Received: 12 March 2023 | Accepted: 14 January 2024

DOI: 10.1111/mice.13164

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING WILEY

RESEARCH ARTICLE

Chained machine learning model for predicting load capacity and ductility of steel fiber–reinforced concrete beams

Torkan Shafighfard[1] | **Farzin Kazemi**[2,3] | **Faramarz Bagherzadeh**[4] | **Magdalena Mieloszyk**[1] | **Doo-Yeol Yoo**[5]

[1]Institute of Fluid Flow Machinery, Polish Academy of Sciences, Gdańsk, Poland

[2]Faculty of Civil and Environmental Engineering, Gdańsk University of Technology, Gdańsk, Poland

[3]Department of Civil, Environmental & Geomatic Engineering, University College London, London, UK

[4]Department of Mathematics and Computer Science, University of Bremen, Bremen, Germany

[5]Department of Architecture and Architectural Engineering, Yonsei University, Seoul, South Korea

Correspondence
Faramarz Bagherzadeh, Mathematics and Computer Science, University of Bremen, Bremen, Germany.
Email: fabagher@uni-bremen.de

Doo-Yeol Yoo, Department of Architecture and Architectural Engineering, Yonsei University, Seoul, South Korea.
Email: dyyoo@yonsei.ac.kr

Funding information
University of Bremen; National Research Foundation of Korea, Grant/Award Number: 2021R1A2C4001503; Yonsei University, Grant/Award Number: 2023-22-0134

Abstract

One of the main issues associated with steel fiber–reinforced concrete (SFRC) beams is the ability to anticipate their flexural response. With a comprehensive grid search, several stacked models (i.e., chained, parallel) consisting of various machine learning (ML) algorithms and artificial neural networks (ANNs) were developed to predict the flexural response of SFRC beams. The flexural performance of SFRC beams under bending was assessed based on 193 experimental specimens from real-life beam models. The ML techniques were applied to predict SFRC beam responses to bending load as functions of the steel fiber properties, concrete elastic modulus, beam dimensions, and reinforcement details. The accuracy of the models was evaluated using the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE) of actual versus predicted values. The findings revealed that the proposed technique exhibited notably superior performance, delivering faster and more accurate predictions compared to both the ANNs and parallel models. Shapley diagrams were used to analyze variable contributions quantitatively. Shapley values show that the chained model prediction of ductility index is highly affected by two other targets (peak load and peak deflection) that show the chained algorithm utilizing the prediction of previous steps for enhancing the prediction of the target feature. The proposed model can be viewed as a function of significant input variables that permit the quick assessment of the likely performance of SFRC beams in bending.

1 | INTRODUCTION

Although the application of fiber-reinforced concrete (FRC) beams turns back to a few decades ago (Adhikary & Mutsuyoshi, 2006; Masuelli, 2013; Soltanzadeh et al., 2015), significant efforts also have been made to increase the strength and ductility of concrete in construction and building structures since sustainable infrastructure is crucial for economic development (Aldwaik & Adeli, 2016). Like other fiber-reinforced composite structures (Çelik & König, 2022; Rafiei & Adeli, 2017b; Shafighfard et al., 2021), it has recently been shown that FRC structures are capable of possessing more exceptional ductility and strength than normal concrete. The ability to predict the structural behavior of steel fiber–reinforced concrete (SFRC) beams is one of the many challenges that researchers face in investigating their performance (Rafiei et al., 2017; Singh, 2016; Venkateshwaran & Tan, 2018).

Among the numerous flexural parameters (Gribniak et al., 2012; Gribniak & Sokolov, 2023), the ductility ratio has drawn the attention of researchers due to its potential to reflect the reaction of structural elements to bending loads. Another important flexural metric is the bending load capacity (peak load), which has been studied by means of numerical simulations, experimental investigations, and machine learning (ML)-based prediction techniques.

Several researchers have conducted numerical and/or analytical studies of SFRC beams to decrease the labor and/or material costs associated with experimental studies (Jeong & Jo, 2021; Júnior & Parvin, 2022). The longitudinal rebar ratio and residual tensile strength are typical variables considered in parametric studies of the flexural performance of SFRC beams. The use of fibers to enhance tensile strength is not more effective than continuous reinforcement in improving the moment capacity of concrete beams, but fiber reinforcement increases stiffness and strength compared to plain RC beams (Mobasher et al., 2015). Tan et al. (2022) conducted a parametric analysis of the effects of material properties of SFRC on bending performance and found that the flexural ductility was negatively affected by a high volume fraction in RC beams. Three-dimensional (3D) modeling of SFRC beams with different fiber aspect ratios, orientations, and beam sizes showed that fibers oriented in the direction of tensile stress due to bending enhanced the peak load more than randomly distributed fibers. In addition, smaller beams with lower fiber reinforcement ratios exhibit higher peak loads (Al-Ahmed et al., 2022).

Experimental studies are typically considered complementary to numerical works (Pereira et al., 2020) specifically to verify the results they provided. Yang et al. (2020) employed steel fiber and fiber reinforcing ratios volume contents as variables in studying the flexural performance of ultra high-performance concrete (UHPC) beams reinforced with steel rebars. It is worth mentioning that steel fiber is the most frequently used fiber type for structural applications of UHPC. However, various types of fibers are made of glass (Xu & Humar, 2006), carbon (Finckh & Zilch, 2012), and polyvinyl alcohol (Zhang, Jin, et al., 2019). The UHPC beams' ductility index was considerably lower than that of high-strength concrete beams at a rebar ratio of 0.79%. The UHPC beams' ductility index was greater than that of plain concrete at a rebar ratio of 1.58%. These findings show that the ductility of UHPC beams is predominantly impacted by the presence of steel fibers at low rebar ratios.

Owing to fracture localization characteristics, beams containing steel fibers had lower ductility index values than beams without fibers. In contrast to SFRC beams, which collapsed through rebar rupture, beams without fibers were crushed in the compression zone. Fiber bridging at the crack surfaces effectively stopped the crack from spreading, but the length and type of fiber had little impact on the cracking response of reinforced UHPC beams (Yoo & Yoon, 2015).

The flexural properties of UHPC beams reinforced with steel fibers were investigated. The flexural properties of UHPC beams reinforced with steel fibers were investigated. The ductility index values obtained were 2.01, 1.79, 2.64, 3.275, and 3.39 for rebar ratios of 0.6%, 0.9%, 1.2%, 1.31%, and 1.96%, respectively. The rebar yielded until flexural failure of the UHPC beams with low rebar ratios, and the steel fiber–reinforced UHPC beams exhibited effective ductile behavior (Yang et al., 2010). It has also been demonstrated that the increase in the peak load under flexure attributable to the addition of steel fibers depends on the ratio between the longitudinal rebar ratio and the toughness of the material.

A relation between SFRC beam behavior and the fiber volume fraction was reported by Meda et al. (2012). Experiments were conducted to assess the effects of the fiber type and volume fraction on the flexural performance of SFRC beams with low longitudinal rebar ratio. The primary factors were found to be the fiber type, fiber volume fraction, and axial reinforcement ratio. The flexural ductility of SFRC beams with low longitudinal rebar ratios was found to decrease as the volume fraction of steel fibers increased (Gümüş & Arslan, 2019).

In another study, SFRC samples exhibited larger peak load values than conventional concrete samples for almost all longitudinal rebar ratios. Up to a longitudinal rebar ratio of 0.40%, the ultimate deflection values of the under-reinforced portions of conventional concrete specimens

were higher than those of the SFRC specimens. However, the ultimate deflections of the SFRC specimens were 50%–300% greater than those of conventional concrete specimens when the longitudinal rebar ratio exceeded 0.40% (Mertol et al., 2015).

The impending Industry 4.0 revolution is anticipated to be characterized by automated design, the use of building information modeling, 3D printing, and the integration of intelligent materials in shaping future infrastructure. Construction and building materials with high ductility and flexural load capacity, such as SFRC beams, have the potential to become part of automated building processes. Recently, ML techniques have become prevalent and have contributed to filling gaps in automation processes owing to the availability of experimental and numerical data (Bagherzadeh, Freitag, et al., 2023; Maeda et al., 2021; Zhang & Yuen, 2021). Hence, the use of ML approaches to address structural engineering challenges has increased significantly in recent years (Kazemi & Jankowski, 2023; Pak et al., 2023). The mechanical properties of structural members, including their shear and compressive strength, have been predicted using various input parameters by Sandeep et al. (2023). It was shown that XGBoost and ANNs provided better performance for shear strength prediction of reinforced beams.

Failure modes have been investigated and predicted using ML algorithms proposed by Solhmirzaei et al. (2020). It was concluded that the KNN model provides superior performance in identifying flexural failure modes, while Genetic programming depicts acceptable predictions of shear capacity of reinforced beams by R^2 = 92%. Furthermore, the random forest ML technique integrated with the SHapley Additive exPlanations (SHAP) method was employed and the preciseness of 84% and 86% were provided (Mangalathu et al., 2020). The seismic response is another factor several researchers considered. This response was predicted for RC structure through several algorithms and lasso regression provided better performance with an accuracy of 81%. The importance of different input variables was also discussed (Mangalathu & Jeon, 2018). Moreover, a novel method called locally weighted least squares support vector machines for regression was proposed to evaluate the seismic response of RC beams. The new technique provided R^2 = 81% for the flexure failure (Luo & Paal, 2019).

Cardoso et al. (2019) investigated the effects of volume fraction and aspect ratio of hooked-end steel fibers on the flexural behavior of SFRC beams. In their study, hooked-end steel fibers with aspect ratios of 45 or 80 were used, with a volume fraction ranging from 0% to 2%. The flexural ductility of SFRC beams decreased as the fiber volume fraction and aspect ratio increased. The reduction, in relation to plain RC beams, was more pronounced for beams with lower reinforcing ratios. The increase in fiber volume fraction and aspect ratio also significantly decreased the crack openings for similar bending moments. Yoo and Moon (2018) reported that the enhancement of ultimate load-carrying capacity by steel fibers was relatively minor, and the ductility index and flexural strength margin, used to guarantee a ductile failure mode, deteriorated with the inclusion of steel fibers. Lower reinforcement ratios and higher fiber volume fractions clearly led to a lower ductility index. Biolzi and Cattaneo (2017) assessed the maximum moment of RC and SFRC beams as a function of the shear span-to-depth ratio. Across all concrete grades, the lowest moment was obtained for the shear span-to-depth ratio of 2.5, while short beams with a shear span-to-depth ratio of 1.5 resulted in a higher moment due to arch action (Biolzi & Cattaneo, 2017; Cardoso et al., 2019; Yoo & Moon, 2018).

A few studies have predicted the effects of various parameters on the load-bearing capacity of RC structures as well as other structural responses (Oh et al., 2017; Rafiei et al., 2016) using new ML techniques. Feng et al. (2020) employed an ensemble ML algorithm to predict the load-carrying capacity of RC columns and found that AdaBoost regression achieved the best prediction performance with the $R^2 = 96\%$. The accuracy of the proposed explainable ML algorithm in predicting the ultimate load of RC columns with steel fibers was studied. An adaptive network-based fuzzy inference system (ANFIS) was found to be useful in predicting the load-bearing capacity of RC columns (Le & Phan, 2020). A hybrid mixture of back-propagation gradient descent method and least-squares was employed to train the ANFIS model. It was concluded that the results given by this model are mostly comparable to studies in the literature.

Other studies have predicted the ductility indices of RC beams using ML methods. Two ML and three regression-based algorithms were used to predict ductility indices under seismic loads, and a comparison of the results revealed that the nonlinear regression model achieved the highest accuracy (Dabiri et al., 2022). Considering the low amount of data, the models implemented in this study provided $R^2 = 76\%$. Various artificial neural networks (ANNs) algorithms have also been developed to predict the displacement ductility of RC columns. The findings revealed that an ANN could be particularly effective in estimating the shear strength of RC columns. The performance of complicated systems can be analyzed using an ANN. ANNs are data-driven models that resolve extremely nonlinear problems (Alam et al., 2020; Rafiei & Adeli, 2017a). They can identify the relationship between the input and output variables from preexisting patterns (historical data sets on the input and output variables of large-scale problems). This relationship is encoded through intermediate layers

of neurons, so-called hidden layers, between the input and output patterns (Guessasma et al., 2004; Zhang, Miyamori et al., 2019).

However, Koçer et al. (2019) found that the ability of an ANNs to predict the displacement ductility of RC beams was inadequate. It is assumed that the influence on the ductility index of assumptions such as yield displacement approximation has a great impact, thus considerably affecting the outcome. Due to the high importance of the ductility index of SFRC beams, and newly developed potentials of stacked ML models, the authors have developed this study. As the flexural parameters are connected, the ML model has to provide simultaneous prediction of these outputs leading to a multioutput problem. The multioutput problem increases the complexity of the problem and authors have investigated different ML model stacking strategies to understand the effects of various input parameters on the flexural performance of the SFRC beams.

2 | RESEARCH SIGNIFICANCE

Reinforced concrete (RC) beams are important structural elements of buildings and bridges. As mentioned before, a significant flexural characteristic of RC beams is their ductility ratio, which is an important predictive factor in the flexural response of RC beams. The maximum load that these beams can withstand is another important factor in their design. ML algorithms offer the potential to predict these characteristics without the complexity of numerical simulations and labor and time costs associated with experimental studies. There is a need for developing a new prediction approach for load carrying capacity (flexural strength) and ductility of SFRC beams with steel rebars. Some international codes, such as ACI 544 and MC2010, proposed equations to predict flexural capacity of SFRC beams. These models need material test results especially flexural load-deflection or load-crack mouth opening displacement (CMOD) curves to establish a tensile stress block for moment calculation. This requires additional work for predicting the load capacity of SFRC beams with steel rebars. Notwithstanding the existence of numerous experimental and numerical studies on RC beams, as well as research studies on the application of ML techniques to RC beam performance modeling, the novel algorithm called chained was developed in this study to cover the multioutput prediction of RC beams under flexure. Several ML-based models, that is, chained, parallel, and ANNs were employed, using the SciKit-Learn library in Python, integrated with hyperparameter tuning and a k-fold cross-validation (CV) method to predict the flexural features of SFRC beams. Also, there is a requirement for understanding the effect of different parameters on the flexural response of SFRC beams.

Finally, the chained algorithm was compared with the parallel and ANN models, and the advantages of this model were demonstrated. Finally, the limitations of these models in predicting the ductility index are discussed. The methodology used in this study is illustrated schematically in Figure 1. It is illustrated that after collecting and preprocessing the data, they went through the quality control section, wherein again they were reviewed, and in case either data or reference mismatch was identified, they were transferred to the preprocessing section. In this step, mismatched data, for example, unit differences, and outliers were removed or corrected, in order to unify the final data set. Reference mismatch occurs when data discrepancy is seen between the data and the corresponding reference. In this case, the references are checked again to ensure the correctness of the data associated with the reference. Also, in case of repeating data, the references are checked again to identify which data are missing from the literature that are not correlated with the corresponding literature.

The main objective of this detailed study is to give influential and impact factors, based on the developed ML algorithm, affecting the load capacity and ductility of SFRC beams, which are beneficial for engineers who consider the use of SFRC in structural members under flexure.

3 | MATERIAL AND METHODS

This section summarizes the synthetic data extraction from the literature, each ML technique used, the feature selection (FS) technique, and the model metrics used.

3.1 | Data collection

After a thorough literature review, a data set was assembled that included various input parameters that influenced the flexural performance of SFRC beams. Nine inputs and three outputs were considered in this study. A total of 193 experimental specimens on SFRC beams were identified in 19 peer-reviewed academic papers (peer-reviewed articles in ScienceDirect, Wiley, Springer, etc.). Only numerical studies validated by experimental studies were included in the data set. The following studies were used for the acquisition of the data set: Yang et al. (2020), Yoo and Yoon (2015), Yang et al. (2010), Kodur et al. (2018), Ashour and Wafa (1993), Yang et al. (2018), Yoo et al. (2017), Gümüş and Arslan (2019), Fallah-Valukolaee et al. (2022), Abbass et al. (2019), Chidambaram and Agarwal (2015), Altun and Aktaş (2013), Hawileh et al. (2018), Cardoso et al. (2019),

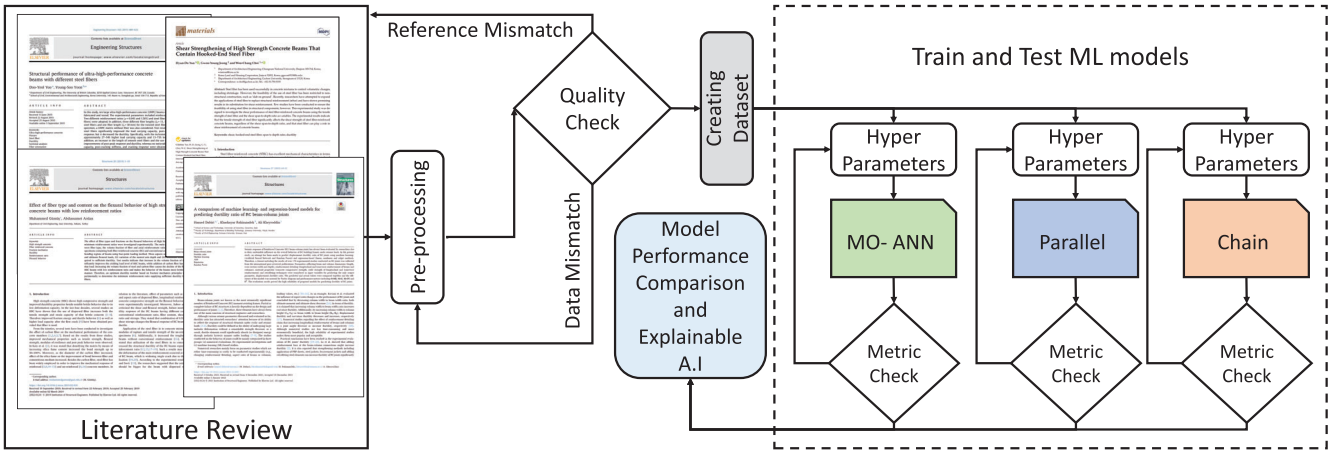

FIGURE 1 Schematic illustration of the utilized methodology.

TABLE 1 Data sample, input and output of machine learning (ML) models.

	Input										Output		
SFT[a]	$\frac{l}{d}$[b] $\left(\frac{mm}{mm}\right)$	ρ_f[c] $\left(\frac{gr}{cm^3}\right)$	v_f[d] (%)	ρ[e] $\left(\frac{mm^2}{mm^2}\right)$	w[f] (mm)	h[g] (mm)	l[h] (mm)	X_t[i] (MPa)	$\frac{A}{ED}$[j] $\left(\frac{mm}{mm}\right)$	E[k] (GPa)	PL[l] (kN)	PD[m] (mm)	$DI\left(\frac{\Delta u}{\Delta y}\right)$[n] $\left(\frac{mm}{mm}\right)$
Smooth	65.0	7.9	2.0	0.94	150	220	2500	2788	5.0	46.7	87.3	28.41	4.40
Smooth	65.0	7.5	2.0	1.96	180	270	2900	2500	4.8	45.5	233	20.82	1.71
Hooked	75.0	7.85	0.5	1.39	170	300	2940	1100	4.0	38.0	215	30.0	3.156
Hooked	64.0	7.85	0.75	1.27	160	290	1750	1200	2.5	39.3	213.7	15.0	1.81
Hooked	43.75	7.85	0.75	2.0	100	100	1000	1200	6.0	28.0	18.2	9.70	13.69
Hooked	54.5	7.8	2.0	1.34	150	230	2200	1100	3.5	24.8	117	6.83	1.34

[a]Steel fiber type.
[b]Fiber aspect ratio.
[c]Fiber density.
[d]Fiber volume fraction.
[e]Longitudinal rebar ratio.
[f]Beam width.
[g]Beam height.
[h]Beam length.
[i]Fiber tensile strength.
[j]Shear span to depth ratio.
[k]Concrete elastic modulus.
[l]Peak load.
[m]Peak Deflection (deflection at the peak load).
[n]Ductility index.

Mirdan and Saleh (2022), Arslan et al. (2017), Qissab and Salman (2018), Abbas et al. (2014), and Yun et al. (2021). Data records, such as graphs that were not expressly mentioned in the literature were meticulously compiled. The data set was split into sub–data sets for model testing (20%) and training (80%) with fivefold CV. Table 1 presents the data records utilized in this research study. The input and output parameters are listed in Table 2, along with their lowest, highest, standard deviation, mean, median, and variance values. It is worth mentioning that in the "Min" section, the value "0" represents a beam without steel fibers. Since some concrete beams did not contain fiber reinforcements, they were considered as zero value. It is worth noting that all of the data collected were under three- or four-point loading and used pin or roller support systems. Some data that showed unrealistic results, such as lower flexural strength although fibers were included, which conflicts with the findings of most studies, were neglected. Different factors were considered carefully as inclusion criteria for constructing the data set. First of all, those works published in prestigious peer-reviewed journals were taken into account. Studies with validation were

TABLE 2 Data set statistical representation.

	Input										Output		
	$\frac{l}{d}$ $\left(\frac{mm}{mm}\right)$	ρ_f $\left(\frac{gr}{cm^3}\right)$	ν_f (%)	ρ $\left(\frac{mm^2}{mm^2}\right)$	w (mm)	h (mm)	L (mm)	X_t (MPa)	$\frac{A}{ED}$ $\left(\frac{mm}{mm}\right)$	E (GPa)	PL (kN)	PD (mm)	DI $\left(\frac{\Delta u}{\Delta y}\right)$ $\left(\frac{mm}{mm}\right)$
Min	0.0	0.0	0.0	0.0	100	100	500	0.0	0.9	14.7	11.4	1.0	1.0
Max	100.0	8.0	2.0	4.87	300	320	4000	3005	6.9	53.5	438.6	119.3	26.0
Std	24.55	2.68	0.78	0.91	49.55	64.51	927.52	746.95	1.43	8.89	65.53	19.88	3.44
Mean	57.73	6.76	0.95	1.19	159.25	222.72	2026	1315.76	3.59	34.82	117.60	21.38	4.47
Median	63.0	7.8	0.75	1.13	150	220	2000	1100	3.5	37.5	100.0	15.0	3.75

highly welcomed. The conclusion section of each work was checked and in case the corresponding work has a firm and clear conclusion, that work was considered to collect the data. Also, those works with legitimate definitions of ultimate displacement were taken into account. The concentration of this study is merely on the steel fibers. Finally works in the English language were considered. Figure 2 presents the histogram of input and output data sets.

The ductility index of RC beams has been determined by dividing the ultimate deflection by yield deflection. The definition of ultimate deflection varied by researchers. On the descending branch of the load–deflection curve, 80% (Shin et al., 1989) and 90% (Manharawy et al., 2022) of the peak load have been used to read Δu (ultimate displacement), while in several studies Δu is read at the point at which the load–deflection curve begins to decline sharply (Bernardo & Lopes, 2004). In this research, the ultimate deflection has been selected in 80% on the descending branch of the load–deflection curve. It should be noted that samples with behavioral anomalies, reportedly attributed to fiber–concrete mixing problems, were ignored. The aim of this study was to develop ML models for multioutput problems to predict the maximum tolerable load, peak deflection, and ductility index. Hence, the framework comprised data on RC beams with steel fibers subjected to flexural loading.

3.2 | FS techniques

In general, using all available features of the data set can increase model complexity, diminish interpretability, and lead to overfitting. Before training the ML model with all the features of the data set, it is needed to perform a sanitary check over the data and perform FS to check if the specific feature is contributing to the prediction of the target feature or not. If an independent feature is too noisy and has little or no contribution to the model performance, it should be eliminated. Also, if two features are highly correlated utilizing both of them leads to multicollinearity issues (Bagherzadeh & Shafighfard, 2022). FS methods are used to eliminate redundant data set dimensions and retain only the most important features. Additionally, using the FS simplifies the model and speeds up the ML training. The contribution of a feature to the accuracy of the ML model is related to relevancy. Three main types of FS methods are typically employed: filter, wrapper, and embedded.

In order to establish how well the features correlate with the result variables, the filter technique chooses features depending on how well they perform in various statistical tests. The model is trained using the wrapper technique that just has a select few aspects. Embedded features combine the filter and wrapper methods. Both the filter and wrapper approaches were used in this study.

3.2.1 | Backward elimination method

By considering a criterion, a wrapper method uses a greedy search to identify the best subset of features that results in a model performing as well as possible. For classification algorithms, the criterion could be the accuracy, precision, $f1$-score, or some other, while in regression algorithms, the criterion could be R^2 or adjusted R^2. Beginning with all features, the backward elimination procedure eliminates one insignificant feature per cycle until the desired significance level is reached. In this study, the library "mlxtend" is used and the underlying model is a linear regression model from Scikit-learn.

3.2.2 | Random forest selection

One of the embedded methods that benefits from the advantages of filter methods and improvements in wrapper methods is the random forest (RF) method. This method uses several decision trees created by random feature extraction and data set observations. Trees are decorrelated because of this arbitrary selection of records and features (bootstrapping), which reduces the possibility of model overfitting (Smith, 2017).

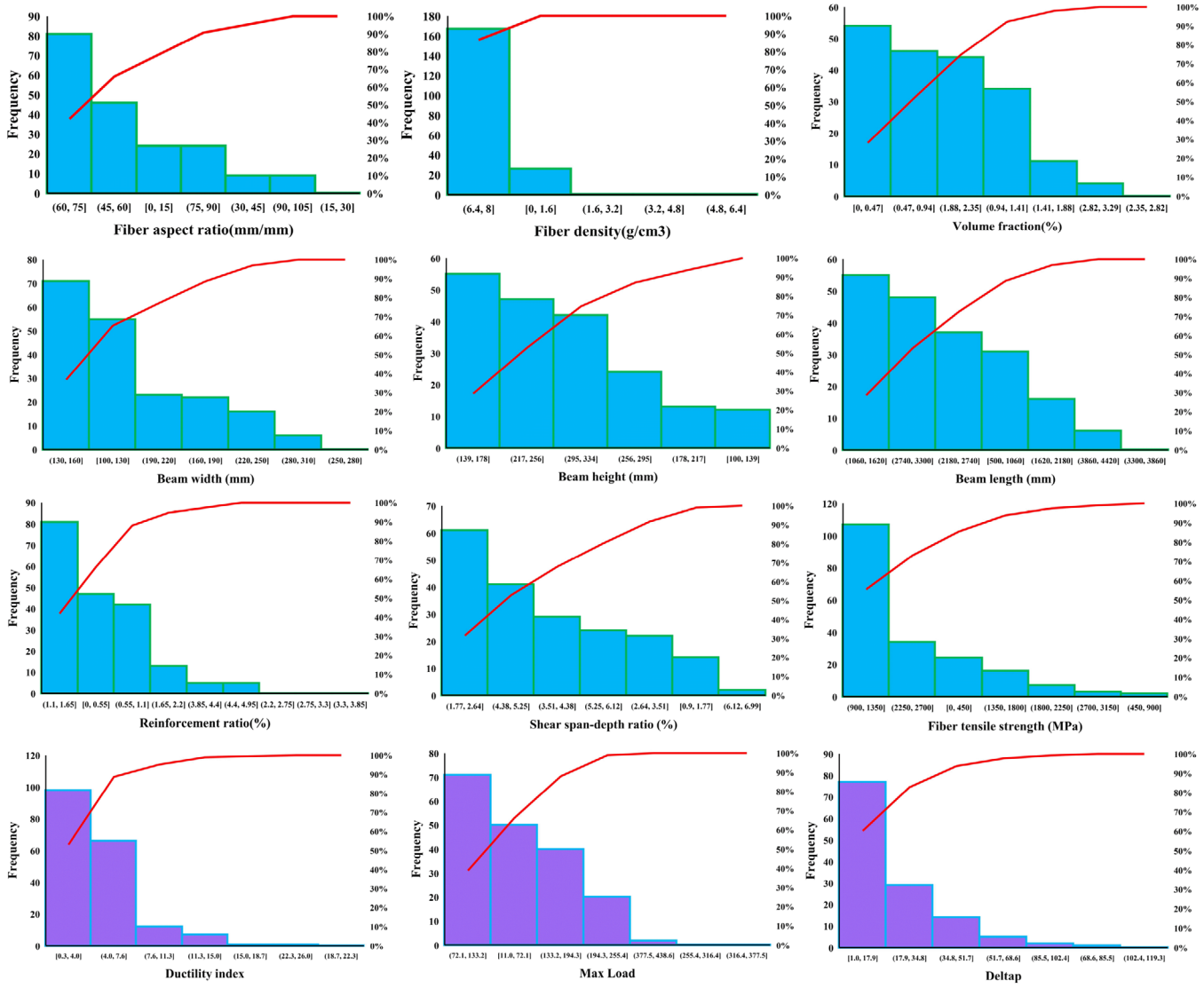

FIGURE 2 Histograms of input and target parameters.

The capacity of random forests to select crucial features for the data set is one of their powerful qualities. The utilization of OOB (Out of Bag) data is crucial in this situation. The method of determining feature importance includes determining the OOB error for each bootstrap's OOB data e_i, mixing up the data points for certain predictor variable x_i, OOB error employing permutation from the permutated bootstrap = p_i based on the mean and standard deviation of (p_i-e_i) over all trees in the population, calculating the variable significance.

$$d_i^n = p_i^n - e_i^n \tag{1}$$

According to the equation, the difference between the per-mutated feature i for model n and the not per-mutated (original OOB) feature i for model n can be used to determine the OOB error for that feature. Analyzing the evaluation indicators in greater detail led to:

$$\bar{d}_i = \frac{1}{n}\sum_{i=1}^{n} d_i^n \tag{2}$$

$$s_i^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(d_i^n - \bar{d}_i^{\,2}\right) \tag{3}$$

wherein Equation (2) refers to average OOB error over total models for feature i, and Equation (3) standard deviation for total tree models for feature i. It is worth mentioning that the more the feature is relevant to the target, the higher the importance it possesses.

Also, in this method, the Gini importance represents the contribution of each feature to the output variables as a subset of data. Gini importance or mean decrease in

impurity calculates each feature's importance as the sum over the number of splits (across all tress) that include the feature, proportionally to the number of samples it splits.

3.2.3 | Pearson correlation coefficient

The covariance of two variables divided by the product of their standard deviations is known as the Pearson correlation coefficient (Sedgwick, 2012). A product moment or the mean of the product of mean-adjusted random variables is included in the form of a definition; hence, the modifier product moment is in the name. The Pearson correlation coefficient is defined as follows:

$$r = \frac{\sum (v_i - \bar{v})(m_i - \bar{m})}{\sqrt{\sum (v_i - \bar{v})^2 \sum (m_i - \bar{m})^2}} \quad (4)$$

where the r, v_i, $\bar{v}$, m_i, $\bar{m}$ are correlation coefficients, values of the v-variable in a sample, mean of the values of v-variable, values of the m-variable in a sample, and mean of the values of m-variable, respectively (Stańczyk & Jain, 2015). Features with a pairwise correlation of more than.80 are susceptible to the overfitting problem (Kazemi et al., 2024; Shafighfard et al., 2022). The beam height had the highest correlation with the load capacity of the SFRC beams. The beam length and fiber aspect ratio also showed a correlation with the maximum tolerable load, although their influence was not noticeable. Unlike the load capacity, the peak deflection did not show any correlation with the beam height but showed a high correlation with the beam length and width. The Pearson correlation coefficient did not indicate a strong correlation between the ductility index and any of the input features. However, the shear-to-span depth ratio affected the displacement ductility index more than the other inputs. Further analysis is required to identify the redundant features and ensure correlations, as discussed in the results and discussion section (Section 4) of this paper. Figure 3 represents the heat map correlation for the input and output variables of this study. It is observed that the fiber aspect ratio has an influence on the load capacity of FRCs, though might not be effective in the ductility index of this material. Fiber density has the same role as fiber aspect ratio per the correlation map and the input parameters utilized in this study. Furthermore, the longitudinal rebar ratio affects the bearing load of the material while the ductility index was not affected by this factor. Changing the steel fiber type also did not show a high influence on the ductility index of FRCs.

3.3 | ML algorithms

In this section, three different ML algorithms including ANNs, parallel and chained algorithms that are introduced in this study for the prediction purpose are discussed in detail. The definition of each algorithm and its architecture is provided.

3.3.1 | Artificial neural network

This study utilized the form of neuron in a feed-forward ANN. A neural network with modified weight magnitudes was used to compute the approximate mathematical nonlinear relationship between the input (I_{mi}) and output (O_{ni}) layers. The inputs were multiplied by these weights (w_{ij}) and the results for each layer of neurons were combined with the biases (B_i), wherein the final outcome was multiplied with an activation function as following:

$$O_{mi} = g\left(\sum_{i=1}^{n} w_{ij} I_{mi} + B_i\right) \quad (5)$$

The sigmoid/logistic activation function was adopted for this study.

One of the most crucial problems in developing neural network models is determining an appropriate network layout with respect to the number of hidden layers and neurons. There are no absolute rules for counting the number of hidden layers or neurons (Tiryaki et al., 2014). The ideal architecture of the ANNs model in this study was determined through experiments with different numbers using the PyTorch software. The typical architecture of the ANNs model used in this study is shown in Figure 4. AutoKeras was utilized to carry out a thorough grid search. The proposed models were reproduced using PyTorch and tested using a fivefold CV and finally, the highest performance was executed from a model with the shape of (10-64-128-256-512-512-1024) having rectified linear unit (ReLU) as the activation function. The trial grid search is shown in Figure 5 and it is observed that increasing the number of layers and neurons does not increase the model performance and the selected model has reached the maximum possible accuracy.

3.3.2 | Gradient boosting machine (GBM)

A GBM is a nonlinear regression algorithm that is considered a useful method for integrating various principal regressors sequentially to construct a group. The performance of this algorithm is noticeably better than other basic regressors (Chun et al., 2021; Liang et al., 2022). The

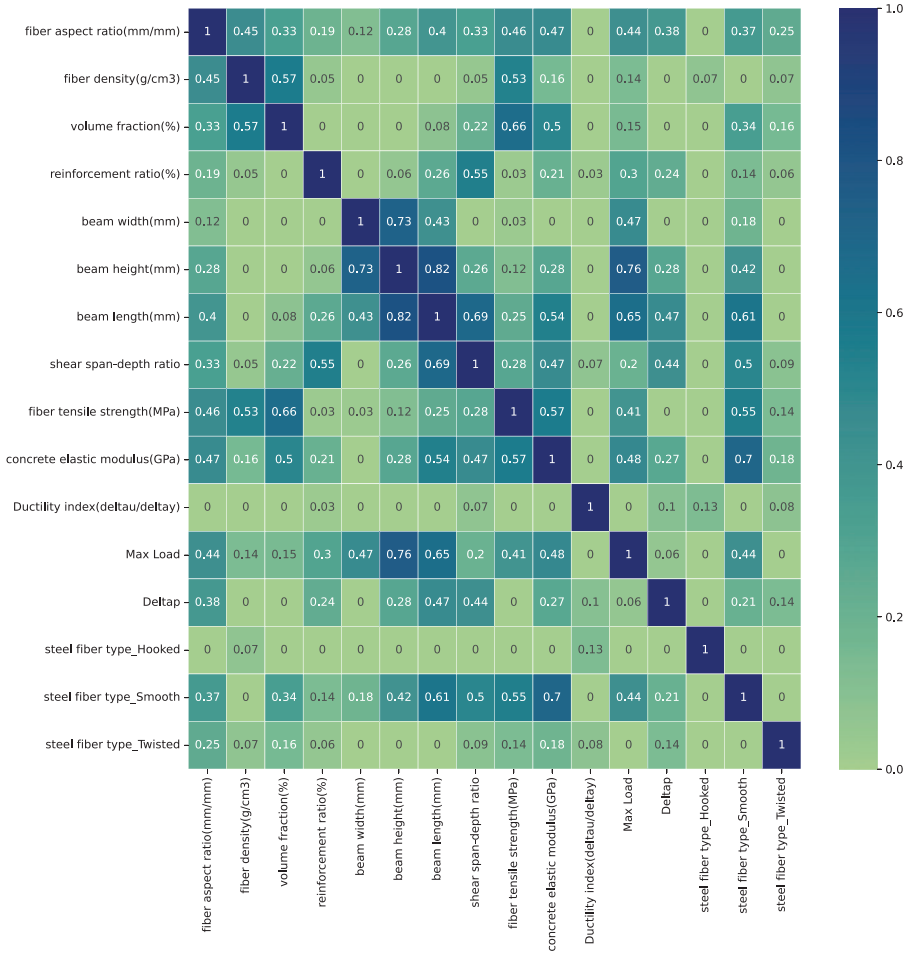

FIGURE 3 Pearson correlation heat map.

GBM includes three constituents: a weak learner, a loss function, and an additive model (Chen & Feng, 2022). The GBM links new basic learners with a negative loss function gradient as follows:

$$G(x) = \sum_{n=1}^{N} \xi_n f_m(x) \tag{6}$$

where $f_m(x)$s are weak learners or basic functions and ξ_n denotes error in the GBM. Fundamental functions are typically represented by decision trees of a fixed size. Another form of an additive algorithm can be written as follows:

$$G_m(x) = \sum_{n=1}^{N} G_{m-1}(x) + \xi_n f_m(x) \tag{7}$$

$f_m(x)$ is chosen to minimize the loss function L in each step, leading to an updated algorithm $G_{m-1}(x)$ and fit $G_{m-1}(x_i)$.

$$G_m(x) = G_{m-1}(x) + \underset{h}{\mathrm{argmin}} \sum_{n=1}^{M} L(y_i, G_{m-1}(x_i) - h(x)) \tag{8}$$

The loss function can be minimized by the GBM through gradient descent, which is a negative gradient of L in $F_{m-1}(x)$. The loss function can be differentiated as follows:

$$G_m(x) = G_{m-1}(x) + \xi_n \underset{h}{\mathrm{min}} \sum_{n=1}^{M} \nabla G L(y_i, G_{m-1}(x_i)) \tag{9}$$

3.3.3 | Chained algorithm

Chained ML models, also known as stacked or ensemble models, are suitable solutions when facing multiple output problems. Chained ML models offer a novel approach by harnessing the collective power of diverse algorithms to enhance predictive performance. These ensembles excel in improving accuracy by combining the strengths of

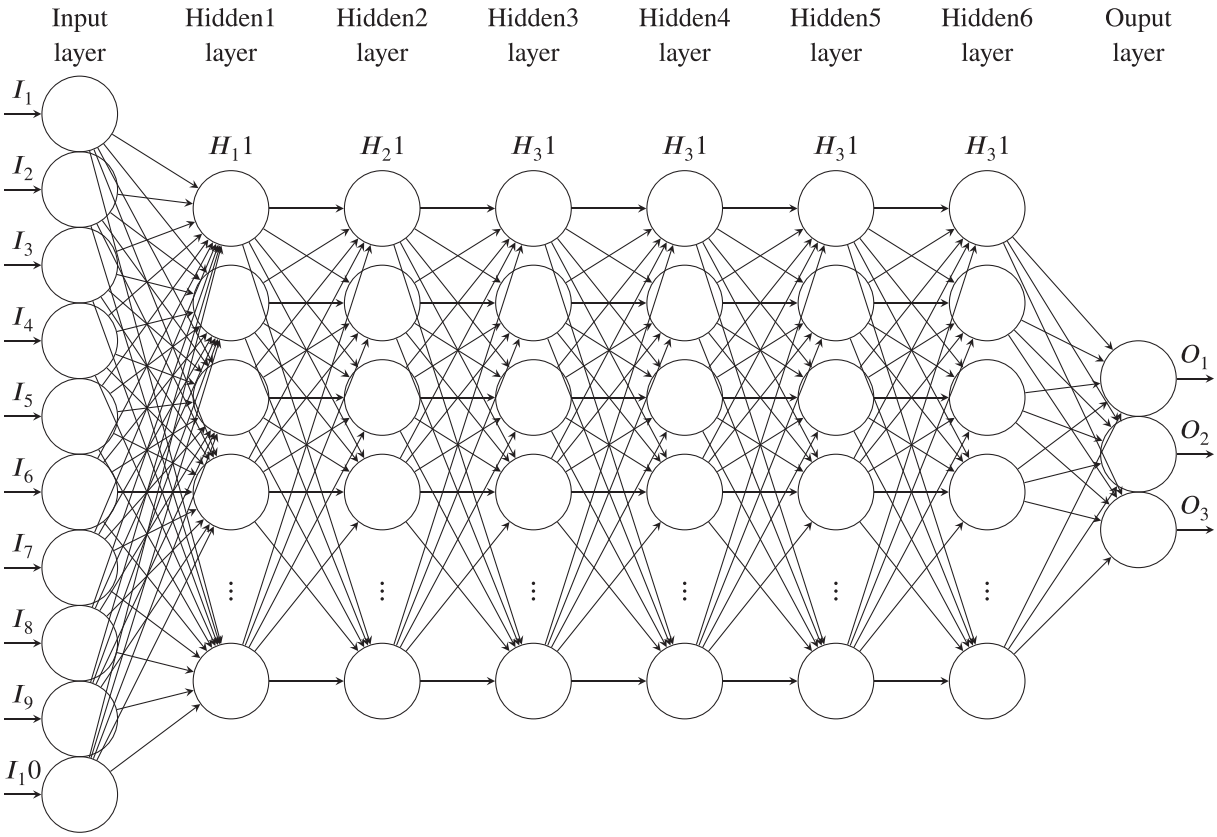

FIGURE 4 Developed artificial neural networks (ANNs) architecture with the shape of (10-64-128-256-512-512-1024-3).

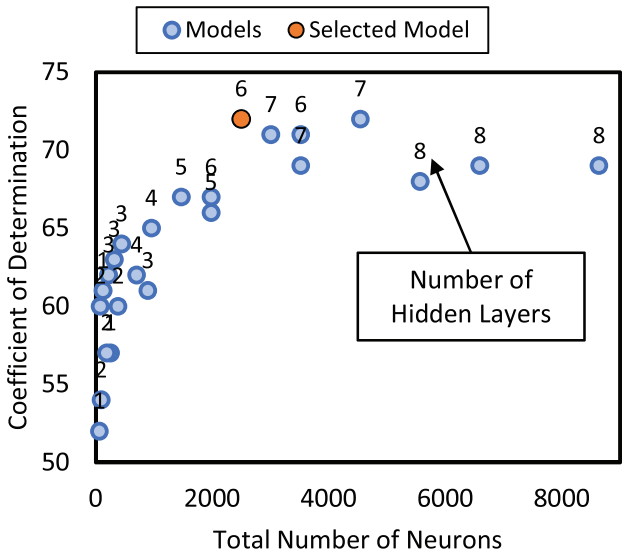

FIGURE 5 Grid search for artificial neural networks (ANNs).

individual models and mitigating weaknesses. By merging various models, they yield diverse perspectives, effectively capturing nuanced patterns within data that a single model might overlook. The chaining process enables the handling of complex problems by allowing each model to focus on different aspects or levels of intricacy, ultimately resulting in a more comprehensive understanding of the data. This approach also aids in reducing overfitting tendencies, as the ensemble learns to generalize from multiple models, thus increasing the robustness of predictions and mitigating biases present in individual algorithms.

A chained algorithm is proposed in this study to overcome the aforementioned challenges. As Figure 6a shows, the chained algorithm contains three different ML models, where each output feature is associated with one algorithm that is trained using the training data set, that is, real values. Independent features (X) and one target feature (Y1) were used to train the initial ML model (GBM1). The second model (GBM2) used Y2 and (X+Y1) for training, and the last model used Y3 and X+Y1+Y2 for training. Therefore, more data were available for GBM2 and GBM3. Unlike simple models, this chained model takes into account the relationship between various target variables, which enhances the prediction accuracy (Adibimanesh et al., 2023).

GBM1 took the independent input features (X) to predict the output ($\hat{Y}_1$) from this ML model. Subsequently, ($\hat{Y}_1$) were concatenated with X, and all were provided to GBM2. Finally, the last output feature ($\hat{Y}_3$) in the

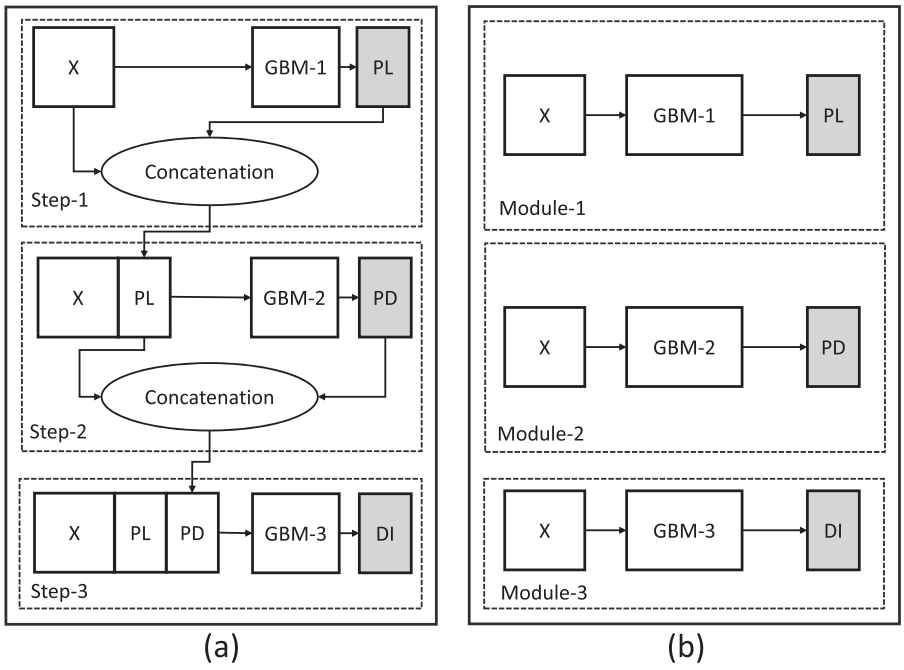

FIGURE 6 Schematic view: (a) chained and (b) parallel.

chained algorithm was predicted using GBM3, which took ($X+\hat{Y}_1+\hat{Y}_2$) as input. A schematic view of the chained model is shown in Figure 6a.

3.3.4 | Parallel structure

One of the simplest methods for constructing a multioutput regression model is to split the problem into several single-output models. This method is known as the parallel (direct) method. One instance of the given model (i.e., RF, GBM, or ANNs), as illustrated in Figure 6b, considers all the independent variables (X). One output (e.g., Y1, Y2, or Y3) is then predicted for each ML model. This model is considered a performance baseline, regardless of its simple structure. It should be noted that the correlation between dependent parameters is neglected in this method. Several types of ML algorithms, such as RF, GBM, and ANNs, were employed. The GBM was chosen for all separate ML models (for each output), as shown in Table 3, because the GBM had the highest R^2 (greater than that of ANNs and RF). A schematic illustration of the parallel model is shown in Figure 6b.

3.4 | Model metrics

The effectiveness of ML models can be evaluated using several widely used techniques. These techniques, referred to as model metrics, demonstrate the predictive power of the model in various scenarios (e.g., on the test data set; Bagherzadeh, Shafighfard, et al., 2023). The coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) were calculated.

3.5 | Shapley additive exPlanations

SHAP is a recently developed method based on game theory used for analyzing model results to better understand the performance of the model prediction. The importance of various features in the models can be assessed using SHAP values. The change in error in a particular disturbance prediction serves as the basis for quantifying the feature importance (Lundberg & Lee, 2017). It is a visualization tool that can be used to make an ML model more explainable by visualizing its output. By calculating each feature's contribution to the prediction, SHAP aims to explain the prediction of an instance of data. Shapley values are calculated using the SHAP explanation method using coalitional game theory. Players in a coalition are the feature values of a data instance. Shapley values show how to equally distribute the prediction among the features. SHAP is formulated as follows:

$$\varphi_k(\mathrm{val}) = \sum_{s \subseteq N \setminus \{i\}} \frac{|s|!\,(n-|s|-1)!}{n!} (\mathrm{val}(s \cup \{i\}) - \mathrm{val}(s)) \tag{10}$$

TABLE 3 Individual model metrics (fivefold cross-validation [CV] average).

		Train			Test		
Model type	**Target**	**MAE**	**RMSE**	R^2	**MAE**	**RMSE**	R^2
Random forest	PD (mm)	.011	.07	.97	.041	.014	.83
	DI $\left(\frac{\text{mm}}{\text{mm}}\right)$	.021	.05	.94	.062	0.012	.41
	PL (kN)	.013	.02	.98	.021	.004	.94
Gradient boosting	PD (mm)	.021	.005	.96	.042	.007	.78
	DI $\left(\frac{\text{mm}}{\text{mm}}\right)$	.028	.007	.88	.081	.008	.47
	PL (kN)	.010	.001	.99	.021	.004	.97
Neural network	PD (mm)	.021	.007	.92	.058	.021	.59
	DI $\left(\frac{\text{mm}}{\text{mm}}\right)$	.032	.005	.81	.078	.023	.35
	PL (kN)	.010	.004	.97	.023	.005	.93

Abbreviations: MAE, mean absolute error; RMSE, root mean square error.

where val is the feature importance to the algorithm target while $\varphi_k(val)$ is the weighted summation of the feature contributions to the model target result throughout all feature combinations. The i is the vector of the feature values for the parts to be interpreted, n is the number of features on the data framework, and s is the subset of the model features, $\frac{s!(n-|s|-1)!}{n!}$ is the weight of $|s|$, and $val(s)$ is the predicted value of $|s|$.

The target of an algorithm was a linear summation of the standard features and SHAP of all features. It was represented as:

$$g(x) = l(x_s) = \varphi_0 + \sum_{t=1}^{n} \varphi_t x_s^i \tag{11}$$

where n indicates the number of features in the data framework, φ_0 is the standard value for the undefined features, φ_t is the SHAP value for feature t, and x_s is the vector of simplified input variables.

4 | RESULTS AND DISCUSSION

The FS results are summarized in the first subsection below. The second subsection describes a data-driven comparative study conducted to identify the optimum algorithm for predicting the target variables. Finally, a parametric study conducted to understand the effect of each parameter on the outputs is described.

4.1 | Feature selection

To understand the effect of each feature and remove trivial features, FS results were obtained while considering different model metrics. Pearson correlation 3.2.3, backward elimination 3.2.1, and RF selection 3.2.2 were considered in this study. The initial model metric was the RMSE, which was adopted in the backward elimination graph. During the backward elimination, the feature with the least contribution to the performance of a linear regression model in a fivefold CV framework was dropped. Thus, the most important features are listed from left to right. As the target features are flexural parameters of SFRC, there is an underlying connection between them, available in the data, so the order of feature importance is the same for all the target features and only the amount of contribution is different.

FS should be performed using another method in addition to the backward technique. The RF method is more complex than previous algorithms. A lower Gini importance value indicates a smaller feature contribution to the target variables. However, it is still not perfect and requires improvement because it does not necessarily consider all correlations in the model.

Figure 7 illustrates the RF FS results. According to Figure 7, the steel fiber type and beam height do not contribute to the training of the data set for the prediction of the maximum load. The remaining factors were significant for training purposes, as indicated by the RF algorithm. In addition, the RF algorithm suggests that the steel fiber type and fiber density may be additional factors in the prediction of peak deflection. Therefore, they can be considered redundant features that may not contribute to the target value. Furthermore, important information regarding the input features for predicting the deflection ductility index was obtained. The longitudinal rebar ratio, fiber aspect ratio, and relative beam geometries could affect the model performance and should be considered to obtain a reasonable level of accuracy, while the steel fiber type could be removed from the subsampled data.

Considering these models and the Pearson correlation map, the steel fiber type and concrete elastic modulus were considered as possible redundant features. Trial-and-error

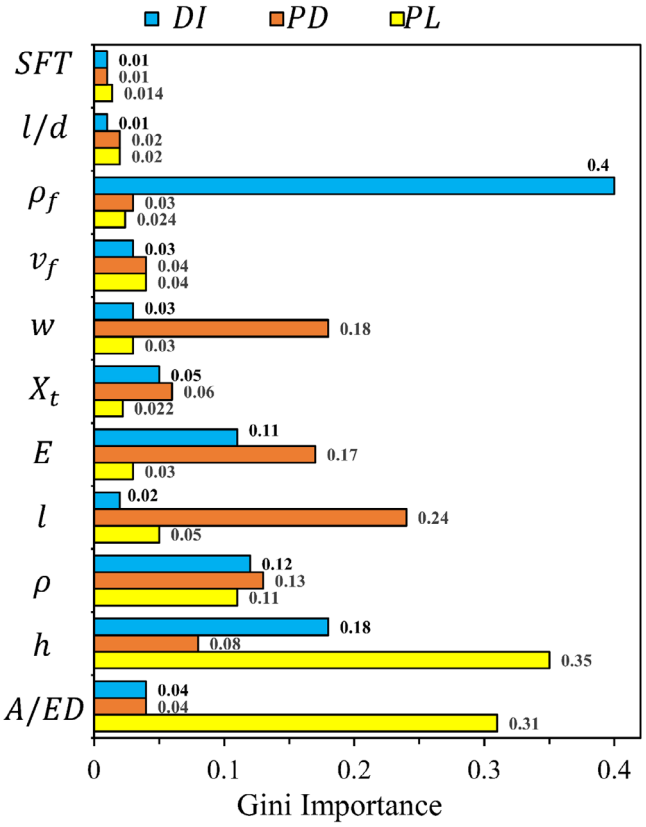

FIGURE 7 Random forest feature selection result.

evaluations of the main developed models (i.e., parallel, ANNs, and chained) were conducted to confirm the redundancy of these features. The overall prediction accuracy was increased by removing the steel fiber type and concrete elastic modulus from the training algorithms. Thus, these were not considered in the final training of the data set for predicting the flexural behavior of the RC beams. FS results are extremely sensitive to the size, distribution, and complexity of the collected data set (from different experiments in literature). Thus, those results shall not be interpreted to describe the physical behavior of the sample. In this study, FS is used only to identify the redundant features considering the given data set.

4.2 | Model evaluation

As previously discussed, an ML model should be developed for each output using a parallel approach. To achieve this, several well-known ML models were trained and tuned for each specific output. Table 3 provides details of these models and their evaluation results for a given output using a fivefold CV testing method. Based on the metric results shown in Table 3, the GBM was selected for use in the parallel model. Therefore, for each output (i.e., peak load, peak deflection, and ductility index), a separate GBM model was developed with no interconnection.

TABLE 4 Model metrics (mean of fivefold cross-validation).

Model type	Train			Test			Inference time
	MAE	RMSE	R^2	MAE	RMSE	R^2	*Log*(*s*)
Parallel	0.021	0.036	0.92	0.04	0.007	0.72	−2.2
Chained	**0.014**	**0.003**	**0.94**	**0.03**	**0.004**	**0.87**	**−2.1**
MO-ANNs	0.002	0.0008	0.98	0.03	0.008	0.75	−1.0

Abbreviations: MAE, mean absolute error; RMSE, root mean square error; MO-ANNs, multioutput artificial neural networks.

Table 3 shows that predicting the ductility index is the most challenging task since the R^2 values are less than those two other targets. In addition, the R^2 values of ductility index in all ML models are less than 50%, and this means the results are not reliable. While the results of peak load and peak deflection with R^2 values of 93% and 83% have acceptable accuracy, respectively. Thus, to create a combination of models in a chained structure, the ductility index should be at the final stage such that the prediction results from the other two outputs (i.e., peak deflection and peak load) can help the final model improve the prediction of the ductility index. The chained model initially predicts the maximum load and then concatenates the prediction results with the inputs. Thereafter, it utilized these data to predict the peak deflection. Next, the peak deflection predictions are added to the inputs, and finally, the ductility index is predicted.

Interestingly, none of these models predicted the ductility index accurately. There are numerous factors that may affect the prediction performance for deflection ductility. One factor is that the ultimate deflection was not defined in various studies. As mentioned earlier, using 80% and 90% of the peak load and a sharp reduction in the peak load as criteria for defining the ultimate deflection can lead to inaccuracies in the model performance since different type of data points are involved. A second factor is that yield points given in different research studies have been read at various points or calculated based on various techniques, such as strain and/or energy-based calculation. In addition, reading these outputs may result in errors because they are not straightforward. These are among the reasons why a parallel model may not accurately predict the ductility index. To overcome this issue, in this research, 80% of the peak load in the descending part of load–deflection curve has been selected as corresponding ultimate deflection, and the data set has been modified regarding this assumption. The results show that using this definition can reduce the error and improve the accuracy of predictions.

The quantitative results of these three modeling approaches are presented in Table 4. The fivefold CV

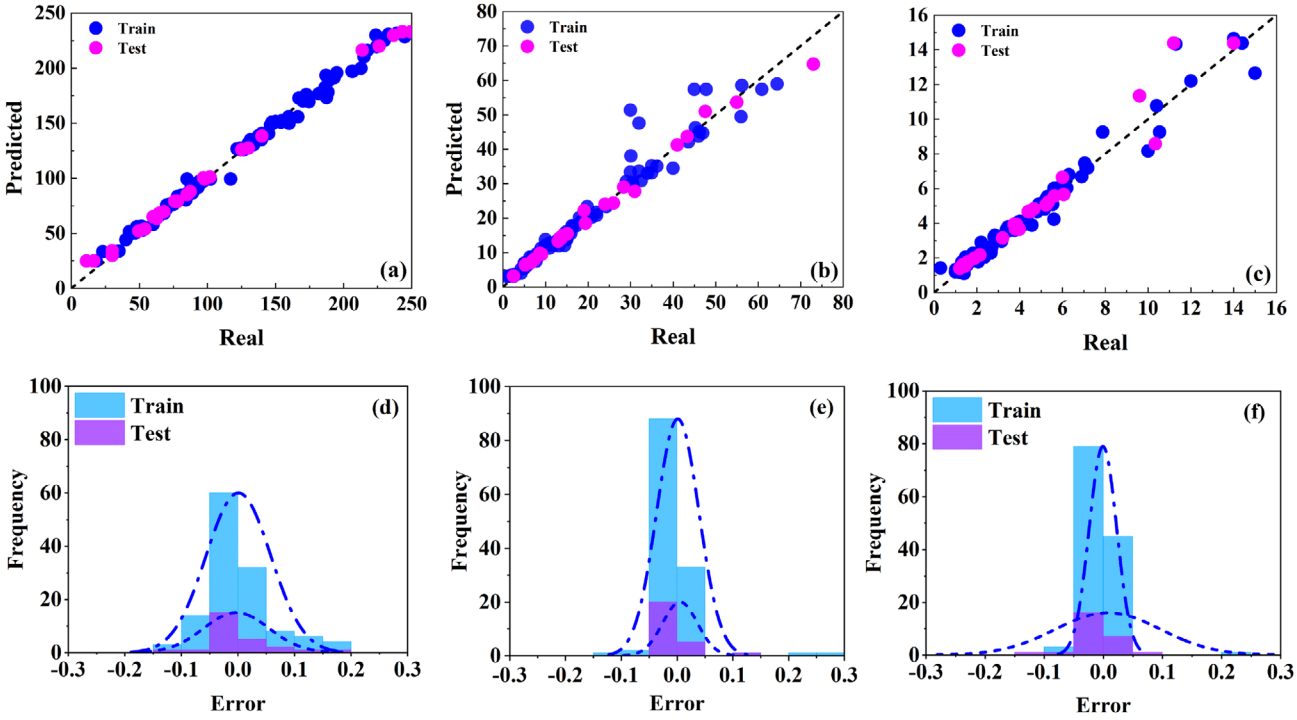

FIGURE 8 Prediction reality of (a) peak load, (b) peak deflection, and (c) ductility index, and error histogram of (d) peak load, (e) peak deflection, and (f) ductility index.

metrics of the chained model reflect the superior performance of this algorithm. As the standardization step is performed, error metrics (in Table 4) are not affected by the magnitude of the target variables. However, averaging the errors of three targets of different magnitudes with different ranges might prevent reflecting the real overall errors of each model. The results confirm that the chained model with accuracy of 94% in training and 87% for testing has the highest accuracy of prediction among two other ML models. In addition, the MO-ANNs can predict the training and testing data points with accuracy of 98% and 75%, respectively, which is higher than the parallel model. Comparing the error metrics can be a good tool to find out about the reliability of the predictions. As it is shown, the chained model has the lowest error values in the testing dataset.

4.3 | Comparative study

An in-depth investigation of the chained algorithm and its performance was conducted. Figure 8 shows the error histogram for each specific target feature. The natural distribution of the error histogram indicates that the model is not biased and has sufficient complexity. Plots of the real values versus the predicted values for each output feature are shown in Figure 8. The training and testing sets are indicated by the blue and pink dots, respectively. A greater number of points clustered near the (45°) line indicates a higher model accuracy. Thus, it is evident that the model accuracy for load capacity is higher than that for the ductility index. In particular, the large output values resulted in large deviations from the experimental values. The chained model underestimated the peak deflection and ductility index, specifically for large target values. However, this is not the case for the maximum load capacity. Furthermore, underestimation of the ductility index was not noticeable. According to scatter plots, it is clear that the accuracy of the chained model for predicting target values is promising. Figure 9 illustrates the predicted residual values corresponding to real values of peak load, peak deflection, and ductility index. Results demonstrate that the predicted residual values are limited between −0.2 and +0.2 for estimating peak load, while for peak deflection and ductility index, the predicted values are limited to −0.05 to +0.05. Moreover, it can be seen that in peak deflection lower than 60 and ductility index lower than 8, the predicted residual values are approximately 0.0, which confirms the higher prediction accuracy of the chained model. Furthermore, for values upper than 60 for peak deflection and 8 for ductility index, it is expected to have greater residual values due to lack of data set.

For a better understanding of the model error occurrences and difficulties, the error versus ductility index is shown in Figure 10. In this figure, the radius of the circle

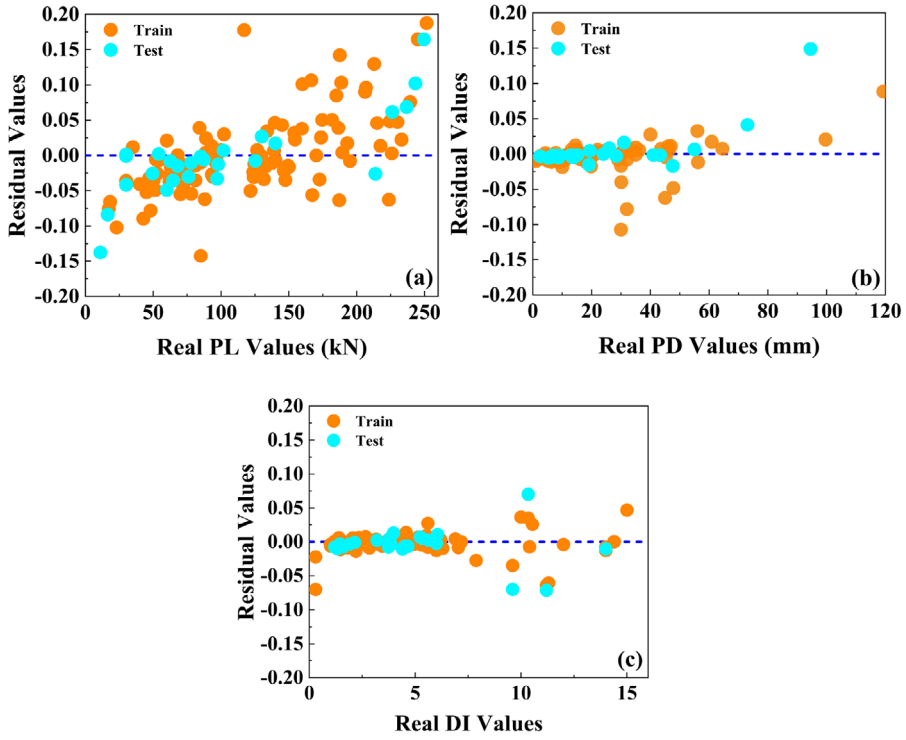

FIGURE 9 Scatter plot of the predicted residual values versus real values of (a) peak load, (b) peak deflection, and (c) ductility index.

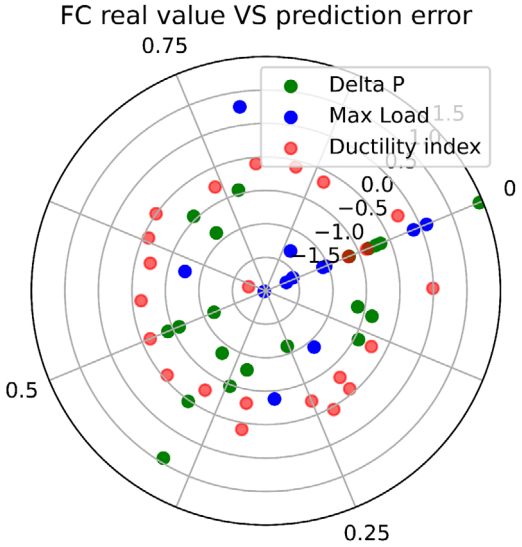

FIGURE 10 Radar diagram of errors versus real target values normalized values.

represents the error values, and the angle is the normalized target value. The deviation of the error, along with the increase in the maximum load, was visible, although the ductility index error and delta p range varied from beginning to end. This is a good indication that more data are needed from a larger range of input features to increase the model accuracy for these two features. The error was calculated based on $(y - \hat{y})$. Underestimation is represented by a positive value, and overestimation is represented by a negative value.

ML inference time is the process of running data points into an ML model to calculate a target variable (e.g., a single numerical score). A computer system with an Intel core-i5 processor running at 2.8 GHz with 8.0 GB RAM was used to obtain the inference time for each model. Table 4 shows the inference time for each algorithm, calculated as the time required to process the array of the test data set (single-batch input). Both the chained and parallel algorithms outperformed the ANNs model in terms of the time required to achieve the target variables. The parallel model provides results slightly sooner than the chained model because of its simplicity. The chained model produced a prediction 27 times faster than the ANNs algorithm.

After assessing the prediction performance of the chained, parallel, and ANN models, a Taylor diagram (Figure 11) was used to analyze the accuracy of the algorithms (Taylor, 2001). This type of diagram is widely employed to evaluate the predictive performance of the ML model in SFRC and has the highest degree of correlation with the experimental data benchmark point. It is worth mentioning that the x- and y-axis is the standard deviation of the output variable. The standard deviation of the simulated pattern is proportional to the radial distance from the origin (grey contours). Also, the narrow black arcs represent the Pearson correlation coefficient, gauging similarity in pattern between the simulated and observed

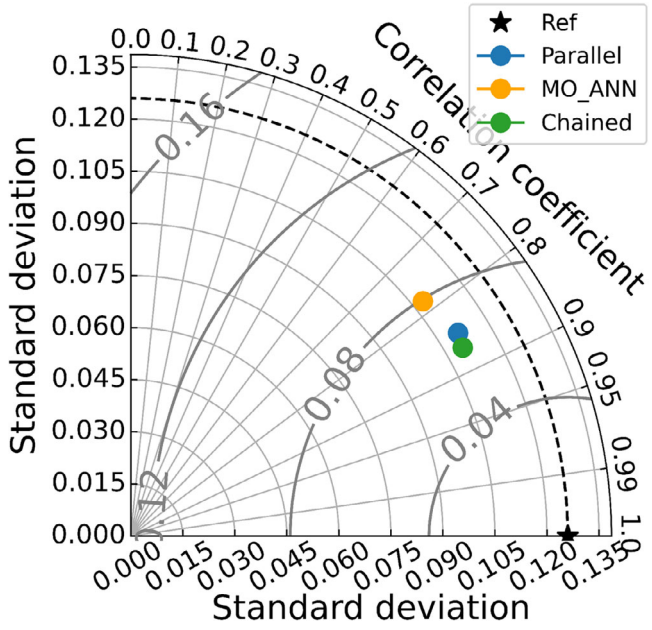

FIGURE 11 Taylor diagram of models (averaged for all outputs on the first fold).

fields. The green curves represent the centered RMSE in the simulated field, which is inversely proportional to the separation from the point on the *x*-axis designated as Figure 11. For example, the MO-ANNs shown by yellow color has a lower correlation coefficient (i.e., around 0.78) with respect to other models with a correlation coefficient of approximately 0.85 and 0.87 for parallel and chained models, respectively. Also, RMSE is higher for MO_ANNs than other techniques.

As discussed earlier in the "Material and Methods" section (Section 3), the underlying ML models between parallel and chained are the same, the main difference is that the chained algorithm utilizes the prediction of one target variable for predicting the other one. As is shown in Figure 11, the multioutput ANNs model has a higher standard deviation error and lower correlation with the target values compared to the parallel and chained models. This can be an indication that the ANNs model was able to establish a link between output neurons, but in a chained framework, this connection is established better leading to higher performance. There are several reasons for these observed results, including the inaccuracy of other techniques, experimental errors, incorrect input data when conducting the test for different records, and different definitions of some output features (e.g., ultimate deflection), which lead to inconsistencies in the data set. There are several limitations to this approach. First, the data set, which is derived from other studies relies on the accuracy of the measurements of other research groups. Second, the chained model is composed of several steps, which makes it more difficult to track the reason behind predictions record by record. Lastly, the overall chain performance is dependent on the performance of each layer and if one of the submodels is not working well, the overall performance will drop relatively.

The model metric results for all three output features for each algorithm were averaged and included in the Taylor diagram. In this figure, three models were tested against the first fold of the fivefold CV. The parallel and chained models exhibited considerably better performance than the MO-ANNs model. The overall performance of the chained model on the fivefold CV was noticeably stronger.

Figure 12 shows the test set target features and their corresponding prediction. On the *X*-axis, the record number does not represent the time and the plot is drawn for absolute value illustration purposes. The absolute errors of prediction of peak load and peak deflection are comparatively lower than ductility index. It can be also observed that the model performance decreases while reaching outliers of data. Meanwhile, results show that the predicted values are aligned with real values in most of record numbers, and the peak load and peak deflection have better fitted values.

4.4 | Sensitivity analysis

The influence of alterations in the individual feature values on the model target was further investigated by producing a beeswarm plot using SHAP, as shown in Figure 13. The *y*-axis indicates the input parameters in ascending order of importance. The blue (low) and red (high) color dots are based on the values of the input parameters. The amount of data distributed at a specific location is indicated by the density of the data. Figure 13a–c shows the global SHAP interpretation of the peak load, peak deflection, and ductility index, respectively. Figure 13a shows that the most significant feature is the beam length, which has a stronger and more positive influence on the peak deflection as the beam length increases. The next most significant feature is the fiber tensile strength, which has a direct effect on the peak deflection, according to most of the data. However, the fiber tensile strength has an inverse influence on the target value. As the value of this parameter decreases, the peak deflection increases, which is consistent with experimental results reported in the literature (Koksal et al., 2012). The shear span-to-depth ratio was directly related to the peak deflection; At higher values, the peak deflection tended to be higher. The longitudinal rebar ratio did not exhibit any clear correlation with peak deflection. However, for some data points, the peak deflection decreased as the longitudinal rebar ratio increased. The fiber aspect ratio and volume fraction also had an inverse effect on the

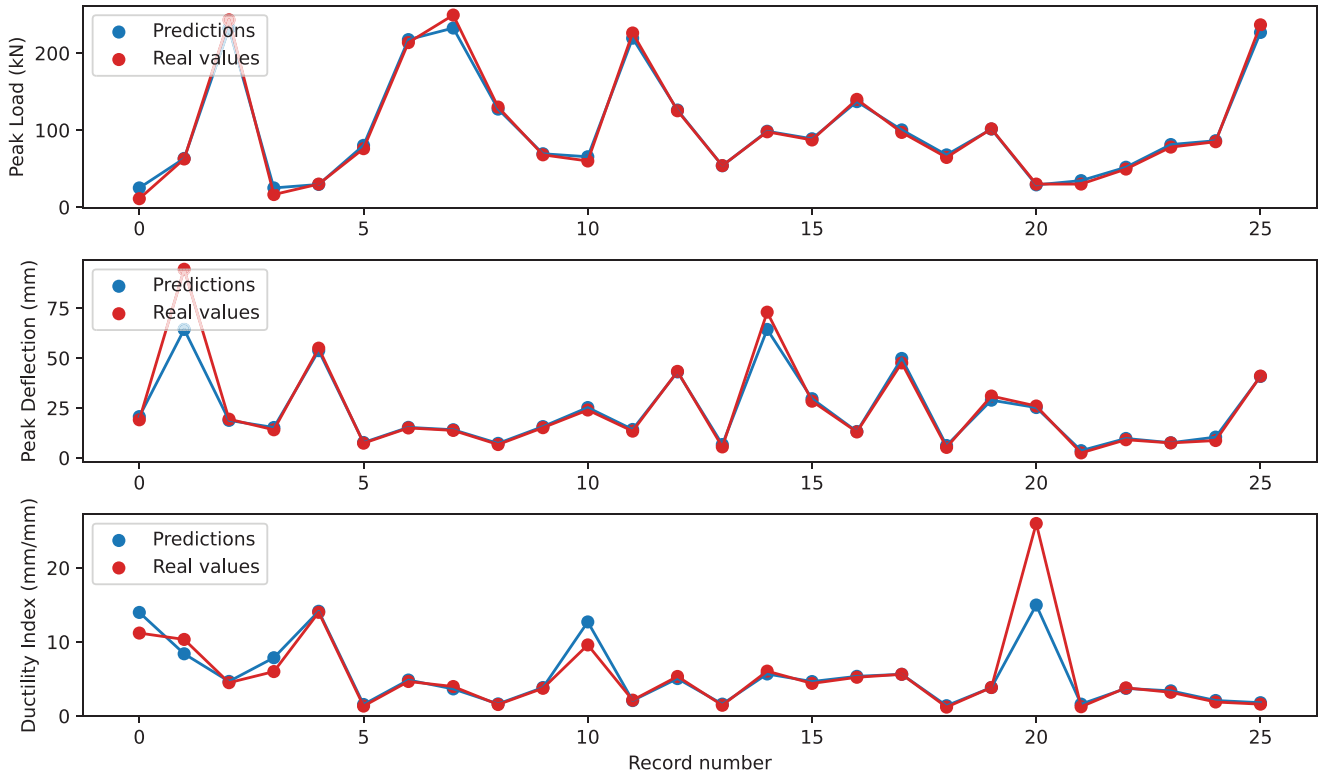

FIGURE 12 Absolute values of test set.

peak deflection. The beam width and height had the least influence on the target value. Figure 13b shows that the longitudinal rebar ratio affects the load capacity more than the other features, and as it increases, the peak load also increases. The longitudinal rebar ratio and beam length are the next most influential features: Both affect the peak load positively and directly. However, the volume fraction has a mostly negative impact on the peak load. The fiber tensile strength, shear span-to-depth ratio, and other factors have little to no effect on the load capacity, as illustrated by the SHAP diagram. Figure 13c shows that the beam height has the greatest effect on ductility index of the SFRC beams. As the beam height increases, the ductility index values of the tested beams decrease. The same effect of the beam width was observed for the ductility index for the majority of data points in the data framework. Although the effect of the longitudinal rebar ratio was not greater than that of the beam width, this factor should be considered carefully when designing SFRC beams for increasing flexural performance. The shear span-to-depth ratio exhibited a relatively positive correlation with the ductility index. The effect of the fiber aspect ratio on the ductility index of the RC beams was unclear, although it appeared to be negative. The fiber aspect ratio and fiber tensile strength were both directly correlated with ductility index. However, their effects were weaker than those of the other factors. More data are required to understand the effects of these factors on ductility index of SFRC beams. In addition, based on the SHAP value, the beam length and fiber volume fraction seem to be less influential factors for the flexural behavior of SFRC beams than the fiber aspect ratio.

It is also observed from Figure 13d–f that ductility index predictions as the final target feature of the multioutput chained model is highly impacted by the result of predictions of prior steps (i.e., peak load and peak deflection predictions), and this is an indication that the chained model is well utilizing the interconnection among the multiple target outputs. The alteration of the SHAP value for the ductility index of the RC beams with respect to the parameters related to the steel fiber is shown in Figure 14. The fiber aspect ratios upper than 75 have positive effects on DI, while the values lower than 55 has negative effects on ductility index (Figure 14a). Figure 14b shows that as the steel fiber tensile strength increases, its effect on the target value also decreases. However, lower values of the longitudinal rebar ratio indicate their potential for predicting the target variable. Longitudinal rebar ratio less than 2% have larger SHAP values than ratios higher than 2%, as shown in Figure 14c. Thus, lower longitudinal rebar ratios have a significant influence on the target variable. The effect of the fiber volume fraction on ductility index is illustrated in Figure 14d. A lower fiber volume fraction (e.g., lower than 1) has a negative effect on the ductility indices of the RC beams, while fiber volume fractions between 1 and 2 have

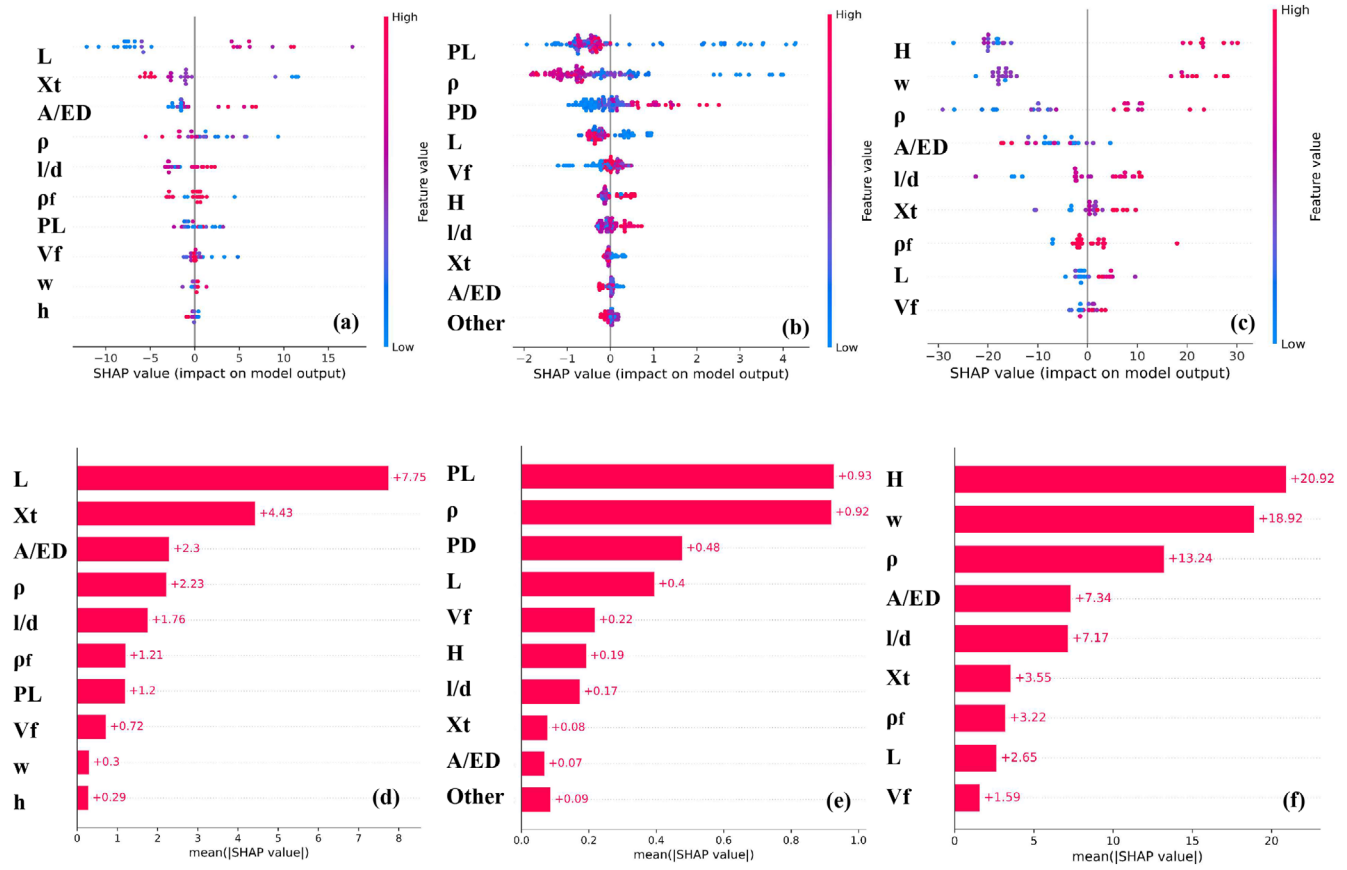

FIGURE 13 SHapley Additive exPlanations (SHAP) global interpretation on (a) peak deflection, (b) ductility index, (c) peak load, (d) peak deflection, (e) ductility index, and (f) peak load.

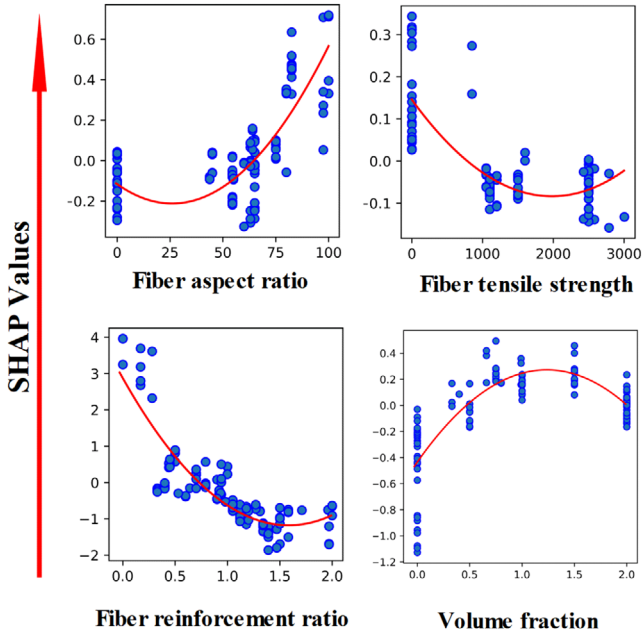

FIGURE 14 SHapley Additive exPlanations (SHAP) values for fiber characteristics of (a) fiber aspect ratio, (b) fiber tensile strength, (c) longitudinal rebar ratio, and (d) volume fraction.

a positive effect on the ductility indices. Overall, the tensile strength and aspect ratio of the steel fibers have a direct effect on the prediction of the target value, whereas the longitudinal rebar ratio and fiber volume fraction are inversely related to the output value. The SHAP results are extremely sensitive to the size, distribution, and complexity of the collected data set (from different experiments in the literature), as well as, the accuracy of developed ML models. Thus, those results shall not be interpreted to describe the physical behavior of the sample. In this study, SHAP is used only to identify the relation between input features and their effect on the accuracy of the prediction. Meanwhile, the descriptions can be used as preliminary behavior assessment of effects of input features on the target values.

5 | CONCLUSIONS

Accurately predicting the peak load, peak deflection, and ductility index of SFRC beams under bending represents a complex multioutput prediction task, navigated through the application of a diverse range of ML models. Within this study, a comprehensive investigation has been

conducted encompassing ANNs, parallel, and chained models to address this intricate challenge. This rigorous exploration involved the assessment of these various ML models against a comprehensive database sourced from pertinent literature. Leveraging a rigorous fivefold CV methodology and appraising performance across four key metrics, this investigation rigorously scrutinized and compared the efficacy of the suggested ML models. A comparative analysis was conducted to determine the most accurate predictive approach. The comparisons demonstrated the superiority of the chained algorithm in terms of predictive performance and overall prediction accuracy. The following conclusions were drawn from this study:

- The proposed chained algorithm is a stacking ML model that uses the output of one model to predict the next target feature. It predicts the flexural performance of SFRC beams well, as indicated by its values of $R^2 = 0.87$, MAE = 0.03, and RMSE = 0.004.
- While the ANNs model demonstrates an ability to leverage interconnections among target features, its utilization of this connectivity was relatively constrained, particularly evident when compared to the proposed chained algorithm. This limitation becomes more apparent in scenarios with a limited number of data points, wherein the ANNs exhibit a reduced capacity to exploit these interconnections, leading to a comparatively lower optimization of overall averaged performance.
- As per the SHAP global diagrams, it became evident that among the various factors considered, the longitudinal rebar ratio and beam length wielded the most substantial influence on the ductility index of SFRC beams. These two parameters emerged as pivotal contributors, significantly shaping and impacting the ductility index within the context of the studied SFRC beams.
- The results of residual values confirmed that the chained model has the highest accuracy for peak deflection values lower than 60 and the ductility index values lower than 8 (i.e., the residual value in this range is approximately 0.0). Therefore, the model has the highest reliability for this range of predictions.
- The efficiency of the chained model in producing accurate predictions was notably superior in terms of speed compared to the ANNs models. This enhanced speed in delivering precise predictions signifies a considerable advantage of the chained model over the ANNs, highlighting its efficacy in achieving accuracy within a more expedited timeframe.

The results of this study contribute to improving the prediction of the flexural performance of SFRC beams without exhaustive simulations or expensive experimental tests. However, further experimental data should be obtained to improve the accuracy of ductility index prediction for more than 8 and peak deflection upper than 60. The future work includes adding experimental data from the authors further to this meta-analysis work and employing dynamic ensemble learning (Rafiei et al., 2022) and self-supervised learning for prediction purposes.

ACKNOWLEDGMENTS

The authors thank the University of Bremen for covering the fee for the publication. This work was also supported by a National Research Foundation of Korea (NRF) grant (No. 2021R1A2C4001503) funded by the Korean Government (MSIT) and the Yonsei University Research Fund of 2023 (2023-22-0134).

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no financial/personal interests that could have affected the reported work in this paper.

DATA AVAILABILITY STATEMENT

The data set and codes are available in the form of Python Jupyter notebook in https://github.com/FarzinKazemi/Chained-Machine-Learning-Model

REFERENCES

Abbas, A. A., Syed Mohsin, S. M., Cotsovos, D. M., & Ruiz-Teran, A. M. (2014). Shear behaviour of steel-fibre-reinforced concrete simply supported beams. *Proceedings of the Institution of Civil Engineers-Structures and Buildings*, *167*(9), 544–558.

Abbass, A., Abid, S., & Özakça, M. (2019). Experimental investigation on the effect of steel fibers on the flexural behavior and ductility of high-strength concrete hollow beams. *Advances in Civil Engineering*, *2019*, 1–31.

Adhikary, B. B., & Mutsuyoshi, H. (2006). Prediction of shear strength of steel fiber RC beams using neural networks. *Construction and Building Materials*, *20*(9), 801–811.

Adibimanesh, B., Polesek-Karczewska, S., Bagherzadeh, F., Szczuko, P., & Shafighfard, T. (2023). Energy consumption optimization in wastewater treatment plants: Machine learning for monitoring incineration of sewage sludge. *Sustainable Energy Technologies and Assessments*, *56*, 103040.

Al-Ahmed, A. H. A., Al-Rumaithi, A., Allawi, A. A., & El-Zohairy, A. (2022). Mesoscale analysis of fiber-reinforced concrete beams. *Engineering Structures*, *266*, 114575.

Alam, K. M. R., Siddique, N., & Adeli, H. (2020). A dynamic ensemble learning algorithm for neural networks. *Neural Computing and Applications*, *32*, 8675–8690.

Aldwaik, M., & Adeli, H. (2016). Cost optimization of reinforced concrete flat slabs of arbitrary configuration in irregular highrise building structures. *Structural and Multidisciplinary Optimization*, *54*, 151–164.

Altun, F., & Aktaş, B. (2013). Investigation of reinforced concrete beams behavior of steel fiber added lightweight concrete. *Construction and Building Materials*, *38*, 575–581.

Arslan, G., Keskin, R. S. O., & Ulusoy, S. (2017). An experimental study on the shear strength of SFRC beams without stirrups. *Journal of Theoretical and Applied Mechanics*, *55*(4), 1205–1217.

Ashour, S. A., & Wafa, F. F. (1993). Flexural behavior of high-strength fiber reinforced concrete beams. *Structural Journal*, *90*(3), 279–287.

Bagherzadeh, F., Freitag, J., Frese, U., & Wilhelms, F. (2023). Ice core micro-CT image segmentation with deep learning and Gaussian mixture model. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–11. https://doi.org/10.1109/TGRS.2023.3334867

Bagherzadeh, F., & Shafighfard, T. (2022). Ensemble machine learning approach for evaluating the material characterization of carbon nanotube-reinforced cementitious composites. *Case Studies in Construction Materials*, *17*, e01537.

Bagherzadeh, F., Shafighfard, T., Khan, R. M. A., Szczuko, P., & Mieloszyk, M. (2023). Prediction of maximum tensile stress in plain-weave composite laminates with interacting holes via stacked machine learning algorithms: A comparative study. *Mechanical Systems and Signal Processing*, *195*, 110315.

Bernardo, L., & Lopes, S. (2004). Neutral axis depth versus flexural ductility in high-strength concrete beams. *Journal of Structural Engineering*, *130*(3), 452–459.

Biolzi, L., & Cattaneo, S. (2017). Response of steel fiber reinforced high strength concrete beams: Experiments and code predictions. *Cement and Concrete Composites*, *77*, 1–13.

Cardoso, D. C., Pereira, G. B., Silva, F. A., Silva Filho, J. J., & Pereira, E. V. (2019). Influence of steel fibers on the flexural behavior of RC beams with low reinforcing ratios: Analytical and experimental investigation. *Composite Structures*, *222*, 110926.

Çelik, F., & König, M. (2022). A sigmoid-optimized encoder–decoder network for crack segmentation with copy-edit-paste transfer learning. *Computer-Aided Civil and Infrastructure Engineering*, *37*(14), 1875–1890.

Chen, S. Z., & Feng, D. C. (2022). Multifidelity approach for data-driven prediction models of structural behaviors with limited data. *Computer-Aided Civil and Infrastructure Engineering*, *37*(12), 1566–1581.

Chidambaram, R. S., & Agarwal, P. (2015). Flexural and shear behavior of geo-grid confined RC beams with steel fiber reinforced concrete. *Construction and Building materials*, *78*, 271–280.

Chun, P. j., Izumi, S., & Yamane, T. (2021). Automatic detection method of cracks from concrete surface imagery using two-step light gradient boosting machine. *Computer-Aided Civil and Infrastructure Engineering*, *36*(1), 61–72.

Dabiri, H., Rahimzadeh, K., & Kheyroddin, A. (2022). A comparison of machine learning-and regression-based models for predicting ductility ratio of RC beam-column joints. *Structures*, *37*, 69–81.

Fallah-Valukolaee, S., Hashemi, S., & Nematzadeh, M. (2022). Effect of steel fiber on flexural performance of bilayer concrete beams with steel and GFRP rebars: Experiments and predictions. *Structures*, *39*, 405–418.

Feng, D. C., Liu, Z. T., Wang, X. D., Jiang, Z. M., & Liang, S. X. (2020). Failure mode classification and bearing capacity prediction for reinforced concrete columns based on ensemble machine learning algorithm. *Advanced Engineering Informatics*, *45*, 101126.

Finckh, W., & Zilch, K. (2012). Strengthening and rehabilitation of reinforced concrete slabs with carbon-fiber reinforced polymers using a refined bond model. *Computer-Aided Civil and Infrastructure Engineering*, *27*(5), 333–346.

Gribniak, V., Kaklauskas, G., Kwan, A. K. H., Bacinskas, D., & Ulbinas, D. (2012). Deriving stress–strain relationships for steel fibre concrete in tension from tests of beams with ordinary reinforcement. *Engineering Structures*, *42*, 387–395.

Gribniak, V., & Sokolov, A. (2023). Standardized RC beam tests for modeling the fiber bridging effect in SFRC. *Construction and Building Materials*, *370*, 130652.

Guessasma, S., Montavon, G., & Coddet, C. (2004). Modeling of the APS plasma spray process using artificial neural networks: Basis, requirements and an example. *Computational Materials Science*, *29*(3), 315–333.

Gümüş, M., & Arslan, A. (2019). Effect of fiber type and content on the flexural behavior of high strength concrete beams with low reinforcement ratios. *Structures*, *20*, 1–10.

Hawileh, R., Nawaz, W., & Abdalla, J. (2018). Flexural behavior of reinforced concrete beams externally strengthened with hard-wire steel-fiber sheets. *Construction and Building Materials*, *172*, 562–573.

Jeong, J. H., & Jo, H. (2021). Deep reinforcement learning for automated design of reinforced concrete structures. *Computer-Aided Civil and Infrastructure Engineering*, *36*(12), 1508–1529.

Júnior, S. A. A., & Parvin, A. (2022). Reinforcement of new and existing reinforced concrete beams with fiber-reinforced polymer bars and sheets—A numerical analysis. *Structures*, *40*, 513–523.

Kazemi, F., & Jankowski, R. (2023). Machine learning-based prediction of seismic limit-state capacity of steel moment-resisting frames considering soil-structure interaction. *Computers & Structures*, *274*, 106886.

Kazemi, F., Shafighfard, T., & Yoo, D. Y. (2024). Data-driven modeling of mechanical properties of fiber-reinforced concrete: A critical review. *Archives of Computational Methods in Engineering*, *30*, 1–30.

Koçer, M., Öztürk, M., & Arslan, M. H. (2019). Determination of moment, shear and ductility capacities of spiral columns using an artificial neural network. *Journal of Building Engineering*, *26*, 100878.

Kodur, V., Solhmirzaei, R., Agrawal, A., Aziz, E. M., & Soroushian, P. (2018). Analysis of flexural and shear resistance of ultra high performance fiber reinforced concrete beams without stirrups. *Engineering Structures*, *174*, 873–884.

Koksal, F., Sahin, Y., & Sahin, M. (2012). Effect of steel fiber tensile strength on mechanical properties of steel fiber reinforced concretes. *Special Publication*, *289*, 1–15.

Le, T. T., & Phan, H. C. (2020). Prediction of ultimate load of rectangular CFST columns using interpretable machine learning method. *Advances in Civil Engineering*, *2020*, 1–16.

Liang, M., Chang, Z., He, S., Chen, Y., Gan, Y., Schlangen, E., & Šavija, B. (2022). Predicting early-age stress evolution in restrained concrete by thermo-chemo-mechanical model and active ensemble learning. *Computer-Aided Civil and Infrastructure Engineering*, *37*(14), 1809–1833.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*, 1–31.

Luo, H., & Paal, S. G. (2019). A locally weighted machine learning model for generalized prediction of drift capacity in seismic vulnerability assessments. *Computer-Aided Civil and Infrastructure Engineering*, *34*(11), 935–950.

Maeda, H., Kashiyama, T., Sekimoto, Y., Seto, T., & Omata, H. (2021). Generative adversarial network for road damage detection. *Computer-Aided Civil and Infrastructure Engineering*, *36*(1), 47–60.

Mangalathu, S., Hwang, S. H., & Jeon, J. S. (2020). Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Engineering Structures*, *219*, 110927.

Mangalathu, S., & Jeon, J. S. (2018). Classification of failure mode and prediction of shear strength for reinforced concrete beam-column joints using machine learning techniques. *Engineering Structures*, *160*, 85–94.

Manharawy, M. S., Mahmoud, A. A., El-Mahdy, O. O., & El-Diasity, M. H. (2022). Experimental and numerical investigation of lightweight foamed reinforced concrete deep beams with steel fibers. *Engineering Structures*, *260*, 114202.

Masuelli, M. (2013). *Fiber reinforced polymers: The technology applied for concrete repair*. BoD–Books on Demand.

Meda, A., Minelli, F., & Plizzari, G. A. (2012). Flexural behaviour of RC beams in fibre reinforced concrete. *Composites Part B: Engineering*, *43*(8), 2930–2937.

Mertol, H. C., Baran, E., & Bello, H. J. (2015). Flexural behavior of lightly and heavily reinforced steel fiber concrete beams. *Construction and Building Materials*, *98*, 185–193.

Mirdan, D., & Saleh, A. R. (2022). Flexural performance of reinforced concrete (RC) beam strengthened by UHPC layer. *Case Studies in Construction Materials*, *17*, e01655.

Mobasher, B., Yao, Y., & Soranakom, C. (2015). Analytical solutions for flexural design of hybrid steel fiber reinforced concrete beams. *Engineering Structures*, *100*, 164–177.

Oh, B. K., Kim, K. J., Kim, Y., Park, H. S., & Adeli, H. (2017). Evolutionary learning based sustainable strain sensing model for structural health monitoring of high-rise buildings. *Applied Soft Computing*, *58*, 576–585.

Pak, H., Leach, S., Yoon, S. H., & Paal, S. G. (2023). A knowledge transfer enhanced ensemble approach to predict the shear capacity of reinforced concrete deep beams without stirrups. *Computer-Aided Civil and Infrastructure Engineering*, *38*, 1520–1535.

Pereira, D. R., Piteri, M. A., Souza, A. N., Papa, J. P., & Adeli, H. (2020). FEMA: A finite element machine for fast learning. *Neural Computing and Applications*, *32*, 6393–6404.

Qissab, M. A., & Salman, M. M. (2018). Shear strength of non-prismatic steel fiber reinforced concrete beams without stirrups. *Structural Engineering and Mechanics*, *67*(4), 347–358.

Rafiei, M. H., & Adeli, H. (2017a). A new neural dynamic classification algorithm. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(12), 3074–3083.

Rafiei, M. H., & Adeli, H. (2017b). A novel machine learning-based algorithm to detect damage in high-rise building structures. *The Structural Design of Tall and Special Buildings*, *26*(18), e1400.

Rafiei, M. H., Gauthier, L. V., Adeli, H., & Takabi, D. (2022). Self-supervised learning for electroencephalography. *IEEE Transactions on Neural Networks and Learning Systems*, *57*, 1–15.

Rafiei, M. H., Khushefati, W. H., Demirboga, R., & Adeli, H. (2016). Neural network, machine learning, and evolutionary approaches for concrete material characterization. *ACI Materials Journal*, *113*(6), 781–789.

Rafiei, M. H., Khushefati, W. H., Demirboga, R., & Adeli, H. (2017). Supervised deep restricted Boltzmann machine for estimation of concrete. *ACI Materials Journal*, *114*(2), 237.

Sandeep, M. S., Tiprak, K., Kaewunruen, S., Pheinsusom, P., & Pansuk, W. (2023). Shear strength prediction of reinforced concrete beams using machine learning. *Structures*, *47*, 1196–1211.

Sedgwick, P. (2012). Pearson's correlation coefficient. *British Medical Journal*, *345*, 41–56.

Shafighfard, T., Bagherzadeh, F., Rizi, R. A., & Yoo, D. Y. (2022). Data-driven compressive strength prediction of steel fiber reinforced concrete (SFRC) subjected to elevated temperatures using stacked machine learning algorithms. *Journal of Materials Research and Technology*, *21*, 3777–3794.

Shafighfard, T., Cender, T. A., & Demir, E. (2021). Additive manufacturing of compliance optimized variable stiffness composites through short fiber alignment along curvilinear paths. *Additive Manufacturing*, *37*, 101728.

Shin, S. W., Ghosh, S. K., & Moreno, J. (1989). Flexural ductility of ultra-high-strength concrete members. *Structural Journal*, *86*(4), 394–400.

Singh, H. (2016). *Steel fiber reinforced concrete: Behavior, modelling and design*. Springer.

Smith, C. (2017). *Decision trees and random forests: A visual introduction for beginners*. Blue Windmill Media.

Solhmirzaei, R., Salehi, H., Kodur, V., & Naser, M. (2020). Machine learning framework for predicting failure mode and shear capacity of ultra high performance concrete beams. *Engineering Structures*, *224*, 111221.

Soltanzadeh, F., Barros, J. A., & Santos, R. (2015). High performance fiber reinforced concrete for the shear reinforcement: Experimental and numerical research. *Construction and Building Materials*, *77*, 94–109.

Stańczyk, U., & Jain, L. C. (2015). *Feature selection for data and pattern recognition*. Springer.

Tan, H., Hou, Z., Li, Y., & Xu, X. (2022). A flexural ductility model for UHPC beams reinforced with FRP bars. *Structures*, *45*, 773–786.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192.

Tiryaki, S., Malkoçoğlu, A., & Özşahin, Ş. (2014). Using artificial neural networks for modeling surface roughness of wood in machining process. *Construction and Building Materials*, *66*, 329–335.

Venkateshwaran, A., & Tan, K. H. (2018). Load-carrying capacity of steel fiber reinforced concrete beams at large deflections. *Structural Concrete*, *19*(3), 670–683.

Xu, H., & Humar, J. (2006). Damage detection in a girder bridge by artificial neural network technique. *Computer-Aided Civil and Infrastructure Engineering*, *21*(6), 450–464.

Yang, I. H., Joh, C., & Kim, B. S. (2010). Structural behavior of ultra high performance concrete beams subjected to bending. *Engineering Structures*, *32*(11), 3478–3487.

Yang, I. H., Joh, C., & Kim, K. C. (2018). A comparative experimental study on the flexural behavior of high-strength fiber-reinforced concrete and high-strength concrete beams. *Advances in Materials Science and Engineering*, *2018*, 49–64.

Yang, I. H., Park, J., Bui, T. Q., Kim, K. C., Joh, C., & Lee, H. (2020). An experimental study on the ductility and flexural toughness of ultrahigh-performance concrete beams subjected to bending. *Materials*, *13*(10), 2225.

Yoo, D. Y., Banthia, N., & Yoon, Y. S. (2017). Experimental and numerical study on flexural behavior of ultra-high-performance

fiber-reinforced concrete beams with low reinforcement ratios. *Canadian Journal of Civil Engineering*, *44*(1), 18–28.

Yoo, D. Y., & Moon, D. Y. (2018). Effect of steel fibers on the flexural behavior of RC beams with very low reinforcement ratios. *Construction and Building Materials*, *188*, 237–254.

Yoo, D. Y., & Yoon, Y. S. (2015). Structural performance of ultra-high-performance concrete beams with different steel fibers. *Engineering Structures*, *102*, 409–423.

Yun, H. D., Jeong, G. Y., & Choi, W. C. (2021). Shear strengthening of high strength concrete beams that contain hooked-end steel fiber. *Materials*, *15*(1), 17.

Zhang, R., Jin, L., Tian, Y., Dou, G., & Du, X. (2019). Static and dynamic mechanical properties of eco-friendly polyvinyl alcohol fiber-reinforced ultra-high-strength concrete. *Structural Concrete*, *20*(3), 1051–1063.

Zhang, Y., Miyamori, Y., Mikami, S., & Saito, T. (2019). Vibration-based structural state identification by a 1-dimensional convolutional neural network. *Computer-Aided Civil and Infrastructure Engineering*, *34*(9), 822–839.

Zhang, Y., & Yuen, K. V. (2021). Crack detection using fusion features-based broad learning system and image processing. *Computer-Aided Civil and Infrastructure Engineering*, *36*(12), 1568–1584.

How to cite this article: Shafighfard, T., Kazemi, F., Bagherzadeh, F., Mieloszyk, M., & Yoo, D.-Y. (2024). Chained machine learning model for predicting load capacity and ductility of steel fiber–reinforced concrete beams. *Computer-Aided Civil and Infrastructure Engineering*, 1–22. https://doi.org/10.1111/mice.13164