.-H., and Wang, Q. (2017). Research on network workload resource prediction based on hybrid model. 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), November 2017, Beijing, China, 891–894.

Gupta, D., Borah, P., Sharma, U. M., and Prasad, M. (2022). Data-driven mechanism based on fuzzy Lagrangian twin parametric-margin support vector machine for biomedical data analysis. *Neural Comput. Appl.* 34 (14), 11335–11345. doi:10.1007/s00521-021-05866-2

Hasanipanah, M., Jamei, M., Mohammed, A. S., Amar, M. N., Hocine, O., and Khedher, K. M. (2022). Intelligent prediction of rock mass deformation modulus through three optimized cascaded forward neural network models. *Earth Sci. Inf.* 15 (3), 1659–1669. doi:10.1007/s12145-022-00823-6

Hasanipanah, M., Meng, D., Keshtegar, B., Nguyen-Thoi, T., and Duc-Kien, T. (2021). Nonlinear models based on enhanced Kriging interpolation for prediction of rock joint shear strength. *Neural Comput. Appl.* 33 (9), 4205–4215. doi:10.1007/s00521-020-05252-4

Hassan, S. H., Aziz, H. A., Daud, N. M., Keria, R., Noor, S. M., Johari, I., et al. (2018). "The methods of waste quantification in the construction sites (A review)," in International Conference on Advances in Civil Engineering and Science Technology (ICACEST), October 2018, Penang, Malaysia.

Hoai-Bao, L., and Quoc-Bao, B. (2021). "Predicting the compressive strength of geopolymer concrete: An empirical model for both recycled and natural aggregates," in 6th International Conference on Geotechnics, Civil Engineering and Structures (CIGOS), October 2021, Hanoi, Vietnam, 793–802.

Huang, J., Koopialipoor, M., and Armaghani, D. J. (2020). A combination of fuzzy Delphi method and hybrid ANN-based systems to forecast ground vibration resulting from blasting. *Sci. Rep.* 10 (1), 19397. doi:10.1038/s41598-020-76569-2

Huang, J., Zhang, J., and Gao, Y. (2021). Intelligently predict the rock joint shear strength using the support vector regression and firefly algorithm. *Lithosphere* 2021. doi:10.2113/2021/2467126

Jamei, M., Hasanipanah, M., Karbasi, M., Ahmadianfar, I., and Taherifar, S. (2021). Prediction of flyrock induced by mine blasting using a novel kernel-based extreme learning machine. *J. Rock Mech. Geotechnical Eng.* 13 (6), 1438–1451. doi:10.1016/j.jrmge.2021.07.007

Jiang, B., Qin, Q., Wang, Q., Li, S., and Yu, H.-c. (2019). Study on mechanical properties and influencing factors of confined concrete arch in underground engineering with complex conditions. *Arabian J. Geosciences* 12, 662. doi:10.1007/s12517-019-4786-1

Kachouh, N., El-Hassan, H., and El-Maaddawy, T. (2021). Influence of steel fibers on the flexural performance of concrete incorporating recycled concrete aggregates and dune sand. *J. Sustain. Cement-Based Mater.* 10 (3), 165–192. doi:10.1080/21650373.2020.1809546

Keikes, L., Kos, M., Verbeek, X. A. A. M., Van Vegchel, T., Nagtegaal, I. D., Lahaye, M. J., et al. (2021). Conversion of a colorectal cancer guideline into clinical decision trees with assessment of validity. *Int. J. Qual. Health Care* 33 (2), mzab051. doi:10.1093/intqhc/mzab051

Kim, J. (2021). Construction and demolition waste management in korea: Recycled aggregate and its application. *Clean Technol. Environ. Policy* 23 (8), 2223–2234. doi:10.1007/s10098-021-02177-x

Korkmaz, M. (2021). A study over the general formula of regression sum of squares in multiple linear regression. *Numer. Methods Partial Differ. Equations* 37 (1), 406–421. doi:10.1002/num.22533

Li, C., Zhou, J., Tao, M., Du, K., Wang, S., Armaghani, D. J., et al. (2022). Developing hybrid ELM-ALO, ELM-LSO and ELM-SOA models for predicting advance rate of TBM. *Transp. Geotech.* 36, 100819. doi:10.1016/j.trgeo.2022.100819

Li, G., Zhang, L., Zhao, F., and Tang, J. (2020). Acoustic emission characteristics and damage mechanisms investigation of basalt fiber concrete with recycled aggregate. *Materials* 13 (18), 4009. doi:10.3390/ma13184009

Li, K., and Lv, Z. (2021). Smooth twin bounded support vector machine with pinball loss. *Appl. Intell.* 51 (8), 5489–5505. doi:10.1007/s10489-020-02085-5

Liang, C., Ma, H., Pan, Y., Ma, Z., Duan, Z., and He, Z. (2019). Chloride permeability and the caused steel corrosion in the concrete with carbonated recycled aggregate. *Constr. Build. Mater.* 218, 506–518. doi:10.1016/j.conbuildmat.2019.05.136

Lin, W. S., Long, Z. B., and Ling, L. (2020). "Iop, Research on conventional triaxial test of recycled aggregate concrete with PVA fiber," in International Conference on Green Development and Environmental Science and Technology (ICGDE), September 2020, Changsha, China.

Liu, B., Feng, C., and Deng, Z. (2019). Shear behavior of three types of recycled aggregate concrete. *Constr. Build. Mater.* 217, 557–572. doi:10.1016/j.conbuildmat.2019.05.079

Luangcharoenrat, C., Intrachooto, S., Peansupap, V., and Sutthinarakorn, W. (2019). Factors influencing construction waste generation in building construction: Thailand's perspective. *Sustainability* 11 (13), 3638. doi:10.3390/su11133638

Nunes, K. R. A., and Mahler, C. F. (2020). Comparison of construction and demolition waste management between Brazil, European Union and USA. *Waste Manag. Res.* 38 (4), 415–422. doi:10.1177/0734242x20902814

Pacheco, J., de Brito, J., Chastre, C., and Evangelista, L. (2021). Eurocode shear design of coarse recycled aggregate concrete: Reliability analysis and partial factor calibration. *Materials* 14 (15), 4081. doi:10.3390/ma14154081

Pallardy, Q. (2022). The effects of atypical diurnal temperature cycles on regression-based downscaling of daily temperature extrema in the Central United States, *Theor. Appl. Climatol.* 151, 1497–1521. doi:10.1007/s00704-022-04338-4

Panhalkar, A. R., and Doye, D. D. (2022). Optimization of decision trees using modified African buffalo algorithm. *J. King Saud University-Computer Inf. Sci.* 34 (8), 4763–4772. doi:10.1016/j.jksuci.2021.01.011

Pani, L., Francesconi, L., Rombi, J., Stochino, F., and Mistretta, F. (2020). "The role of parent concrete in recycled aggregate concrete," in 20th International Conference on Computational Science and Its Applications (ICCSA), July 2020, Cagliari, Italy, 368–378.

Peng, J., Liu, W., Bretz, F., and Hayter, A. J. (2022). Simultaneous confidence tubes for comparing several multivariate linear regression models. *Biometrical J.* 64 (2), 290–300. doi:10.1002/bimj.202000148

Qi, C., Fourie, A., Ma, G., Tang, X., and Du, X. (2018). Comparative study of hybrid artificial intelligence approaches for predicting hangingwall stability. *J. Comput. Civ. Eng.* 32 (2). doi:10.1061/(asce)cp.1943-5487.0000737

Qiao, J., Zhang, L., Zhao, J., and Zhou, Z. (2018). "Iop, test study on impermeability of recycled concrete," in 4th International Conference on Applied Materials and Manufacturing Technology (ICAMMT), May 2018, Nanchang, China, Nanchang Hangkong Univ, Natl Def Key Discipline Lab Light Alloy Proc Sci. .

Rezaei, K., and Rezaei, H. (2022). An improved firefly algorithm for numerical optimization problems and it's application in constrained optimization. *Eng. Comput.* 38 (4), 3793–3813. doi:10.1007/s00366-021-01412-9

Rossi Salamanca-Neto, C. A., Marcheafave, G. G., Mattos, G. J., Moraes, J. T., Schwarzova-Peckova, K., and Sartori, E. R. (2021). Boron-doped diamond film and multiple linear regression-based calibration applied to the simultaneous electrochemical determination of paracetamol, phenylephrine hydrochloride, and loratadine in fixed-dose combinations. *Microchem. J.* 162, 105831. doi:10.1016/j.microc.2020.105831

Shariati, M., Mafipour, M. S., Ghahremani, B., Azarhomayun, F., Ahmadi, M., Nguyen-Thoi, T., et al. (2020). A novel hybrid extreme learning machine–grey wolf optimizer (ELM-GWO) model to predict compressive strength of concrete with partial replacements for cement. *Eng. Comput.* 38, 757–779. doi:10.1007/s00366-020-01081-0

Shmlls, M., Abed, M., Horvath, T., and Bozsaky, D. (2022). Multicriteria based optimization of second generation recycled aggregate concrete. *Case Stud. Constr. Mater.* 17, e01447. doi:10.1016/j.cscm.2022.e01447

Sibyan, H., Svajlenka, J., Hermawan, H., Faqih, N., and Arrizqi, A. N. (2022). Thermal comfort prediction accuracy with machine learning between regression analysis and naive bayes classifier. *Sustainability* 14 (23), 15663. doi:10.3390/su142315663

Vilventhan, A., Ram, V. G., and Sugumaran, S. (2019). Value stream mapping for identification and assessment of material waste in construction: A case study. *Waste Manag. Res.* 37 (8), 815–825. doi:10.1177/0734242x19855429

Wang, P. F., Liu, Y., Gao, S. Y., Luo, S., Jia, J., Yang, H., et al. (2020). "Iop, prediction of construction waste in representative cities of north and south based on grey model," in International Conference on Green Development and Environmental Science and Technology (ICGDE), September 2020, Changsha, China.

Wang, Q., Chen, L., Hu, R., Ren, Z., He, Y., Liu, D., et al. (2020). An empirical study on waste generation rates at different stages of construction projects in China. *Waste Manag. Res.* 38 (4), 433–443. doi:10.1177/0734242x19886635

Wang, X., Zhang, H., Wang, Y., and Yang, S. (2019). ELM-based AFL-SLFN modeling and multiscale model-modification strategy for online prediction. *Processes* 7 (12), 893. doi:10.3390/pr7120893

Wang, Y., Lin, Z., Tang, C., and Hao, W. (2021). Influencing factors on the healing performance of microcapsule self-healing concrete. *Materials* 14, 4139. doi:10.3390/ma14154139

Weglorz, M., Ajdukiewicz, A., and Kliszczewicz, A. (2018). "Assessment of recycled concrete aggregate properties required for structural concretes," in 64th Scientific Conference of the Committee-for-Civil-Engineering-of-the-Polish-Academy-of-Sciences and the Science-Committee-of-the-Polish-Association-of-Civil-Engineers (PZITB) (KRYNICA), January 2018, Krynica Zdroj, Poland.

Wei, W., Hui, M., Zhang, B., Scherer, R., and Damasevicius, R. (2021). Research on decision tree based on rough set. *J. Internet Technol.* 22 (6), 1385–1394. doi:10.53106/160792642021112206015

Xiao, J. Z., Fan, Y. H., and Tawana, M. M. (2013). Residual compressive and flexural strength of a recycled aggregate concrete following elevated temperatures. *Struct. Concr.* 14 (2), 168–175. doi:10.1002/suco.201200037

Xiao, J., Zhang, H., and de Brito, J. (2022). M&S highlight: Limbachiya, et al. (2000), Use of recycled aggregate in high-strength concrete. *Mater. Struct.* 55 (2), 40. doi:10.1617/s11527-021-01871-y

Xie, R., Ge, R., Li, Z., Qu, G., Zhang, Y., Xu, Y., et al. (2022). Synthesis and influencing factors of high-performance concrete based on copper tailings for efficient solidification of heavy metals. *J. Environ. Manag.* 325, 116469. doi:10.1016/j.jenvman.2022.116469

Xiong, E., Li, T., Wang, S., and Fan, Y. (2018). Influence of vibration on performance of recycled concrete. *Civ. Eng. Journal-Stavebni Obz.* 27 (4), 587–598. doi:10.14311/cej.2018.04.0047

Xu, J., Wang, Y., Ren, R., Wu, Z., and Ozbakkaloglu, T. (2020). Performance evaluation of recycled aggregate concrete-filled steel tubes under different loading conditions: Database analysis and modelling. *J. Build. Eng.* 30, 101308. doi:10.1016/j.jobe.2020.101308

Yin, C., (2021). "Ieee Comp, Experimental study on mix proportion of recycled concrete with super plasticizer," in 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), June 2021, Beihai, China, 488–491.

Yuan, X. Z., Tian, Y. Z., Ahmad, W., Ahmad, A., Usanova, K. I., Mohamed, A. M., et al. (2022). Machine learning prediction models to evaluate the strength of recycled aggregate concrete. *Materials* 15 (8), 2823. doi:10.3390/ma15082823

Zhang, W., Liu, X.-h., Huang, Y., and Tong, M.-N. (2022). Reliability-based analysis of the flexural strength of concrete beams reinforced with hybrid BFRP and steel rebars. *Archives Civ. Mech. Eng.* 22, 171. doi:10.1007/s43452-022-00493-7

Zhu, W., Rad, H. N., and Hasanipanah, M. (2021). A chaos recurrent ANFIS optimized by PSO to predict ground vibration generated in rock blasting. *Appl. Soft Comput.* 108, 107434. doi:10.1016/j.asoc.2021.107434

10.3389/fmats.2023.1159079

Appendix

TABLE A1 Abbreviated term.

FS	flexural strength
SVM	support vector machine
DT	decision tree
FA	firefly algorithm
MLR	multiple linear regression
SVM-FA	firefly algorithm and support vector machine hybrid machine learning model
DT-FA	firefly algorithm and decision tree hybrid machine learning model
MLR-FA	firefly algorithm and multiple linear regression hybrid machine learning model
WC	effective water-cement ratio
AC	aggregate-cement ratio
RCA	recycled concrete aggregate replacement ration
NMR	nominal maximum recycled concrete aggregate size
NMN	nominal maximum normal aggregate size
BDR	bulk density of recycled concrete aggregate
BDN	bulk density of normal aggregate
WAR	water absorption of recycled concrete aggregate
WAN	water absorption of normal aggregate
RMSE	root mean square error
R	correlation coefficient
Std.	standard deviation