

Article

Predicting Mechanical Properties of High-Performance Fiber-Reinforced Cementitious Composites by Integrating Micromechanics and Machine Learning

Pengwei Guo ¹, Weina Meng ¹, Mingfeng Xu ², Victor C. Li ³ and Yi Bao ^{1,*} 

¹ Department of Civil, Environmental and Ocean Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA; pguo@stevens.edu (P.G.); weina.meng@stevens.edu (W.M.)

² School of Civil and Transportation Engineering, Hebei University of Technology, Tianjin 300401, China; xumingfeng@hebut.edu.cn

³ Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109, USA; vcli@umich.edu

* Correspondence: yi.bao@stevens.edu

Abstract: Current development of high-performance fiber-reinforced cementitious composites (HPFRCC) mainly relies on intensive experiments. The main purpose of this study is to develop a machine learning method for effective and efficient discovery and development of HPFRCC. Specifically, this research develops machine learning models to predict the mechanical properties of HPFRCC through innovative incorporation of micromechanics, aiming to increase the prediction accuracy and generalization performance by enriching and improving the datasets through data cleaning, principal component analysis (PCA), and K-fold cross-validation. This study considers a total of 14 different mix design variables and predicts the ductility of HPFRCC for the first time, in addition to the compressive and tensile strengths. Different types of machine learning methods are investigated and compared, including artificial neural network (ANN), support vector regression (SVR), classification and regression tree (CART), and extreme gradient boosting tree (XGBoost). The results show that the developed machine learning models can reasonably predict the concerned mechanical properties and can be applied to perform parametric studies for the effects of different mix design variables on the mechanical properties. This study is expected to greatly promote efficient discovery and development of HPFRCC.

Keywords: ductility; high-performance fiber-reinforced cementitious composites (HPFRCC); machine learning; mechanical properties; micromechanics model



Citation: Guo, P.; Meng, W.; Xu, M.; Li, V.C.; Bao, Y. Predicting Mechanical Properties of High-Performance Fiber-Reinforced Cementitious Composites by Integrating Micromechanics and Machine Learning. *Materials* **2021**, *14*, 3143. <https://doi.org/10.3390/ma14123143>

Academic Editor: Hyeong-Ki Kim

Received: 28 April 2021

Accepted: 4 June 2021

Published: 8 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

High-performance fiber-reinforced cementitious composites (HPFRCC) feature high tensile strength and ductility, strain-hardening property, and long-term durability [1]. Representative HPFRCC include engineered cementitious composites (ECC) [2–4] and ultra-high-performance concretes (UHPC) [5–8]. ECC feature high ductility, dense cracks, and self-control of crack width, and is designed by mechanistically tuning the matrix, fibers, and fiber–matrix interface [1]. Recently, ECC has achieved multifunctionality, such as self-healing, self-sensing, self-cleaning, and air-purifying [9,10]. With self-healing crack width control, ECC possesses extreme durability [11]. Typically, with the use of a medium volume (~2%) of polymer fibers, ECC can achieve a tensile strain capacity of 4% or higher [12,13]. UHPC feature high compressive and tensile strengths and are designed by maximizing the particles packing density. Under standard curing conditions, UHPC can achieve compressive strengths higher than 120 MPa. Uncracked UHPC has excellent durability due to the dense microstructure. The superior properties are based on proper mix design. For example, UHPC is designed to densify the microstructures through maximizing the particle

packing and chemistry [5], and ECC is designed to mechanistically tune the matrix, fibers, and fiber–matrix interface [1]. Due to their superior mechanical properties, ECC and UHPC have been used to improve the load capacity and resilience of various civil engineering structures such as bridges and buildings under earthquake [14,15], fatigue [16,17], and fire [18,19]. The reported applications have shown that HPFRCC may significantly improve the resilience and sustainability of structures.

Current development of HPFRCC mainly relies on intensive experiments. A micromechanics model was developed to describe the mechanisms of the unique tensile properties and crack resistance, and generate effective strategies for improving the post-cracking behavior, thus, greatly promoting the development of HPFRCC [20–22]. However, intense experiments are still needed to determine multiple essential parameters in the micromechanics model. For example, single-fiber pullout tests are needed to characterize the fiber–matrix interfacial properties [4]. Since the experiments of cementitious materials typically take long time, more efficient methods are desired. Based on the micromechanics model, semi-empirical models have been presented to link the tensile properties and mix design variables [23]. However, the semi-empirical models consider only a limited number of variables (e.g., fiber length, diameter, and volume).

Recently, machine learning methods have been applied to predict material properties, which can reduce time and cost for discovering new materials [24,25]. Compared with the conventional regression-based data-driven methods [26], machine learning methods are capable of dealing with complicated datasets with various input and output variables [27] while achieving desired accuracy [28]. Machine learning has been applied to predict the compressive strength [29–32] and the modulus of elasticity [33–36] of concrete. For example, an artificial neural network (ANN) was used to predict the compressive strength by using the water-to-cement ratio, the fly ash content, and the aggregate content [30]. Support vector regressor (SVR) and classification and regression tree (CART) models were used to predict the compressive strength and the modulus of elasticity of concrete [31–34]. Recently, ANN was used to predict the compressive strength and the tensile strength of ECC [37], and the extreme gradient boosting (XGBoost) algorithm was used to predict electrical resistivity of concrete [38].

Despite the above advances in using machine learning for predicting concrete properties, the following challenges have been identified: (1) A large amount of data are required to achieve an acceptable prediction accuracy, but there is insufficient data in most cases, particularly in the course of developing new materials. (2) There is no machine learning model for predicting ductility, which is a critical property of HPFRCC. (3) There is a lack of knowledge on how to select appropriate machine learning models and the variables for HPFRCC. (4) It is unclear how to improve the quality of the dataset for developing the machine learning models. These challenges have hindered wider acceptance and applications of machine learning methods.

This paper proposes to predict the mechanical properties of HPFRCC by integrating micromechanics and machine learning, aiming to achieve high prediction accuracy while limiting the dataset size. The ductility (i.e., strain capacity) of HPFRCC is predicted for the first time, in addition to the compressive and tensile strengths. The main objectives and contributions of this research are to: (1) develop a novel method to incorporate the micromechanics model for automated prediction of the mechanical properties of HPFRCC with a high prediction accuracy; (2) enable the prediction of ductility (i.e., tensile strain capacity) with a reasonable prediction accuracy; (3) develop innovative methods to achieve a high prediction accuracy and the generalization performance through improving the dataset; and (4) compare the performance of different machine learning models for prediction of HPFRCC. To this end, this study investigates four different machine learning methods: the ANN, SVR, and CART, XGBoost, which are used to develop machine learning models to predict the tensile strength, the tensile strain capacity, and the compressive strength, respectively. Two strategies are presented to utilize micromechanics, and multiple

innovative methods are proposed to improve the dataset. This study attempts to provide an alternative method to promote the development of HPRCC.

2. Methodology

2.1. Machine Learning Models

This section introduces the ANN, SVR, and CART. ANN links the input variables (e.g., mix design) to the output variables (e.g., mechanical properties). Figure 1a shows a typical ANN consisting of three types of layers, including an input layer, one or multiple hidden layers, and an output layer [39]. Each layer has one or multiple variables, and the relationships of the variables in different layers are described by using weights and bias. Given a dataset with known mix designs and the corresponding mechanical properties, the weights and the bias are determined to minimize the discrepancy between the predicted and real mechanical properties through an optimization process [40], which is known as the training process. Once the ANN is trained, the relationships between the layers are determined, so the ANN can be used to predict the mechanical properties using the mix design variables. SVR links the input variables (e.g., mix design) to the output variable (e.g., compressive strength) using a regression relationship [41].

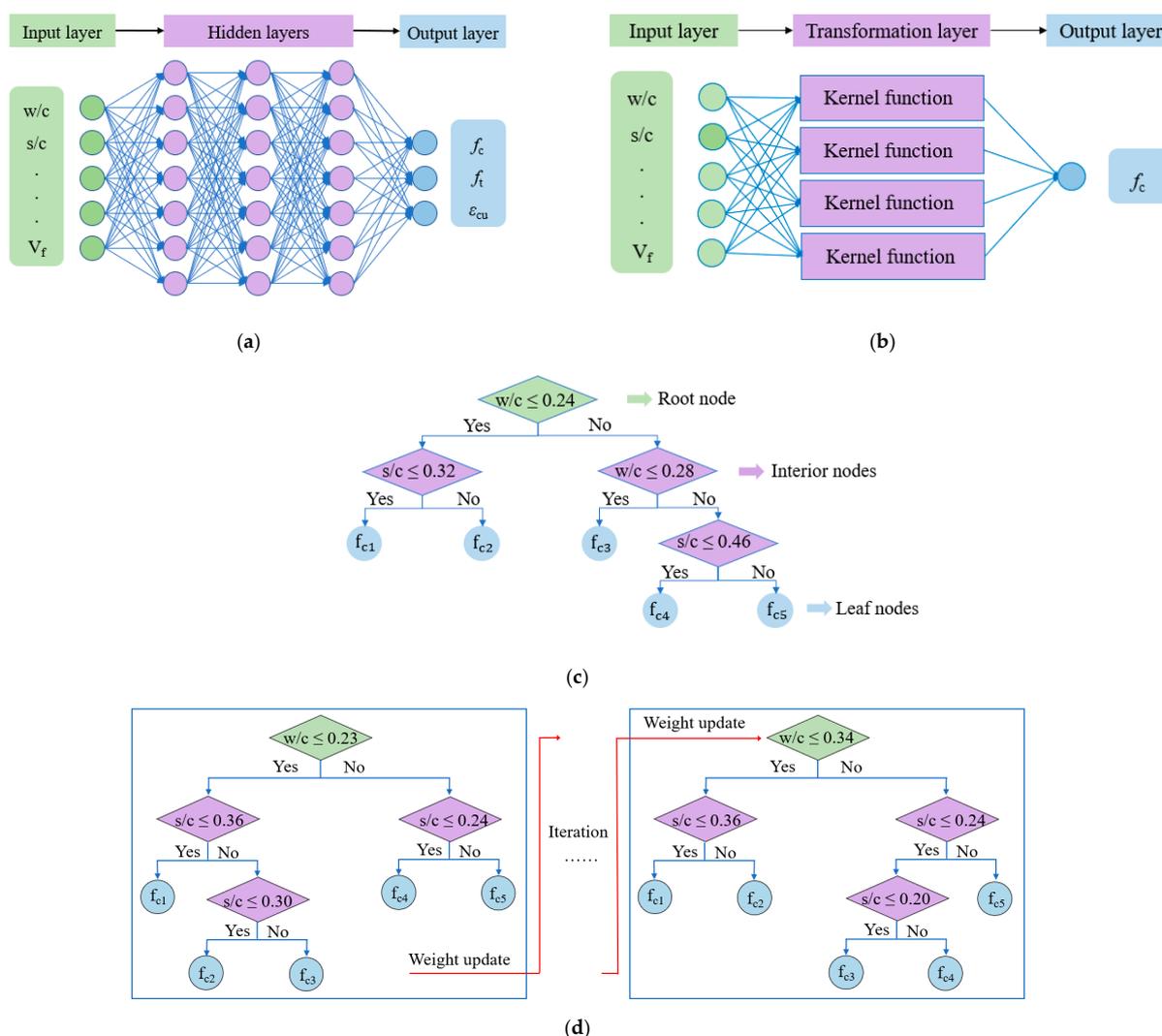


Figure 1. Machine learning models: (a) ANN; (b) SVR; (c) CART; and (d) XGBoost. w/c is the water-to-cement ratio; s/c is the sand-to-cement ratio; V_f is the fiber content; f_c is the compressive strength; f_t is the tensile strength; and ε_{cu} is the tensile strain capacity (i.e., ductility).

Figure 1b illustrates an application of using SVR to predict the compressive strength. Compared with ANN, SVR also contains three types of layers [42], and uses weights and bias to relate the layers. SVR has one output variable at a time. To consider multiple mechanical properties, SVR can be operated for multiple times. Compared with the conventional regression methods, SVR employs kernel functions that enable the model to solve complex, non-linear problems because the relationships between some variables cannot be described using linear functions [43]. CART relates the input variables to the mechanical properties using a tree structure [40], as shown in Figure 1c. The tree is composed of a root node, multiple interior nodes, and multiple leaf nodes [44]. CART describes the relationships between the input and the output variables by splitting the values of the input variables into subgroups, and determines the splitting pathway through a training process by using a given dataset. The splitting operation of the tree is terminated when the termination criterion is met. The splitting schemes are determined through the training process of the CART model [45]. Similar to SVR, CART has one output variable at a time. Typically, a single tree model (e.g., CART model) cannot provide accurate predictions due to the relatively simple architecture and limited prediction capability. Therefore, the extreme gradient boosting trees were presented to ensemble multiple tree models. The XGboost method can continuously add a new tree and fit the discrepancy between the real value and predicted value from the last iteration, as shown in Figure 1d.

2.2. Dataset

The development of the machine learning models is based on datasets that are needed to relate the input and the output variables of the models. The size and the quality of the dataset are significant for the accuracy and generalization performance of the machine learning models.

2.2.1. Overview

Figure 2 shows the proposed flowchart for establishing the datasets used to develop the machine learning models. First, the variables informed by the mix design and the micromechanics model of HPFRCC in published references are preliminary selected to form a dataset, designated as Dataset 1. Considering that there are limited data of HPFRCC in the published references and the test results of the tensile strain capacity usually show significant scatters, the micromechanics model is used to generate more results of the tensile strain capacity for data augmentation, forming another dataset, designated as Dataset 2. Then, data cleaning is performed to identify and remove anomalous data in Dataset 1 and Dataset 2. The cleaned datasets are further processed through data normalization. The normalized datasets are tested to check whether multicollinearity occurs. If multicollinearity occurs, a principal component analysis (PCA) will be performed to reduce the dimensionality of the datasets and eliminate the multicollinearity problem. The novelties of the procedures include: (i) utilization of micromechanics model for variable selection and data augmentation; (ii) data cleaning and normalization; and (iii) adoption of PCA.

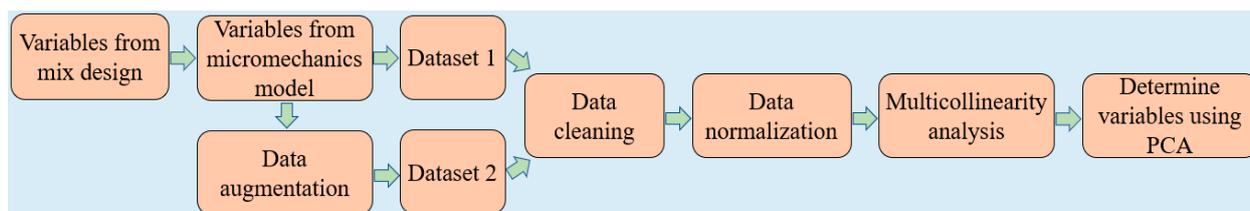


Figure 2. Flowchart of the proposed method for the selection and improvement of the variables.

Currently, there is no consensus on the selection of variables for predicting material properties using machine learning methods. Different scholars selected different variables to predict the same type of properties. For example, in reference [46], the compressive

strength was predicted by using the w/c , the aggregate-to-cement ratio, the fine aggregate content, and the superplasticizer content as the input variables, while in reference [25] the compressive strength was predicted by using the w/c , the fly ash content, the aggregate-to-cement ratio, the micro silica content, and the superplasticizer content.

A micromechanics model [47] was developed to design HPFRCC in order to achieve the desired tensile properties, in particular, the post-cracking strain-hardening properties and the superior ductility and toughness. The micromechanics model informs two criteria that are essential for achieving strain-hardening behavior: energy criterion and stress criterion. Figure 3 shows the stress-crack curve for strain-hardening cementitious composites (e.g., ECC) [22].

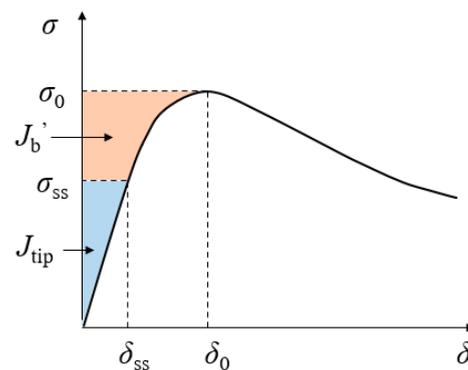


Figure 3. Illustration of the stress-crack opening curve for HPFRCC. The orange area represents the complementary energy (J_b'), and the blue area represents the crack tip toughness (J_{tip}).

The energy criterion for steady-state crack propagation can be expressed in Equation (1) [20]:

$$J_{tip} = \sigma_{ss}\delta_{ss} - \int_0^{\delta_{ss}} \sigma(\delta)d\delta \quad (1)$$

where J_{tip} is the toughness of the matrix, and $J_{tip} = K_m^2/E_m$; E_m is the modulus of elasticity of the matrix; K_m is the fracture toughness of the matrix, which can be tested using beams with a notch under three-point bending [22]; σ_{ss} is the tensile strength under steady-state crack propagation process; and δ_{ss} is the corresponding crack width [48].

The toughness of the matrix must be less than the complementary energy from the fiber bridging [20]. The upper limit for steady-state crack propagation condition can be expressed as:

$$J_{tip} = \frac{K_m^2}{E_m} \leq \sigma_0\delta_0 - \int_0^{\sigma_0} \sigma(\delta)d\delta \equiv J_b' \quad (2)$$

where σ_0 is the peak stress, and δ_0 is the corresponding crack opening width.

The complementary energy can be calculated by Equation (3) [22]:

$$J_b' = V_f \frac{L_f}{d_f} \left(\frac{\tau_0^2 L_f^2}{6d_f E_f} - 2G_d \right) \quad (3)$$

where V_f , L_f , d_f , and E_f are respectively the volume ratio, length, diameter, and elastic modulus of the fibers; τ_0 and G_d are respectively the frictional bond and chemical bond strengths [49].

The micromechanics model shows that the tensile properties of HPFRCC are associated with the following parameters: (1) the properties of the chopped fibers: the volume ratio (V_f), the fiber length (L_f), the fiber diameter (d_f), and the elastic modulus (E_f); (2) the properties of the cementitious matrix: the elastic modulus (E_m) and the fracture toughness (K_m); and (3) the fiber–matrix interface properties: the frictional bond strength (τ_0) and the chemical bond strength (G_d) [12]. Therefore, the fiber properties (V_f , L_f , d_f , E_f) are

also considered as the input variables of the machine learning methods, in addition to the variables typically used for conventional concrete.

Therefore, a total of 14 variables are selected, as listed in Table 1. The variables are categorized into: (1) the mix design variables of HPFRCC: the cement-to-binder ratio, the fly ash-to-binder ratio, the ground-granulated blast slag-to-binder ratio, the limestone powder-to-binder ratio, the rice husk-to-binder ratio, the metakaolin-to-binder ratio, the silica fume-to-cement ratio, the water-to-binder ratio, the sand-to-binder ratio, the superplasticizer content, and the fiber content; and (2) the physical properties of the fibers: the fiber length, the fiber diameter, the elastic modulus. A total of 387 experimental data are collected from published papers [4,13,18,22,23,48,50–66] to form a dataset, designated as Dataset 1. The output variables for Dataset 1 are the compressive strength (f_c), the tensile strength (f_t), and the tensile strain capacity (ϵ_{cu}) at 28 days.

Table 1. Description of input variables.

Number	Variable	Range	Unit	Mean	Standard Deviation
1	Cement-to-binder ratio	0.152–1	1	0.463	0.212
2	Fly ash-to-binder ratio	0–0.848	1	0.362	0.306
3	Slag-to-binder ratio	0–0.808	1	0.12	0.211
4	Rice husk-to-binder ratio	0–0.360	1	0.004	0.028
5	Limestone-to-binder ratio	0–0.577	1	0.022	0.080
6	Metakaolin-to-binder ratio	0–0.094	1	0.001	0.008
7	Silica fume-to-binder ratio	0–0.206	1	0.014	0.035
8	Sand-to-binder ratio	0–1.40	1	0.41	0.19
9	Water-to-binder ratio	0.11–0.80	1	0.27	0.08
10	Superplasticizer content	0–2.7	%	0.78	0.59
11	Fiber volume	0–3.0	%	1.9	0.5
12	Fiber length	6–27	mm	11.5	3.6
13	Fiber diameter	12–39	μm	34.2	8.3
14	Fiber elastic modulus	4–200	GPa	56.1	34.6

2.2.2. Dataset Augmentation

Based on the micromechanics, a semi-empirical model was proposed to predict the tensile strain capacity (ϵ_{cu}) of HPFRCC by using three fiber parameters, as shown in Equation (4) [23]:

$$\epsilon_{cu} = 6.6 \ln \left(\frac{L_f}{d_f} V_f \right) - 10.7 \quad (4)$$

where L_f is the fiber length; d_f is the fiber diameter; and V_f is the fiber content. The R^2 of Equation (4) was 0.95, indicating a strong correlation [23].

Therefore, the semi-empirical model is used to generate more data to enlarge the dataset used to develop the machine learning models. Specifically, Equation (4) is used to generate 70 data by varying the values of L_f , d_f , and V_f . The generated data are used to supplement the data in Dataset 1, forming a larger dataset for the prediction of tensile strain capacity, designated as Dataset 2. Compared with Dataset 1, Dataset 2 has the same types of variables but is larger.

2.2.3. Dataset Cleaning

In general, there are anomalous data in the dataset formed by collecting test data from different sources, due to the errors generated in tests, data documentation, and so on. This study proposes to identify and remove anomalous data from dataset through a cluster analysis. Specifically, the anomalous data are identified from the analysis of data distribution, as elaborated in [67]. For each variable, when the data follows a normal distribution, 99.7% of the entire dataset should be within three times standard deviations

(3σ), as shown in Equation (5). In this study, the data outside the range determined by the normal distribution are considered as anomalous data, as depicted by:

$$P(|x - \mu| > 3\sigma) \leq 0.3\% \quad (5)$$

where x denotes a data; μ is the expectation; and σ is the standard deviation.

2.2.4. Dataset Normalization

The raw data extracted from literature often have different units and scales of magnitude. For example, the water-to-binder ratio is 0.25, while the modulus of elasticity of fibers can be up to 100 GPa. The significant discrepancy of numeric values of different variables may highly affect the results of machine learning models. Therefore, in this study, all the input data are normalized to the range of -1 to 1 , as shown in Equation (6):

$$x^* = \frac{x - \mu}{\sigma} \quad (6)$$

where x is the original data; x^* is the normalized data; μ is the mean value; and σ is the standard deviation. The distribution of data is kept the same before and after the normalization [68]. The dataset is divided into training and testing datasets with the same random seed.

2.2.5. Multicollinearity and Principal Component Analysis

Multicollinearity may occur in high-dimension analysis and compromise the statistical significance of independent variables [69]. According to [70], multicollinearity occurs when the absolute value of the Pearson correlation coefficient is higher than 0.7. When multicollinearity occurs, this study performs a PCA [71], which is an unsupervised learning method to reduce the dimensionality of the dataset and avoid multicollinearity through eigenvalue decomposition. The PCA aims to extract the main variables by evaluating the significance of the variables on the mechanical properties. The significance is reflected by the variance as defined in Equation (7):

$$\lambda = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1} \quad (7)$$

where λ is the variance; X_i is the i th sample; and \bar{X} is the average value of all the samples.

A cumulative variance ratio is defined in Equation (8) [72]:

$$\text{Cumulative variance ratio} = \frac{\sum_{j=1}^k \lambda_j}{\sum_{j=1}^n \lambda_j} \quad (8)$$

where k is the optimal dimensionality of the input variables, and n is the total dimensionality of the input variables. According to [72], the cumulative variance ratio is the ratio of the sum of the variances for the principal components to the total variances for all components, and the cumulative variance ratio should be greater than 0.99.

2.3. Hyperparameter Tuning

Hyperparameters are the key parameters of machine learning methods. For example, the hyperparameters of ANN include the number of variables in each hidden layer and the learning rate. This study proposes to combine a grid search method [73] and K-fold cross-validation method to optimize hyperparameters and prevent overfitting and underfitting. Figure 4 illustrates the proposed hyperparameter tuning or optimization method. For instance, the number of variables in a hidden layer of an ANN is described as $H = \{20, 21, \dots, 100\}$, and the learning rate is expressed as $\eta = \{0.1, 0.01, 0.001\}$. The grid search method tests and selects the H and η values that yield the lowest error. The K-fold cross-validation is used to improve the generalization performance of the machine learning

models. A training dataset is divided into K folds (K = 10) with comparable sizes. One fold is randomly selected as the validation set, and the other folds are used to train the model. By using K-fold cross-validation method, all data can participate in the training process.

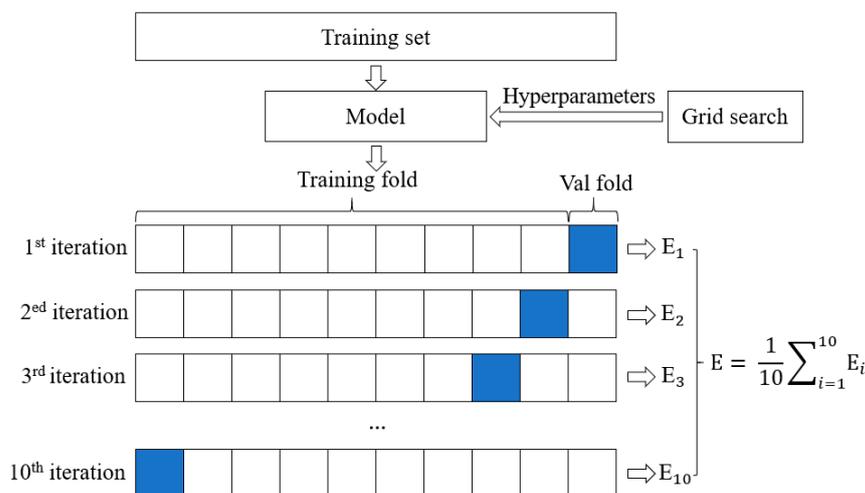


Figure 4. K-fold cross-validation in the training. E_1 to E_{10} are the loss functions during the training process, and E is the cross-validation loss. The grid search method is used to find the optimum E .

2.4. Performance Evaluation

To evaluate the prediction accuracy, three typical performance metrics are used to assess the correlation between the predicted value (Y_{pre}) and the actual value (Y_{actual}) of the four different machine learning models, which are the mean squared error (MSE), Pearson correlation coefficient (R), and coefficient of determination (R^2), as defined in Equations (9)–(11) [74–76]:

$$MSE = \frac{1}{n} \cdot \sum_{i=1}^n (Y_{pre} - Y_{actual})^2 \quad (9)$$

$$R = \frac{\sum_{i=1}^n (Y_{pre} - \overline{Y_{pre}}) \cdot (Y_{actual} - \overline{Y_{actual}})}{\sqrt{\sum_{i=1}^n (Y_{pre} - \overline{Y_{pre}})^2} \cdot \sqrt{\sum_{i=1}^n (Y_{actual} - \overline{Y_{actual}})^2}} \quad (10)$$

$$R^2 = \frac{\sum_{i=1}^n (Y_{pre} - \overline{Y_{pre}}) \cdot (Y_{actual} - \overline{Y_{actual}})}{\sum_{i=1}^n (Y_{actual} - \overline{Y_{actual}})^2} \quad (11)$$

where n is the data number.

2.5. Innovation of the Proposed Methodology

Figure 5 shows the innovations for the prediction of the mechanical properties of HPFRCC. With the challenges identified in the introduction section, novel methods are proposed for improving the dataset used to develop the machine learning models, including data collection, data augmentation, data cleaning, multicollinearity analysis, and variable selection through PCA. Two strategies are proposed to utilize the micromechanics model: (1) Strategy 1 (variable selection): use the theoretical model to screen the variables; and (2) Strategy 2 (data augmentation): use the model to generate more data that supplement the experimental data. These two strategies are elaborated in Section 2.2.1. The PCA method is proposed to finalize the selection of variables and avoid multicollinearity. K-fold cross-validation and grid search are combined to optimize the hyperparameters. Finally, the prediction accuracy is evaluated to select the best machine learning models for the different mechanical properties of HPFRCC.

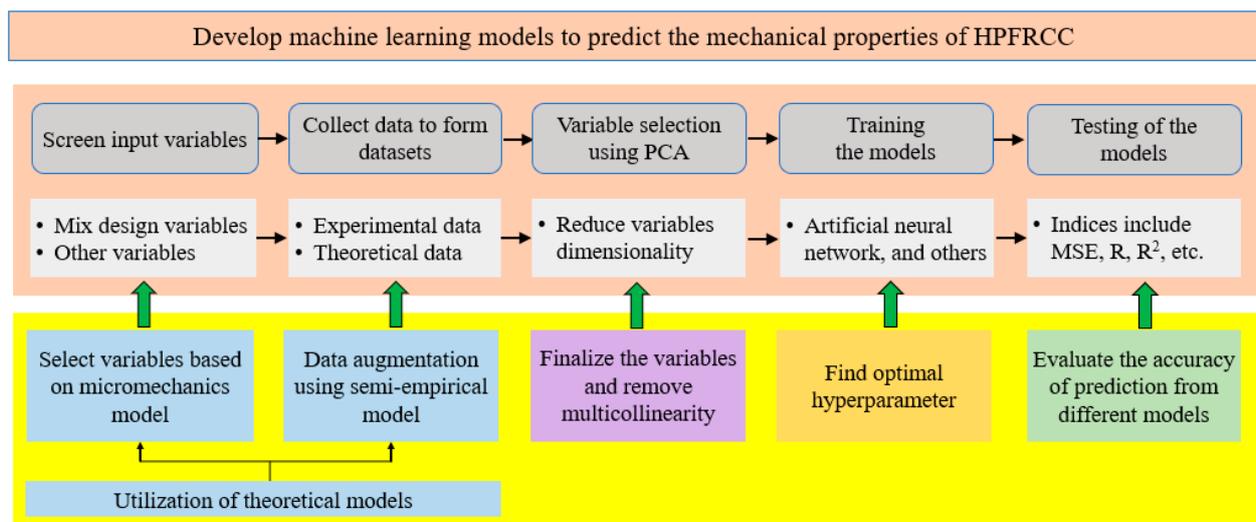


Figure 5. Illustration of the innovation of methodology for prediction of the mechanical properties of HPFRCC by integrating data-driven and model-based methods.

3. Results and Discussions

3.1. Anomalous Data

Table 2 shows the data anomaly detection results. It should be noted that only the items that contained anomalous data are listed. According to the analysis total of 23 data are removed from Dataset 1 and Dataset 2. For example, data with a water-to-binder ratio (w/b) of 0.8 were identified as anomalous data, consistent with the knowledge of typical HPFRCC with low w/b (<0.35).

Table 2. Results of data anomaly detection.

Items	Number of Anomalous Data
Sand-to-binder ratio	7
Water-to-binder ratio	6
Superplasticizer content	6
Fiber length	2
Fiber elastic modulus	2

The dataset sizes of the different mechanical properties are different because different papers reported different properties. For example, a significant number of papers only reported the tensile properties of HPFRCC. After performing data cleaning, the numbers of data for the compressive strength, the tensile strength, and the tensile strain capacity are respectively 238, 247, and 266 in Dataset 1. Dataset 2 are established by incorporating the data generated by the micromechanics model for data augmentation, containing 317 data for predicting the tensile strain capacity.

3.2. Variable Selection

Figure 6a–c show that the Pearson correlation coefficients off the diagonal can be higher than 0.7, indicating that multicollinearity can occur if all the variables are used. Thus, the PCA is performed to reduce the dimensionality and eliminate multicollinearity for the datasets. Figure 6d–f show the results of the variance and the variance ratio for the datasets used to predict the three mechanical properties. With the threshold (0.99) of the cumulative variance ratio, the dimensionality of the input variables is reduced from 14 to 12 for the three datasets. The first 12 components with a high cumulative variance ratio are selected to construct the dataset. The correlation matrix after reducing the dimensionality of dataset is shown in Figure 6g. Because the correlation of each pair of variables is small

(less than 0.01), the correlation matrices of the compressive strength, tensile strength, and tensile strain capacity look the same.

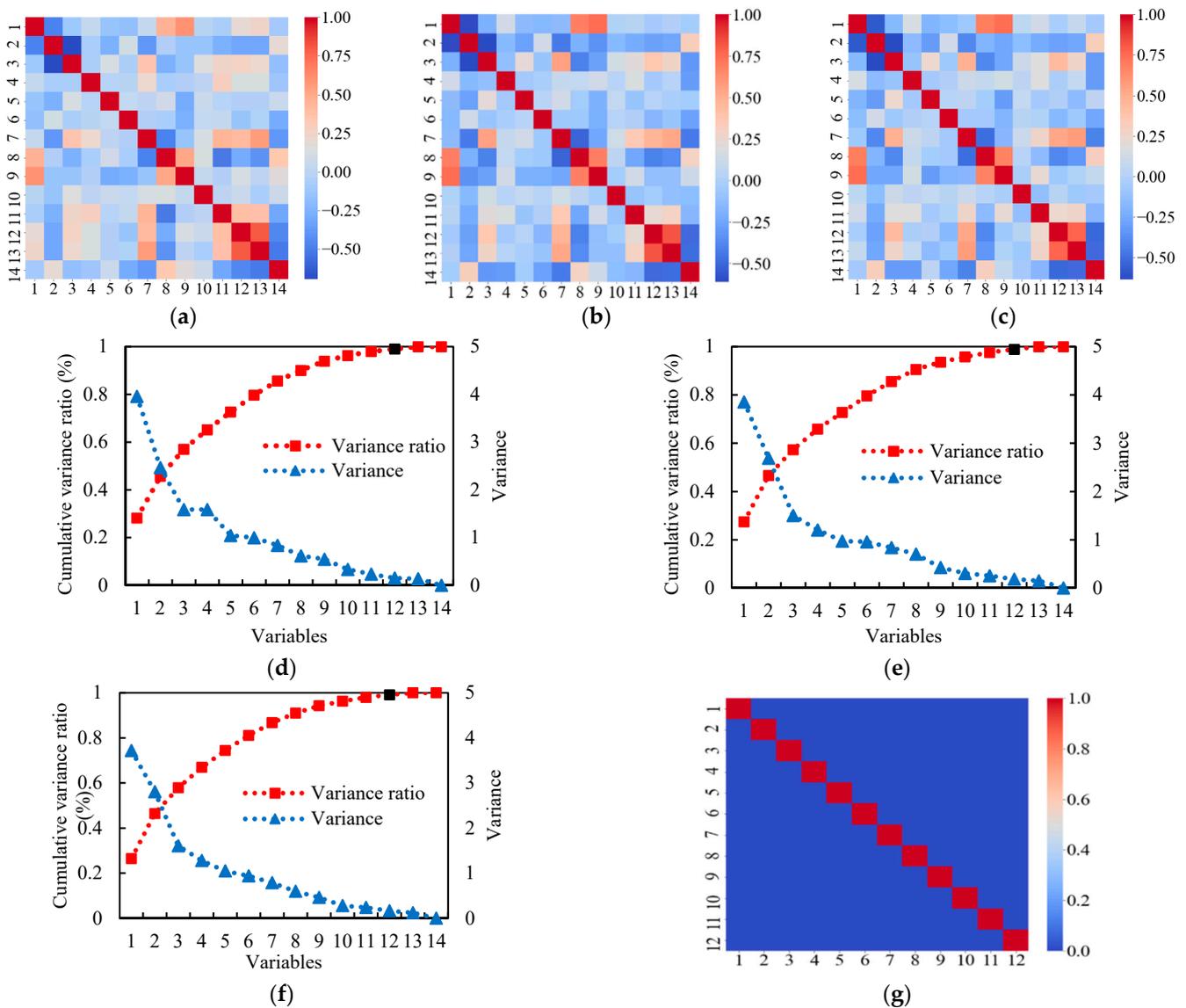


Figure 6. Multicollinearity analysis results. The correlation matrices for the compressive strength (a), tensile strength (b), and tensile strain capacity (c). Variance ratio and variance for the compressive strength (d), tensile strength (e), and tensile strain capacity (f). The correlation matrix for each pair of variables after PCA is shown in (g).

After the datasets are improved by the data cleaning and PCA, the datasets are used to train and test the machine learning models. Specifically, 75% of data are used for training, and 25% of data are used for testing of the machine learning models.

3.3. Hyperparameter Tuning

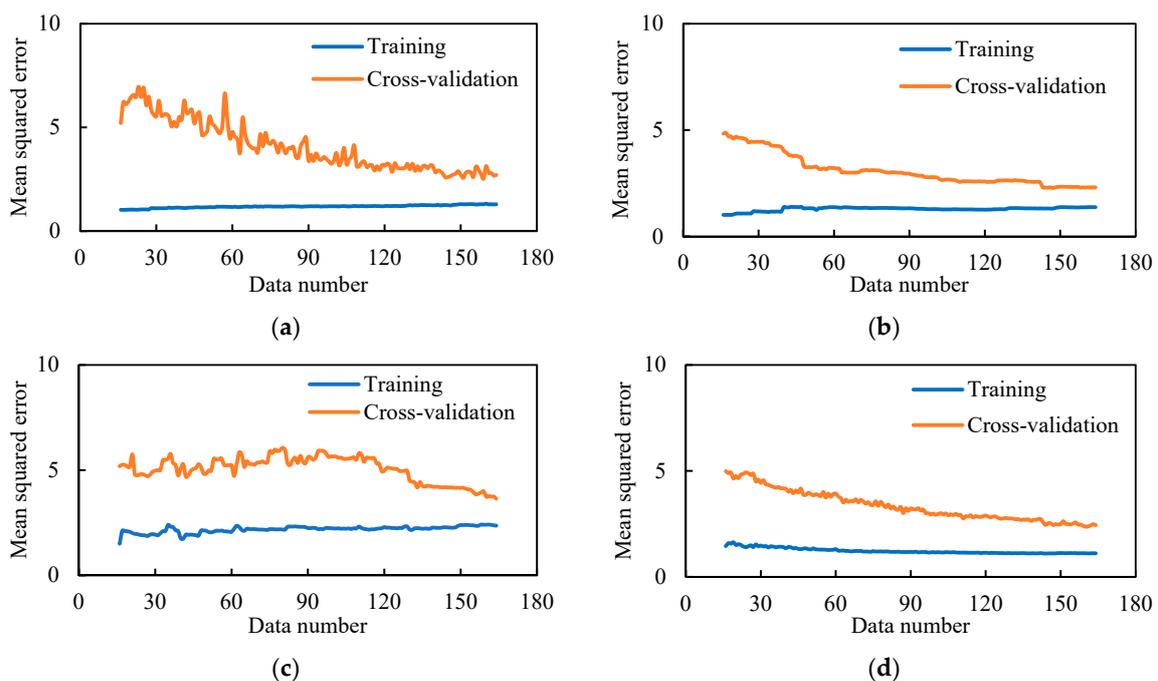
Table 3 lists the optimal hyperparameters of the machine learning models for the different properties. For the same machine learning method, the optimal hyperparameters are different for the different properties. Therefore, different models must be used to predict the different properties.

Table 3. The optimal hyperparameters for the machine learning methods.

Method	Hyperparameter	Range	Optimal Values for Different Properties		
			Compressive Strength	Tensile Strength	Tensile Strain Capacity
ANN	Hidden layer size	15–100	90	40	41
	Learning rate	0.0001–1.0	0.001	0.001	0.001
SVR	C	1–40	37	12	6
	Gamma	0.1–1.0	0.6	0.2	0.1
	Epsilon	0.1–1.0	0.1	0.2	0.2
CART	Maximum depth	2–10	4	4	4
	Maximum leaf nodes	2–10	8	9	7
	Minimum samples leaf	2–10	2	3	9
	Minimum samples split	2–10	6	9	2
XGBoost	Learning rate	0.001–1.0	0.1	0.1	0.1
	Estimator number	20–3000	1000	100	1877
	Gamma	0–10	0.667	0.333	0
	Maximum depth	1–10	2	5	8
	Column sample by tree	0–10	1	1.0	1.0
	Subsample ratio	0–1.0	0.3	0.3	0.3
	Lambda	0–100	33.3	11.1	16.7
Alpha	0–10	2.2	2.0	2.0	

3.4. Training Process

The optimal hyperparameters listed in Table 3 are used to train the machine learning models. In the training process, the MSE values of the different machine learning methods are changed, as shown in Figure 7. As the data number increases, the MSE of the training dataset increases because it becomes more difficult for the machine learning model to fit the data; the MSE of the cross-validation decreases, meaning that the generalization performance of the machine learning model continues to be improved; the MSE of the cross-validation curve gets close to but is larger than the MSE of the training dataset, indicating that overfitting or underfitting does not occur.

**Figure 7.** Learning curve for the training process (tensile strength) using: (a) ANN, (b) SVR, (c) CART, and (d) XGBoost.

3.5. Prediction Results of Mechanical Properties

Based on the trained machine learning models, the compressive strength, tensile strength, and tensile strain capacity can be predicted. The prediction results are compared with the actual test results, as shown in Table 4.

Table 4. Comparison of the predicted and the actual values of the mechanical properties.

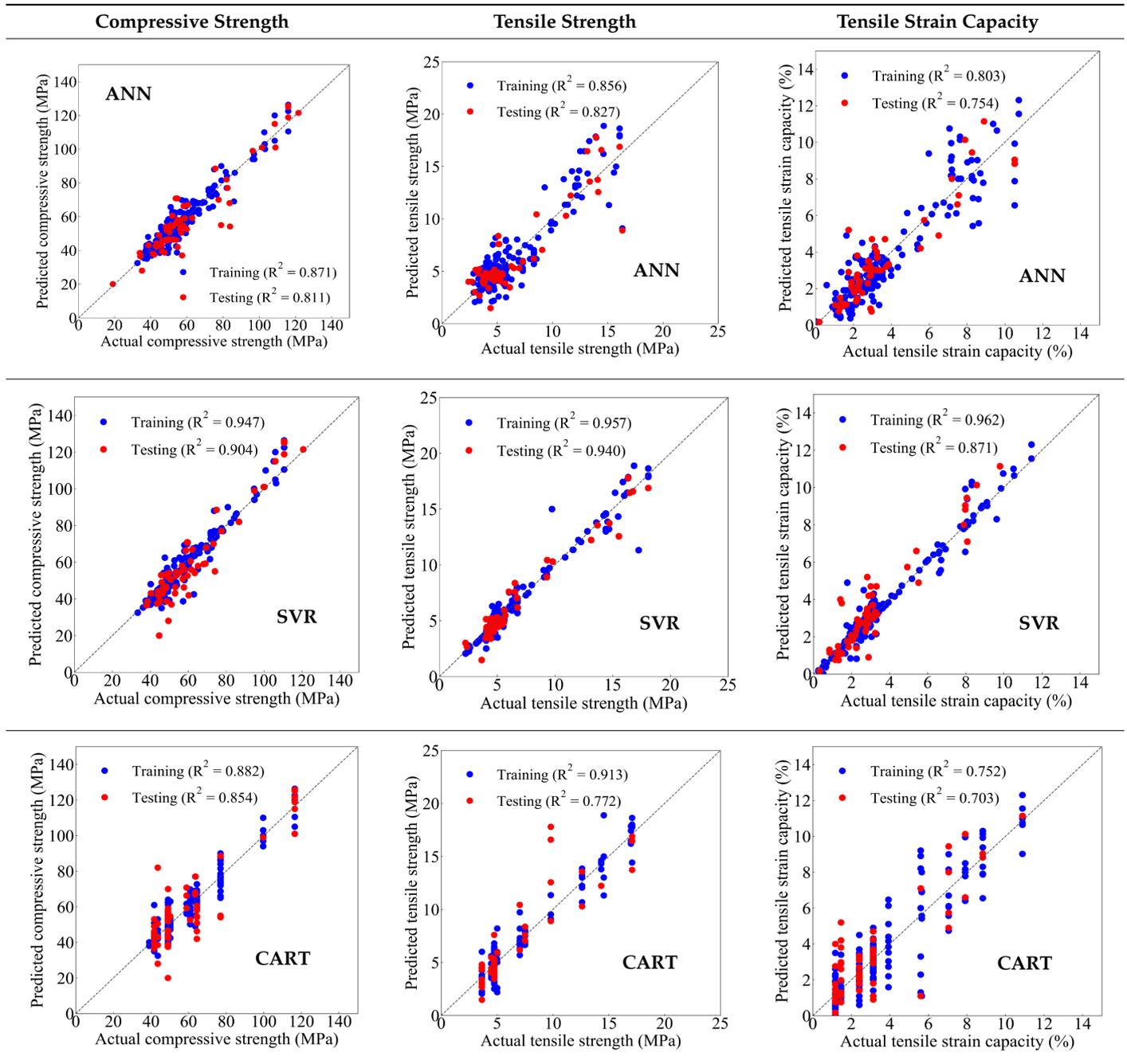
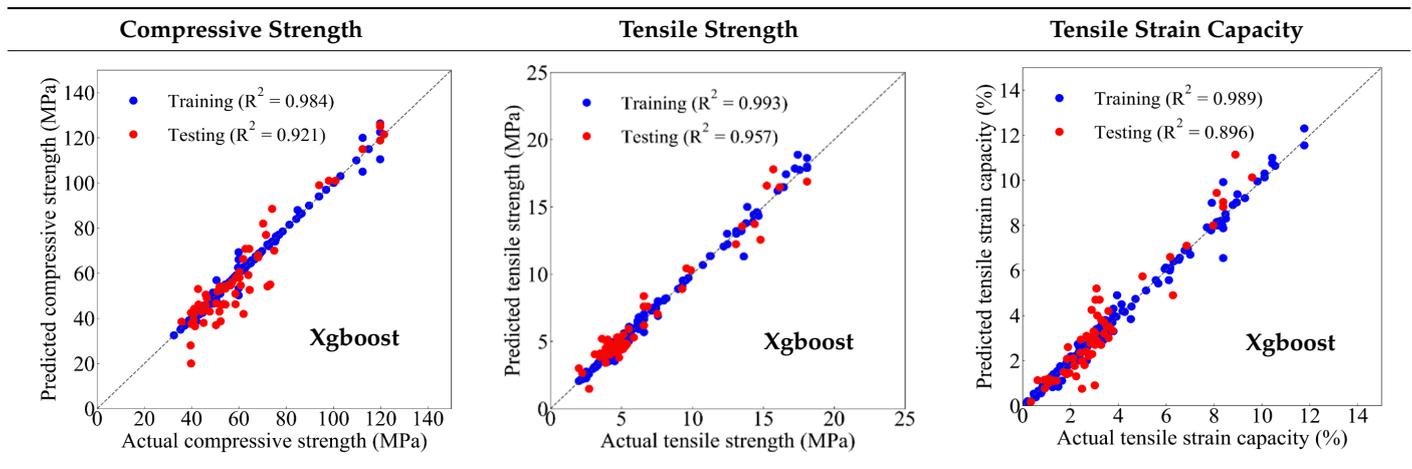


Table 4. Cont.



The prediction accuracy is reflected by the R^2 value, and a large R^2 value indicates a high prediction accuracy. The results corresponding to the training and the testing datasets are respectively considered in the comparison. Among the four machine learning methods, the XGBoost method shows the highest accuracy for all the three investigated mechanical properties, followed by the SVR method and then the ANN method. The CART method shows the lowest accuracy for all the three properties. With the XGBoost method, the R^2 values of the compressive strength, tensile strength, and tensile strain capacity are 0.984, 0.993, and 0.989, respectively, for the training dataset; and the R^2 values of the compressive strength, tensile strength, and tensile strain capacity are 0.921, 0.957, and 0.896, respectively, for the testing dataset. The high accuracy of the XGBoost model can be attributed to its architecture, as shown in Figure 1d, which can better represent the relationship between input and output variables.

The predicted results of compressive strength, tensile strength, and tensile strain capacity from the ANN, SVR, CART, and XGBoost models are summarized in Table 5. For the prediction of the compressive strength, the XGBoost model exhibits the highest accuracy: $R^2 = 0.921$, $R = 0.966$, and $MSE = 45.57$. For the prediction of the tensile strength, the XGBoost model shows the highest accuracy: $R^2 = 0.957$, $R = 0.980$, $MSE = 0.602$. For the prediction of tensile strain capacity, the XGBoost model also shows the highest accuracy: $R^2 = 0.896$, $R = 0.955$, and $MSE = 0.617$. Although XGBoost shows the highest accuracy, the prediction accuracy for the tensile strain capacity is relatively low (lower than 0.90), compared with the accuracy of the compressive strength and the tensile strength. Further improvement is needed for the tensile strain capacity.

Table 5. Evaluation of prediction results from the machine learning models.

Model	Set	Evaluation	Compressive Strength	Tensile Strength	Tensile Strain Capacity
ANN	Training	R^2	0.871	0.856	0.803
		R	0.933	0.925	0.882
		MSE	59.006	2.282	0.913
	Testing	R^2	0.811	0.827	0.754
		R	0.916	0.911	0.876
		MSE	69.513	2.498	0.925
SVR	Training	R^2	0.947	0.957	0.962
		R	0.973	0.979	0.981
		MSE	24.140	0.631	0.765
	Testing	R^2	0.904	0.940	0.871
		R	0.952	0.978	0.944
		MSE	35.228	0.663	0.962

Table 5. Cont.

Model	Set	Evaluation	Compressive Strength	Tensile Strength	Tensile Strain Capacity
CART	Training	R ²	0.882	0.913	0.752
		R	0.928	0.958	0.868
		MSE	52.823	1.299	1.723
	Testing	R ²	0.854	0.772	0.703
		R	0.733	0.880	0.836
		MSE	100.754	3.258	1.886
XGBoost	Training	R ²	0.984	0.993	0.989
		R	0.992	0.996	0.996
		MSE	6.268	0.130	0.063
	Testing	R ²	0.921	0.957	0.896
		R	0.966	0.980	0.955
		MSE	45.570	0.602	0.617

3.6. Effect of Supplemental Data

To further improve the prediction accuracy for the tensile strain capacity, Dataset 2 which includes the supplemental data generated from the semi-empirical model is used to train the machine learning models. After data augmentation, the dataset for the prediction of tensile strain capacity increases from 247 to 317. The correlation map for the variables is plotted in Figure 8. In Figure 8a, when the 14 variables are used, the multicollinearity occurs. In Figure 8b, the dataset is improved by the PCA to reduce the dimensionality from 14 to 12 and remove the multicollinearity.

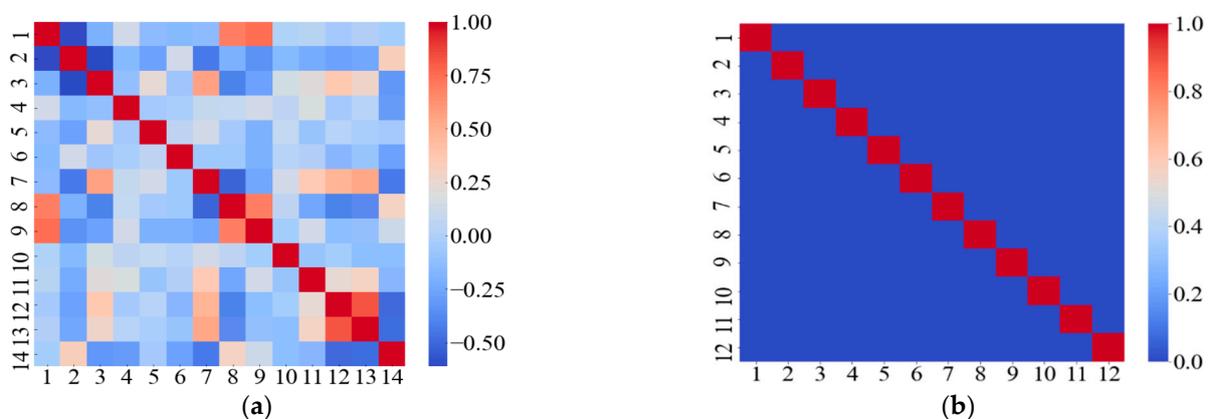


Figure 8. Results of the correlation matrices for input variables: (a) before PCA dimensionality reduction, and (b) after PCA dimensionality reduction.

Then, the improved Dataset 2 is adopted to train the predictive models using the four machine learning methods, and the evaluation results for the training and the testing dataset are shown in Table 6. Compared with Dataset 1, Dataset 2 improves the R² of testing dataset from 0.754 to 0.868 for the ANN model, from 0.871 to 0.907 for the SVR model, from 0.703 to 0.817 for the CART model, and from 0.896 to 0.912 for the XGBoost model, respectively. Therefore, the prediction performance for four machine learning models is improved by using the proposed dataset augmentation method based on the utilization of the micromechanics model.

Table 6. Evaluation of prediction results using enlarged dataset.

Model	Datasets	Evaluation	Tensile Strain Capacity	
			Dataset 1	Dataset 2
ANN	Training	R ²	0.803	0.958
		R	0.882	0.994
		MSE	0.913	0.102
	Testing	R ²	0.754	0.868
		R	0.876	0.948
		MSE	0.925	0.673
SVR	Training	R ²	0.962	0.971
		R	0.981	0.986
		MSE	0.765	0.234
	Testing	R ²	0.871	0.907
		R	0.944	0.954
		MSE	0.962	0.608
CART	Training	R ²	0.752	0.833
		R	0.868	0.972
		MSE	1.723	0.450
	Testing	R ²	0.703	0.817
		R	0.836	0.910
		MSE	1.886	1.190
XGBoost	Training	R ²	0.989	0.987
		R	0.996	0.994
		MSE	0.063	0.102
	Testing	R ²	0.896	0.912
		R	0.955	0.968
		MSE	0.617	0.673

3.7. Implementation of the Predictive Models

In this section, the XGBoost models are used to predict the compressive strength, tensile strength, and tensile strain capacity. In [77], as metakaolin was used to partially replace fly ash at a percentage of 0 to 40%, the compressive strength was increased from 55.3 MPa to 72.7 MPa because the metakaolin was more reactive. The trained XGBoost model is used to predict the compressive strength, as shown in Figure 9a. The results show that the model can reasonably predict the compressive strength. In [22], as fly ash was used to partially replace cement, the tensile strength of the mixture was changed. The trained XGBoost model is used to predict the tensile strength, as shown in Figure 9b. The results show that the model can reasonably predict the tensile strength. In [63], as slag was used to partially replace cement at a percentage of 0 to 30%, the tensile strain capacity was changed. The XGBoost models that are respectively trained using Dataset 1 and Dataset 2 are used to predict the tensile strain capacity, as shown in Figure 9c. The results show that the model can reasonably predict the tensile strain capacity. These results show that the developed machine learning models are promising for parametric studies on the effects of the mix design variables on the mechanical properties.

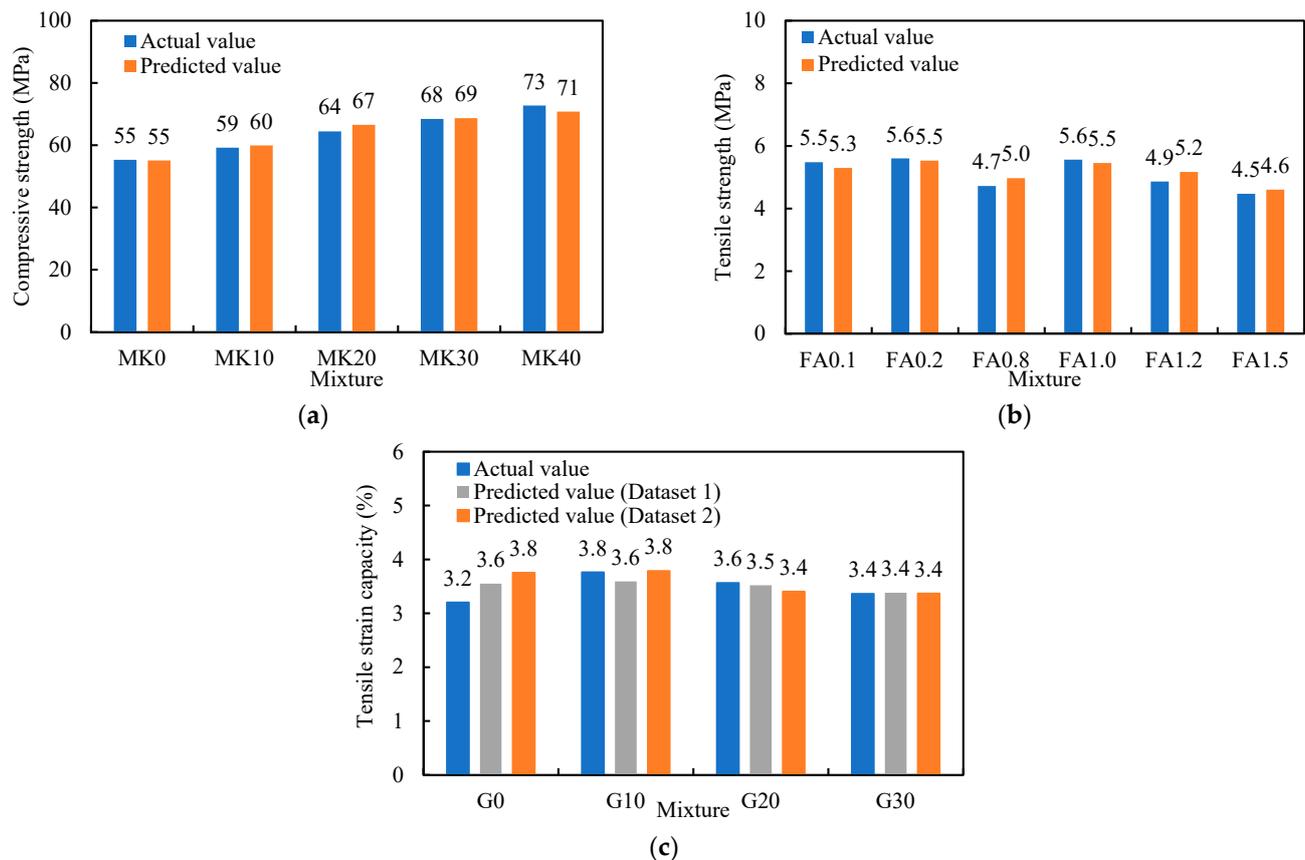


Figure 9. Comparison of the prediction against the test results: (a) compressive strength [77]; (b) tensile strength [22]; and (c) tensile strain capacity [63].

4. Conclusions

This research develops a new paradigm for prediction of the mechanical properties of HPFRCC by integrating the micromechanics and machine learning. Two strategies are presented to utilize micromechanics. Multiple methods are proposed to improve the prediction accuracy through improving the datasets. Four machine learning models are compared and used to predict the compressive strength, tensile strength, and tensile strain capacity of HPFRCC.

Based on the above investigations, the following conclusions can be drawn:

- The proposed methods provide reasonable prediction accuracy for the tensile strain capacity (or ductility), as well as the compressive and tensile strengths of HPFRCC. Among the investigated machine learning methods, the XGBoost method shows the highest prediction accuracy for all the investigated mechanical properties. With the training dataset, R^2 of the compressive strength, tensile strength, and ductility reached 0.984, 0.993, and 0.989, respectively. With the testing dataset, R^2 of the compressive strength, tensile strength, and ductility reached 0.921, 0.957, and 0.896, respectively.
- The prediction accuracy for the tensile strain capacity can be further improved by using the supplemental data generated from the micromechanics model. With the addition of only 70 more data, the R^2 values of the tensile strain capacity is increased from 0.896 to 0.912 for the training results.
- The predictive models are implemented to predict the mechanical properties of HPFRCC. The comparison of the prediction and test results further proves the prediction accuracy of the developed models. The implementation also demonstrates possible use cases of the predictive models for replacing or supplementing the experimental tests in the development and optimization of HPFRCC.

Future research is needed to investigate the performance of the proposed method for prediction of the other important properties of HPFRCC, such as the fresh properties (e.g., flowability) and the durability, and more research is needed to test the applicability of the method for other composites. It is envisioned that the developed prediction method can be used to facilitate optimization of the mix design of HPFRCC, so as to maximize the mechanical properties, the cost-effectiveness, and the durability, while minimizing the environmental impacts (e.g., carbon footprint and energy consumption).

Author Contributions: Conceptualization, Y.B.; methodology, P.G. and W.M.; software, P.G.; validation, M.X.; formal analysis, P.G.; investigation, P.G.; resources, Y.B.; data curation, P.G.; writing—original draft preparation, P.G.; writing—review and editing, W.M., M.X., V.C.L., and Y.B.; visualization, P.G. and Y.B.; supervision, P.G.; project administration, P.G.; funding acquisition, W.M. and Y.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science Foundation, grant number CMMI-2046407.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The language of the paper was edited by Brian Katat.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, S.; Li, V.C. Polyvinyl alcohol fiber reinforced engineered cementitious composites: Material design and performances. In Proceedings of the International Workshop on HPFRCC Structural Applications, Honolulu, HI, USA, 23–26 May 2006; Available online: <http://hdl.handle.net/2027.42/84790> (accessed on 1 March 2021).
2. Zhang, Z.; Qian, S.; Ma, H. Investigating mechanical properties and self-healing behavior of micro-cracked ECC with different volume of fly ash. *Constr. Build. Mater.* **2014**, *52*, 17–23. [[CrossRef](#)]
3. Pan, Z.; Wu, C.; Liu, J.; Wang, W.; Liu, J. Study on mechanical properties of cost-effective polyvinyl alcohol engineered cementitious composites (PVA-ECC). *Constr. Build. Mater.* **2015**, *78*, 397–404. [[CrossRef](#)]
4. Kim, J.-K.; Kim, J.-S.; Ha, G.J.; Kim, Y.Y. Tensile and fiber dispersion performance of ECC (engineered cementitious composites) produced with ground granulated blast furnace slag. *Cem. Concr. Res.* **2007**, *37*, 1096–1105. [[CrossRef](#)]
5. Meng, W.; Valipour, M.; Khayat, K.H. Optimization and performance of cost-effective ultra-high performance concrete. *Mater. Struct.* **2016**, *50*, 1–16. [[CrossRef](#)]
6. Meng, W.; Khayat, K.H. Improving flexural performance of ultra-high-performance concrete by rheology control of suspending mortar. *Compos. Part B Eng.* **2017**, *117*, 26–34. [[CrossRef](#)]
7. Meng, W.; Khayat, K. Effects of saturated lightweight sand content on key characteristics of ultra-high-performance concrete. *Cem. Concr. Res.* **2017**, *101*, 46–54. [[CrossRef](#)]
8. Meng, W.; Samaranayake, A.; Khayat, K.H. Factorial design and optimization of ultra-high-performance concrete with lightweight sand. *ACI Mater. J.* **2018**, *115*, 129–138. [[CrossRef](#)]
9. Xu, M.; Bao, Y.; Wu, K.; Xia, T.; Clack, H.L.; Shi, H.; Li, V.C. Influence of TiO₂ incorporation methods on NO_x abatement in Engineered Cementitious Composites. *Constr. Build. Mater.* **2019**, *221*, 375–383. [[CrossRef](#)]
10. Xu, M.; Clack, H.; Xia, T.; Bao, Y.; Wu, K.; Shi, H.; Li, V. Effect of TiO₂ and fly ash on photocatalytic NO_x abatement of engineered cementitious composites. *Constr. Build. Mater.* **2020**, *236*, 117559. [[CrossRef](#)]
11. Sahmaran, M.; Yildirim, G.; Erdem, T.K. Self-healing capability of cementitious composites incorporating different supplementary cementitious materials. *Cem. Concr. Compos.* **2013**, *35*, 89–101. [[CrossRef](#)]
12. Guo, P.; Meng, W.; Nassif, H.; Gou, H.; Bao, Y. New perspectives on recycling waste glass in manufacturing concrete for sustainable civil infrastructure. *Constr. Build. Mater.* **2020**, *257*, 119579. [[CrossRef](#)]
13. Xu, M.; Bao, Y.; Wu, K.; Shi, H.; Guo, X.; Li, V.C. Multiscale investigation of tensile properties of a TiO₂-doped Engineered Cementitious Composite. *Constr. Build. Mater.* **2019**, *209*, 485–491. [[CrossRef](#)]
14. Zhang, Y.; Deng, M.; Dong, Z. Seismic response and shear mechanism of engineered cementitious composite (ECC) short columns. *Eng. Struct.* **2019**, *192*, 296–304. [[CrossRef](#)]
15. Li, X.; Wang, J.; Bao, Y.; Chen, G. Cyclic behavior of damaged reinforced concrete columns repaired with high-performance fiber-reinforced cementitious composite. *Eng. Struct.* **2017**, *136*, 26–35. [[CrossRef](#)]
16. Leung, C.K.; Cheung, Y.N.; Zhang, J. Fatigue enhancement of concrete beam with ECC layer. *Cem. Concr. Res.* **2007**, *37*, 743–750. [[CrossRef](#)]

17. Liu, Y.; Zhang, Q.; Bao, Y.; Bu, Y. Static and fatigue push-out tests of short headed shear studs embedded in Engineered Cementitious Composites (ECC). *Eng. Struct.* **2019**, *182*, 29–38. [[CrossRef](#)]
18. Li, X.; Bao, Y.; Wu, L.; Yan, Q.; Ma, H.; Chen, G.; Zhang, H. Thermal and mechanical properties of high-performance fiber-reinforced cementitious composites after exposure to high temperatures. *Constr. Build. Mater.* **2017**, *157*, 829–838. [[CrossRef](#)]
19. Li, X.; Bao, Y.; Xue, N.; Chen, G. Bond strength of steel bars embedded in high-performance fiber-reinforced cementitious composite before and after exposure to elevated temperatures. *Fire Saf. J.* **2017**, *92*, 98–106. [[CrossRef](#)]
20. Yang, E.-H.; Wang, S.; Yang, Y.; Li, V.C. Fiber-bridging constitutive law of engineered cementitious composites. *J. Adv. Concr. Technol.* **2008**, *6*, 181–193. [[CrossRef](#)]
21. Spagnoli, A.; Yang, E.-H.; Li, V.C. Micromechanical modelling of multiple fracture in engineered cementitious composites. In Proceedings of the 17th European Conference Fracture, Brno, Czech Republic, 2–5 September 2008; pp. 2407–2414. Available online: <https://deepblue.lib.umich.edu/handle/2027.42/84800> (accessed on 1 March 2021).
22. Guo, P.; Bao, Y.; Meng, W. Review of using glass in high-performance fiber-reinforced cementitious composites. *Cem. Concr. Compos.* **2021**, *120*, 104032. [[CrossRef](#)]
23. Yu, K.-Q.; Lu, Z.-D.; Dai, J.-G.; Shah, S.P. Direct tensile properties and stress–strain model of UHP-ECC. *J. Mater. Civ. Eng.* **2020**, *32*, 04019334. [[CrossRef](#)]
24. Ghafari, E.; Bandarabadi, M.; Costa, H.; Júlio, E. Design of UHPC using artificial neural networks. *Brittle Matrix Compos.* **2012**, *10*, 61–69. [[CrossRef](#)]
25. Prasad, B.R.; Eskandari, H.; Reddy, B.V. Prediction of compressive strength of SCC and HPC with high volume fly ash using ANN. *Constr. Build. Mater.* **2009**, *23*, 117–128. [[CrossRef](#)]
26. Abbas, H.; Al-Salloum, Y.A.; Elsanadedy, H.M.; Almusallam, T.H. ANN models for prediction of residual strength of HSC after exposure to elevated temperature. *Fire Saf. J.* **2019**, *106*, 13–28. [[CrossRef](#)]
27. Abu Yaman, M.; Elaty, M.A.; Taman, M. Predicting the ingredients of self compacting concrete using artificial neural network. *Alex. Eng. J.* **2017**, *56*, 523–532. [[CrossRef](#)]
28. Akande, K.O.; Owolabi, T.O.; Twaha, S.; Olatunji, S. Performance comparison of SVM and ANN in predicting compressive strength of concrete. *IOSR J. Comput. Eng.* **2014**, *16*, 88–94. [[CrossRef](#)]
29. Hammoudi, A.; Moussaceb, K.; Belebchouche, C.; Dahmoune, F. Comparison of artificial neural network (ANN) and response surface methodology (RSM) prediction in compressive strength of recycled concrete aggregates. *Constr. Build. Mater.* **2019**, *209*, 425–436. [[CrossRef](#)]
30. Naderpour, H.; Rafiean, A.H.; Fakharian, P. Compressive strength prediction of environmentally friendly concrete using artificial neural networks. *J. Build. Eng.* **2018**, *16*, 213–219. [[CrossRef](#)]
31. Azimi-Pour, M.; Eskandari-Naddaf, H.; Pakzad, A. Linear and non-linear SVM prediction for fresh properties and compressive strength of high volume fly ash self-compacting concrete. *Constr. Build. Mater.* **2020**, *230*, 117021. [[CrossRef](#)]
32. Young, B.A.; Hall, A.; Pilon, L.; Gupta, P.; Sant, G. Can the compressive strength of concrete be estimated from knowledge of the mixture proportions? New insights from statistical analysis and machine learning methods. *Cem. Concr. Res.* **2019**, *115*, 379–388. [[CrossRef](#)]
33. Yan, K.; Shi, C. Prediction of elastic modulus of normal and high strength concrete by support vector machine. *Constr. Build. Mater.* **2010**, *24*, 1479–1485. [[CrossRef](#)]
34. Behnood, A.; Olek, J.; Glinicki, M.A. Predicting modulus elasticity of recycled aggregate concrete using M5' model tree algorithm. *Constr. Build. Mater.* **2015**, *94*, 137–147. [[CrossRef](#)]
35. Demir, F. Prediction of elastic modulus of normal and high strength concrete by artificial neural networks. *Constr. Build. Mater.* **2008**, *22*, 1428–1435. [[CrossRef](#)]
36. Cao, Y.F.; Wu, W.; Zhang, H.L.; Pan, J.M. Prediction of the elastic modulus of self-compacting concrete based on SVM. *Appl. Mech. Mater.* **2013**, *357–360*, 1023–1026. [[CrossRef](#)]
37. Hossain, K.M.A.; Anwar, M.S.; Samani, S.G. Regression and artificial neural network models for strength properties of engineered cementitious composites. *Neural Comput. Appl.* **2018**, *29*, 631–645. [[CrossRef](#)]
38. Dong, W.; Huang, Y.; Lehane, B.; Ma, G. XGBoost algorithm-based prediction of concrete electrical resistivity for structural health monitoring. *Autom. Constr.* **2020**, *114*, 103155. [[CrossRef](#)]
39. Hajnayeb, A.; Ghasemlooia, A.; Khadem, S.; Moradi, M. Application and comparison of an ANN-based feature selection method and the genetic algorithm in gearbox fault diagnosis. *Expert Syst. Appl.* **2011**, *38*, 10205–10209. [[CrossRef](#)]
40. Chou, J.-S.; Tsai, C.-F.; Pham, A.-D.; Lu, Y.-H. Machine learning in concrete strength simulations: Multi-nation data analytics. *Constr. Build. Mater.* **2014**, *73*, 771–780. [[CrossRef](#)]
41. Yaseen, Z.M.; Deo, R.C.; Hilal, A.; Abd, A.M.; Bueno, L.C.; Salcedo-Sanz, S.; Nehdi, M.L. Predicting compressive strength of lightweight foamed concrete using extreme learning machine model. *Adv. Eng. Softw.* **2018**, *115*, 112–125. [[CrossRef](#)]
42. Yu, L.; Wang, S.; Lai, K.K. Forecasting foreign exchange rates using an SVR-based neural network ensemble. In *Advances in Banking Technology and Management*; IGI Global: Hershey, PA, USA, 2008; pp. 261–277.
43. Amari, S.; Wu, S. Improving support vector machine classifiers by modifying kernel functions. *Neural Netw.* **1999**, *12*, 783–789. [[CrossRef](#)]

44. Gordon, L. Using classification and regression trees (CART) in SAS®enterprise miner TM for applications in public health. In Proceedings of the SAS Global Forum, San Francisco, CA, USA, 28 April–1 May 2013. Available online: <https://support.sas.com/resources/papers/proceedings13/089-2013.pdf> (accessed on 1 March 2021).
45. Lewis, R.J. An introduction to classification and regression tree (CART) analysis. In Proceedings of the Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, San Francisco, CA, USA, 22–25 May 2000. Available online: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.95.4103> (accessed on 1 March 2021).
46. Behnood, A.; Golafshani, E.M. Machine learning study of the mechanical properties of concretes containing waste foundry sand. *Constr. Build. Mater.* **2020**, *243*, 118152. [[CrossRef](#)]
47. Li, V.C. Engineered cementitious composites (ECC)-tailored composites through micromechanical modeling. In *Fiber Reinforced Concrete: Present and the Future*; Canadian Society for Civil Engineering: Montreal, QC, Canada, 1998. Available online: <http://hdl.handle.net/2027.42/84667> (accessed on 1 March 2021).
48. Li, V.C.; Wu, C.; Wang, S.; Ogawa, A.; Saito, T. Interface tailoring for strain-hardening polyvinyl alcohol-engineered cementitious composite (PVA-ECC). *ACI Mater. J.* **2002**, *99*, 463–472. [[CrossRef](#)]
49. Ohno, M. Green and Durable Geopolymer Composites for Sustainable Civil Infrastructure. Ph.D. Thesis, University of Michigan, Ann Harbor, MI, USA, 2017. Available online: <http://hdl.handle.net/2027.42/140947> (accessed on 1 March 2021).
50. Yu, K.; Ding, Y.; Liu, J.; Bai, Y. Energy dissipation characteristics of all-grade polyethylene fiber-reinforced engineered cementitious composites (PE-ECC). *Cem. Concr. Compos.* **2020**, *106*, 103459. [[CrossRef](#)]
51. Said, S.; Razak, H.A. The effect of synthetic polyethylene fiber on the strain hardening behavior of engineered cementitious composite (ECC). *Mater. Des.* **2015**, *86*, 447–457. [[CrossRef](#)]
52. Zhou, J.; Qian, S.; Beltran, M.G.S.; Ye, G.; Van Breugel, K.; Li, V.C. Development of engineered cementitious composites with limestone powder and blast furnace slag. *Mater. Struct.* **2009**, *43*, 803–814. [[CrossRef](#)]
53. Bao, Y.; Li, V.C. Feasibility study of lego-inspired construction with bendable concrete. *Autom. Constr.* **2020**, *113*, 103161. [[CrossRef](#)]
54. Lepech, M.D.; Li, V.C.; Robertson, R.E.; Keoleian, G.A. Design of green engineered cementitious composites for improved sustainability. *ACI Mater. J.* **2008**, *105*, 567. [[CrossRef](#)]
55. Zheng, Y.; Zhang, L.; Xia, L. Investigation of the behaviour of flexible and ductile ECC link slab reinforced with FRP. *Constr. Build. Mater.* **2018**, *166*, 694–711. [[CrossRef](#)]
56. Li, X.; Xu, Z.; Bao, Y.; Cong, Z. Post-fire seismic behavior of two-bay two-story frames with high-performance fiber-reinforced cementitious composite joints. *Eng. Struct.* **2019**, *183*, 150–159. [[CrossRef](#)]
57. Ding, Y.; Yu, J.-T.; Yu, K.; Xu, S.-L. Basic mechanical properties of ultra-high ductility cementitious composites: From 40 MPa to 120 MPa. *Compos. Struct.* **2018**, *185*, 634–645. [[CrossRef](#)]
58. Lin, J.-X.; Song, Y.; Xie, Z.-H.; Guo, Y.-C.; Yuan, B.; Zeng, J.-J.; Wei, X. Static and dynamic mechanical behavior of engineered cementitious composites with PP and PVA fibers. *J. Build. Eng.* **2020**, *29*, 101097. [[CrossRef](#)]
59. Ding, Y.; Yu, K.; Yu, J.-T.; Xu, S.-L. Structural behaviors of ultra-high performance engineered cementitious composites (UHP-ECC) beams subjected to bending-experimental study. *Constr. Build. Mater.* **2018**, *177*, 102–115. [[CrossRef](#)]
60. Yu, K.; Wang, Y.; Yu, J.; Xu, S. A strain-hardening cementitious composites with the tensile capacity up to 8%. *Constr. Build. Mater.* **2017**, *137*, 410–419. [[CrossRef](#)]
61. Wang, Y.; Liu, F.; Yu, J.; Dong, F.; Ye, J. Effect of polyethylene fiber content on physical and mechanical properties of engineered cementitious composites. *Constr. Build. Mater.* **2020**, *251*, 118917. [[CrossRef](#)]
62. Yu, K.-Q.; Dai, J.-G.; Lu, Z.-D.; Poon, C.-S. Rate-dependent tensile properties of ultra-high performance engineered cementitious composites (UHP-ECC). *Cem. Concr. Compos.* **2018**, *93*, 218–234. [[CrossRef](#)]
63. Zhu, Y.; Yang, Y.; Yao, Y. Use of slag to improve mechanical properties of engineered cementitious composites (ECCs) with high volumes of fly ash. *Constr. Build. Mater.* **2012**, *36*, 1076–1081. [[CrossRef](#)]
64. Turk, K.; Nehdi, M.L. Coupled effects of limestone powder and high-volume fly ash on mechanical properties of ECC. *Constr. Build. Mater.* **2018**, *164*, 185–192. [[CrossRef](#)]
65. Zhou, Y.; Xi, B.; Sui, L.; Zheng, S.; Xing, F.; Li, L. Development of high strain-hardening lightweight engineered cementitious composites: Design and performance. *Cem. Concr. Compos.* **2019**, *104*, 103370. [[CrossRef](#)]
66. Yu, K.; Zhu, W.; Ding, Y.; Lu, Z.-D.; Yu, J.-T.; Xiao, J.-Z. Micro-structural and mechanical properties of ultra-high performance engineered cementitious composites (UHP-ECC) incorporation of recycled fine powder (RFP). *Cem. Concr. Res.* **2019**, *124*, 105813. [[CrossRef](#)]
67. Li, Z.; Yue, J.; Hu, L.; Li, D.; Fu, Z. Weighted least square fitting based abnormal aquaculture water quality perception data elimination. *Sens. Lett.* **2012**, *10*, 529–534. [[CrossRef](#)]
68. Zhang, C.; Ma, Y. *Ensemble Machine Learning: Methods and Applications*; Springer: Berlin, Germany, 2012.
69. Friedrich, R.J. In defense of multiplicative terms in multiple regression equations. *Am. J. Political Sci.* **1982**, *26*, 797. [[CrossRef](#)]
70. Tabachnick, B.G.; Fidell, L.S.; Ullman, J.B. *Using Multivariate Statistics*; Pearson: Boston, MA, USA, 2007; Volume 5. Available online: <https://www.pearsonhighered.com/assets/preface/0/1/3/4/0134790545.pdf> (accessed on 1 March 2021).
71. Sulaiman, M.S.; Abood, M.M.; Sinnakaudan, S.K.; Shukor, M.R.; You, G.Q.; Chung, X.Z. Assessing and solving multicollinearity in sediment transport prediction models using principal component analysis. *ISH J. Hydraul. Eng.* **2019**, 1–11. [[CrossRef](#)]

72. PCA Whitening. Stanford Website. Available online: <http://ufldl.stanford.edu/tutorial/unsupervised/PCAWhitening/> (accessed on 1 March 2021).
73. Bao, Y.; Liu, Z. A fast grid search method in support vector regression forecasting time series. In Proceedings of the International Conference on Intelligent Data Engineering and Automated Learning, Yangzhou, China, 12–14 October 2006; Volume 4224, pp. 504–511.
74. Pham, B.T.; Son, L.H.; Hoang, T.-A.; Nguyen, D.-M.; Bui, D.T. Prediction of shear strength of soft soil using machine learning methods. *Catena* **2018**, *166*, 181–191. [[CrossRef](#)]
75. Boddy, R.; Smith, G. *Statistical Methods in Practice*; Wiley: Hoboken, NJ, USA, 2009.
76. Zhang, D. A coefficient of determination for generalized linear models. *Am. Stat.* **2017**, *71*, 310–316. [[CrossRef](#)]
77. Ismail, M.K.; Hassan, A.A.A.; Lachemi, M. Performance of self-consolidating engineered cementitious composite under drop-weight impact loading. *J. Mater. Civ. Eng.* **2019**, *31*, 04018400. [[CrossRef](#)]