AUT Journal of Civil Engineering

AUT J. Civil Eng., 9(2) (2025) 143-158
DOI: 10.22060/ajce.2025.24055.5914

Estimation of Fiber Pullout Curve from Cement Base Matrix Using Machine Learning

Ali Hossein pour[1], Meysam Jalali[1]*, Hosein Naderpour[2,3]

[1] Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran.
[2] Faculty of Civil Engineering, Semnan University, Semnan, Iran.
[3] Department of Civil Engineering, Toronto Metropolitan University, Ontario, Canada.

ABSTRACT: Pull-out tests and corresponding curves help analyze fiber pull-out behavior, which significantly influences the tensile and flexural properties of fiber-reinforced concrete. Artificial intelligence techniques provide an efficient means to predict fiber pull-out curves. In this study, four models based on convolutional neural networks (CNN), long short-term memory networks (LSTM), and extreme gradient boosting (XGBoost) are used to predict fiber pull-out curves: CNN1D, CNN2D, LSTM, and XGBoost. A dataset of 502 experimental samples was compiled, primarily from laboratory studies conducted by one of the authors of this paper. Fiber aspect ratio, loading rate, type and quantity of cement, amount of: binder, silica fume, sand, gravel, quartz, superplasticizer, water, water-to-cement ratio, water-to-binder ratio, curing age, fiber embedded length, end-type of the fibers, the pitch of spirals, number of twists, fiber inclination angle, fiber length and diameter, and concrete compressive strength are among the used input parameters. To represent the expected curve, the model output comprises 1000 pairs of data (pull-out force vs slip). Among the tested models, XGBoost demonstrated superior performance with the lowest mean absolute error (8.39) and highest R^2 value (0.71), making it the optimal choice for predicting fiber pull-out curves. The parameters "embedded length", "number of twists", "Silica fume", and "pitch of spirals" were more important and influenced the model and prediction accuracy, as can be seen from the evaluation of the feature importance graph.

Review History:
Received: Apr. 10, 2025
Revised: Jun. 02, 2025
Accepted: Jun. 26, 2025
Available Online: Aug. 27, 2025

Keywords:
Fiber Pull-out Curve
Machine Learning
Convolutional Neural Networks (CNN)
Long Short-term Memory Networks (LSTM)
Extreme Gradient Boosting (XG-Boost)

1- Introduction

The study of fiber pull-out behavior in cement matrices is an important topic for materials and construction engineers, as it significantly affects the mechanical properties and durability of concrete. Cement-based matrices have low tensile strength and ductility, necessitating reinforcement. However, the incorporation of fibers into these matrices significantly improves their tensile strength and ductility. The use of fibers in cement matrices not only increases crack resistance but also improves energy absorption and the durability of the cement composite.

Numerous laboratory studies have been carried out on this topic, each of which deals with different aspects of the topic. For example, Zeighami et al. [1] tested the pullout behavior of hooked steel fibers at various angles and found that most pullout forces occurred at angles between thirty and forty-five degrees. Isla et al. [2] investigated the pullout tests of various types of steel fibers. A significant variability in the pull-out force was identified and emphasizing that the relationship between matrix and fiber strength must be taken into account in the FRC construction. Zanjani et al. [3] showed that three-dimensional fibers increase pullout energy and resistance when pullouted compared to two-dimensional fibers. Wu et al. [4] investigated the results of matrix composition and fiber shape with regard to the bending properties of ultra-high-performance concrete (UHPC), and found that the use of deformed fibers can improve the bonding properties of the interfaces. McSwain et al. [5] investigated the effect of compressive stress on the pullout behavior of fibers and showed a positive correlation between peak force and confining stress, though total work did not exhibit such a relationship. Denga et al. [6] analyzed the pullout behavior of hook-shaped and straight fibers using analytical models and came to the conclusion that hybrid fibers significantly increased the pullout force. Cao et al. [7] investigated the effects of the loading rate on pullout behavior and found that overall performance depends on loading rates. Feng et al. [8] examined various types of metal fibers and found that reduced relative spacing reduces average bond strength. Chun et al. [9] showed that the surface roughness caused by sanding positively influences slip-hardening behavior.

In recent years, new numerical and analytical models have been developed to study fiber pullout behavior. Zhang & Yu [10] recommended a numerical model for reproducing the pullout behavior of a single fiber embedded in a cement

*Corresponding author's email: mjalali@shahroodut.ac.ir

Copyrights for this article are retained by the author(s) with publishing rights granted to Amirkabir University Press. The content of this article is subject to the terms and conditions of the Creative Commons Attribution 4.0 International (CC-BY-NC 4.0) License. For more information, please visit https://www.creativecommons.org/licenses/by-nc/4.0/legalcode.

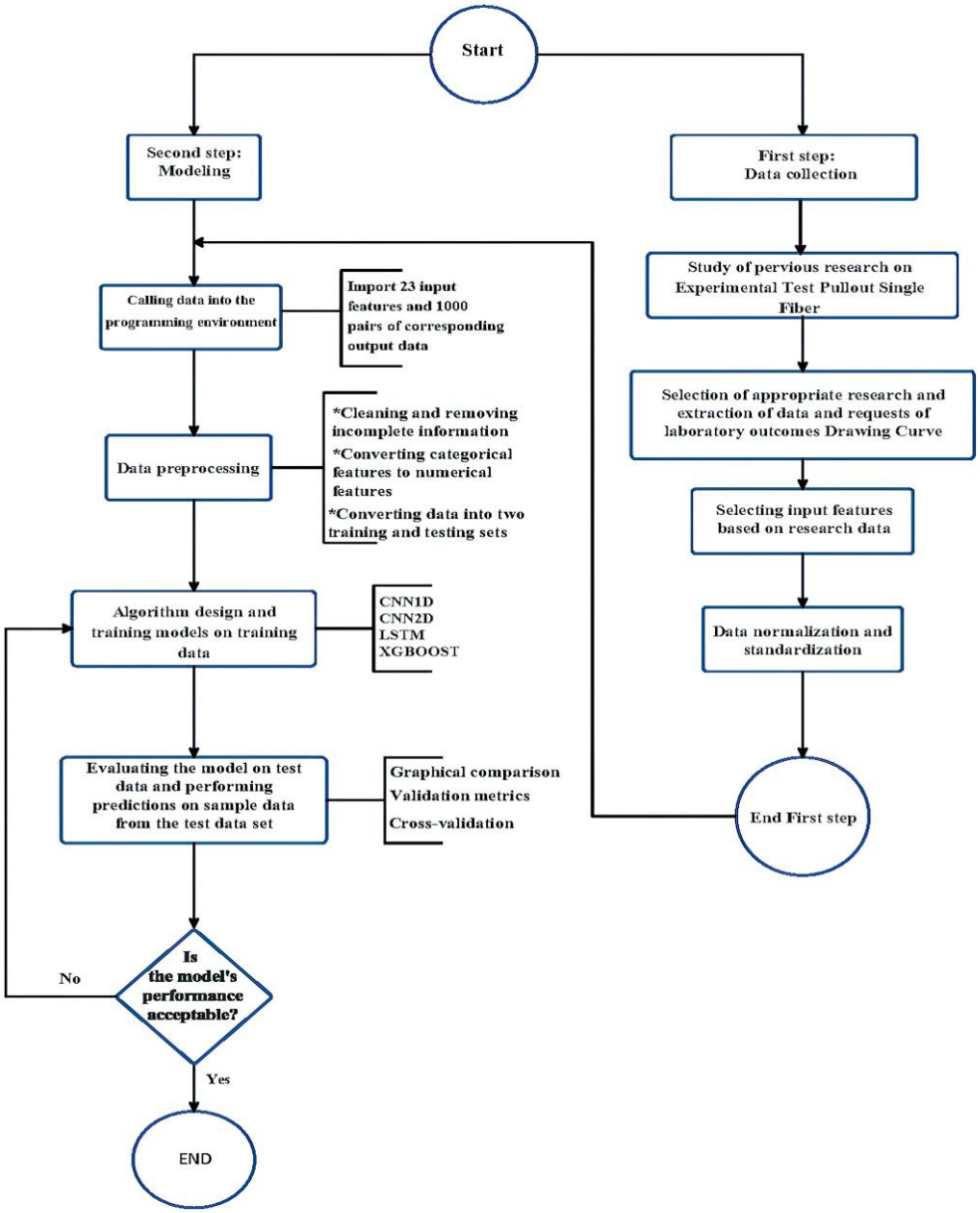


Fig. 1. Flowchart of the paper process

matrix, taking into account the gradual deterioration of bond strength at the interface. In addition, Hoppe [11] used finite element models to simulate the interface separation between fibers and matrix.

These studies underscore the critical importance of studying fiber pullout behavior, especially in light of advances in science and artificial intelligence that require similar research in this area, as shown by several studies that used AI techniques for this purpose. Hemmatian et al. [12] presented an AI model for predicting maximum fiber pullout force and associated slip. 382 experimental information points were used to develop and train artificial neural network models (ANN). The results suggested that the proposed model provided logical predictions and could offer a data-driven approach to optimize fiber-reinforced cement composites. Huang et al. [13] commissioned six unique models, including ANN, gradient-boosted decision trees (GBDT), and XGBoost to predict the bond strength between fibers and matrix, with XGBoost showing advanced predictive performance with an R^2 value of 0.98.

Over the past decades, significant advances have been made through experimental and numerical investigations to characterize the pullout response. While several artificial intelligence studies have successfully predicted key parameters such as the maximum pullout force and the corresponding slip, no research to date has addressed the

prediction of the full pullout curve.

The pullout curve provides a wealth of information beyond peak force values, enabling a more detailed understanding of the fiber-matrix interaction mechanisms. Analysis of the slip-pullout curve allows for the extraction of numerous critical parameters, including bond strength, frictional resistance, and energy absorption characteristics of both the fiber and the surrounding matrix. For example, a steep initial slope in the pullout curve is indicative of a strong interfacial bond between fiber and matrix, whereas a more gradual slope may reflect weaker bonding or increased frictional losses during fiber extraction [14]. Such detailed characterization is essential for elucidating the micromechanical behavior of composites and for optimizing their performance in various engineering applications. This study aims to fill this gap by developing ML-based models for accurate pullout curve prediction.

2- Cement-Based Matrix

A cement-based matrix, such as concrete and mortar, consists of three components: cement, sand, and aggregates combined with water. Although these materials have high compressive strength, they behave weakly when exposed to tension. Therefore, fibers are usually added to cement-mixed materials to improve their mechanical and durable properties [15]. The use of fibers (whether steel, polypropylene, or recycled spring) in concrete is an effective solution to improve mechanical properties, increase durability, and enhance concrete's resistance in harsh conditions, especially after a fire [16]. The use of steel and polypropylene fibers, especially in optimal combination, can be an effective solution for increasing the strength, durability, and economic efficiency in the production of concrete sleepers [17]. Adding different types of fibers can improve the technical properties of concrete against pressure, tension, and heat [18]. The use of steel fibers in lightweight plastic concrete can largely compensate for the mechanical resistance weakness of this type of concrete, while polypropylene and glass fibers are more suitable for improving ductility and reducing surface cracking and do not have a significant effect on mechanical resistance at high temperatures [19].

3- Fiber pull-out behavior

The pullout behavior of fibers in cementitious matrices is influenced by factors such as fiber type, orientation, and matrix properties.

Fig. 2 shows the possible failure modes of the fiber-reinforced matrix: fiber bridging, debonding, pullout, and failure.

By carrying out a fiber pullout test, a pullout-slip curve (Fig. 3) is determined from which important parameters, such as maximum slip, peak pullout force, primary and secondary stiffness, pull-out energy, failure force, bond strength, corresponding slip at failure, etc., can be extracted.

Two important features of pull-slip curves include: Debond process (At the beginning of the test process, fiber debonding begins from the cement matrix with increasing load [14, 21]), Pullout process (After the debonding step, the fibers begin to pull out of the matrix [22]).

4- Background of the Machine learning architecture used

CNN and LSTM models were selected due to their strong performance in time series and sequential data modeling, while XGBoost was included for its high accuracy and interpretability in regression tasks.

4- 1- Convolutional neural network (CNN)

One of the most important deep learning methods is convolutional neural networks (CNNs), which are trained using an effective multi-layer technique. In general, a CNN consists of the three most important layers: the Convolutional Layer, the Pooling Layer, and the Fully Connected Layer (flattened layer). Each layer performs a different task. (Fig. 4) [23].

Convolutional neural networks (CNNs) are increasingly being used in civil engineering to monitor structural health and identify faults. Key findings include: Object detection [24], Crack detection structural health monitoring [25], Image and defect detection [26], and Cost efficiency [27].

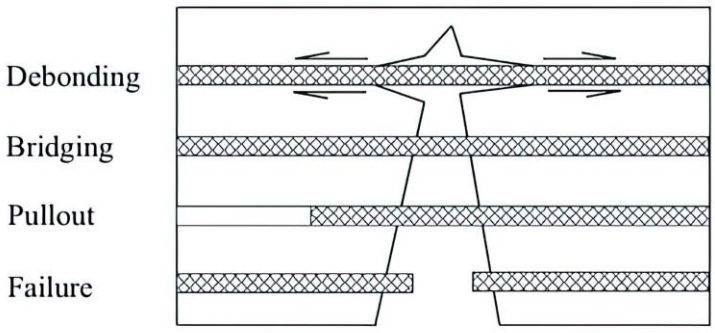


Fig. 2. Mechanisms of fiber–matrix interactions [20].

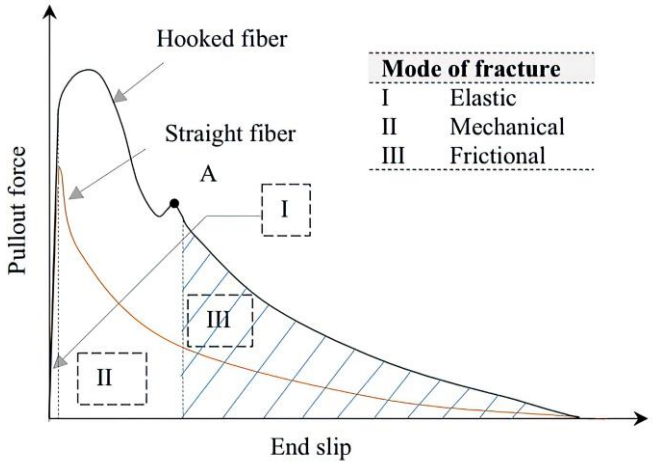


Fig. 3. Pullout modes of fracture for steel fibers [20].

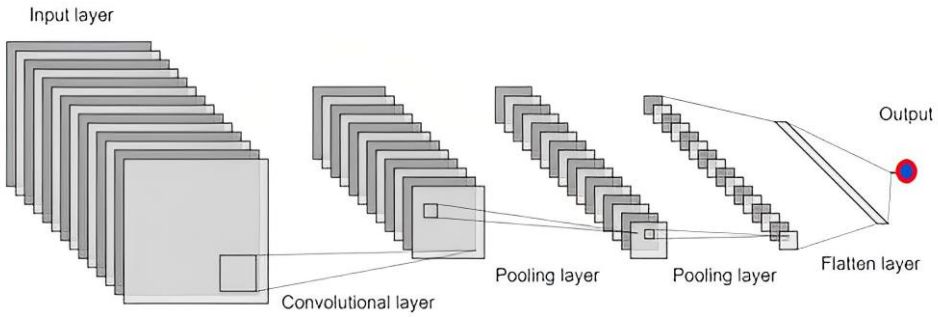


Fig. 4. Fundamental components of the CNN [23].

4- 2- Long short-term memory (LSTM)

Long Short-Term Memory (LSTM) networks (Fig. 5) are a type of recurrent neural network (RNN) specifically designed for learning long-term dependencies in sequence data [28].

The architecture of an LSTM unit comprises several critical components:

Memory Cell: This is the center of the LSTM, which is responsible for storing data over time [35].

Gates: LSTMs use 3 types of gates to regulate data flow: Input gate (Controls how much of the brand-new data from the current entry should be entered into the memory cell [28]), Forget gate (Decides which records must be deleted from the storage cell [28]), Output gate (Determines which data sets must be output from the memory cell to the next layer [28]).

A study showed that the combination of LSTM with optimization algorithms drastically complements the accuracy of deformation forecasts for concrete dams and thus effectively overcomes challenges associated with transient and complicated influencing factors [29]. In addition, LSTM was used to predict the compressive strength of concrete with mineral components. It proved to be powerful in processing complex data sets and in achieving certain predictions that are essential for optimal management of production substances [30]. The results suggest that LSTM can effectively predict loads in different situations, which contributes to advanced methods for structural health monitoring [31].

4- 3- Extreme gradient boosting (XGBoost)

XGBoost (Fig. 6), described as a fairly streamlined and superior implementation of gradient boosting, is often used to fix supervised learning problems with excessive accuracy. He uses classification and regression trees (CART) as a basic learner [32].

Initially, XGBoost generates several vulnerable trees that are personally low in accuracy, rather than trying to build a single tree with medium accuracy [32]. The next tree is trained on the remaining errors from the previous trees,

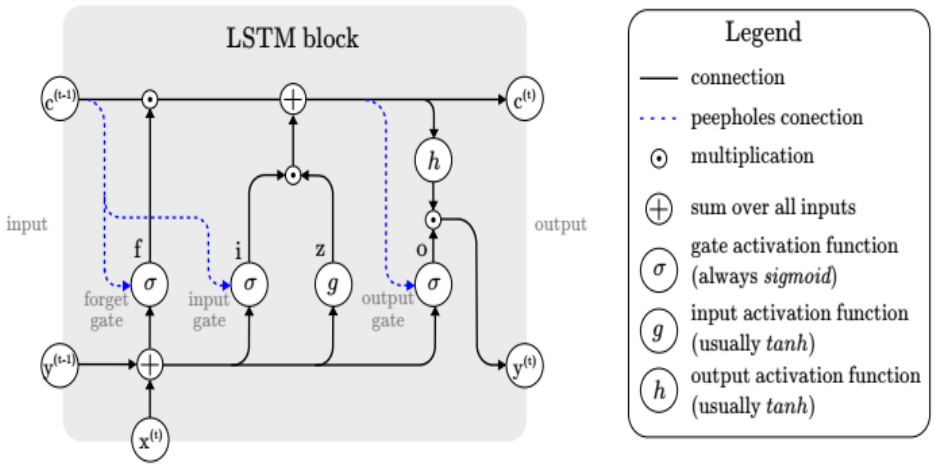


Fig. 5. Architecture of a typical LSTM [28].

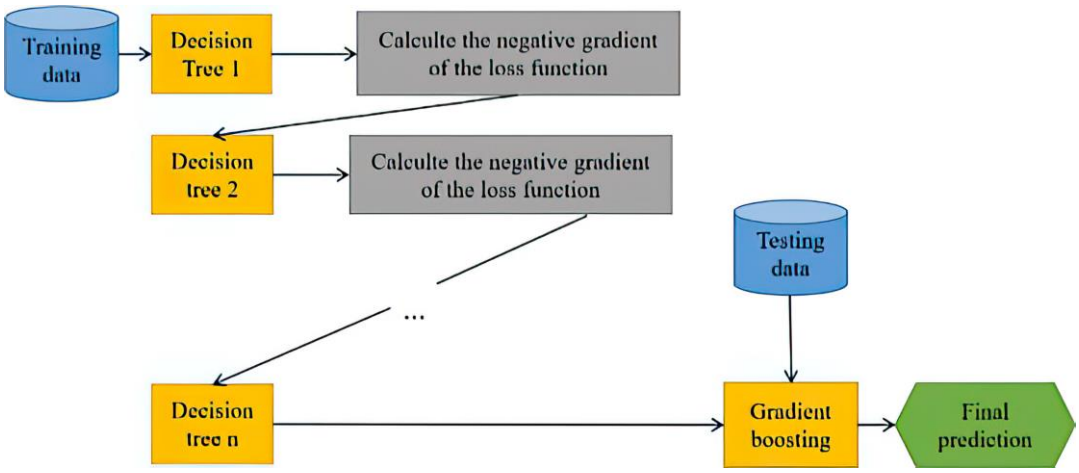


Fig. 6. XGBoost model [33].

which efficiently reduces prediction errors over the course of repetitions [34]. XGBoost includes innovative techniques such as regularized boosting, tree pruning, and parallel processing, which improve its accuracy and efficiency [34].

XGBoost is used to predict the compressive strength of UHPC and other types of concrete and is characterized by high accuracy (R^2 values around 0.90 to 0.93), efficiency, and low error fees compared to other machine learning models, making it a valuable tool for optimizing mix designs and ensuring structural integrity [35, 36]. In addition, XGBoost is used to assess the behavior of superplasticized cement paste, which provides advanced predictions for rheological properties such as yield stress and plastic viscosity [37]. A model based entirely on XGBoost has been further developed to predict the mechanical properties of steel bridge deck systems. It provides accurate forecasts that help with structural design [38]. Furthermore, XGBoost, mixed with particle swarm optimization, is used to expect the axial compressive strength and modulus of rock samples [39].

5- Data collection

Collection and preparation of information from previous experimental studies to predict the fiber pullout curve of the cement-based matrix using ML.

5- 1- Identifying data sources

Data required for algorithm design are obtained through the review of published articles about fibers and cement matrices. Three articles' experimental data are used in this study; two of these were conducted by one of the current

Table 1. The values of each input feature are extracted from the articles.

Ref.	binder kg/m3	Silica fume (microsilica) kg/m3	Sand kg/m3	Gravel kg/m3	Ground quartz kg/m3	Water kg/m3	SP(Superplasticizer) kg/m3	W/C	W/B	Curing age(week)	$f\,c'$ (MPa)
Ataee et al. [40]	0	0, 93, 203	0, 813.13	0, 665.29	0, 1560.9, 1283	168, 148.8, 181	0, 7.13, 12.17	0.42, 0.24	0.42, 0.209, 0.179	20, 12	47.7, 79.29, 109.82
Peyvandi et al. [41]	880, 667	0	988, 792	0	0	188, 250	0	0.5371, 0.431	0.153, 0.2	1	20.6, 30.8
Zeighami et al. [1]	0	25	158	0	0	25	1.2	0.25	0.2	4	40.4

Table 2. Continued values of each input feature are extracted from articles.

Ref.	embedded Length (mm)	Angle of inclination	Material	End-Type	pitch of spirals(mm)	Twisted Number in 10 mm	Length (mm)	Diameter (mm)	Aspect ratio (lf/df)	Loading rate (mm/s)	cement kg/m3	Type Cement
Ataee et al. [40]	10, 15, 20	0	Steel, hook end Steel, Twin-twisted straight	Yes Hook No Hook	0	0, 2, 3, 4	35,40	0.38, 0.41, 0.5, 0.537	70, 74, 92, 97	0.13	400, 620, 811	II
Peyvandi et al. [41]	10, 15, 20	0, 15, 30, 45, 60	Steel, Spirally Deformed	No Hook	4.44	0	40	0.4	100	0.025	350, 580	II
Zeighami et al. [1]	10, 15, 20, 25	0	Steel, hook end	Yes Hook	0	0	50	1	50	0.13	100	II

paper's authors.

S1. Ataee et al. [40]: The pullout behavior of twisted fibers from cement matrices with different strengths was investigated. 330 samples were taken from 54 series of twisted fibers and 12 series of hookend fibers.

S2. Peyvandi et al. [41]: Carrying out experiments on engineered spiral fibers, in which various parameters such as matrix strength and fiber inclination angle were analyzed. A total of 180 samples were obtained.

S3. Zeighami et al. [1]: Investigated the pull-out behavior of inclined hook-end fibers, testing specific angles and embedding lengths. 100 tests were then carried out.

A total of 502 samples are available for analysis: 302 from Ataee et al., 180 from Peyvandi et al., and 20 from Zeighami et al.

5- 2- Selecting Input Features

Based on the concrete compositions used in the three abovementioned articles, the main features of the selected were defined as follows:

Material (type of material and geometric shape of the fibers), Matrix compressive strength (fc'), Fiber Diameter, Fiber length, Embedment Length, Number of Twists, Inclination Angle, Length-to-Diameter Ratio, end-type of the fibers, pitch of spirals, loading rate, kind cement, quantity cement, binder quantity, silica fume (micro silica), sand, gravel, quartz, superplasticizer, water, water-to-cement ratio (W/C), water-to-binder ratio(W/B), curing age in weeks.

These 23 features were selected to model and predict the pullout behavior of fibers from the cement-based matrix in various configurations. In Tables 1 and 2, the values extracted from The articles are offered for each input function.

5- 3- Output parameters

The output parameters include Excel files derived from the results of the pull-out tests. There is one problem with using them in synthetic intelligence algorithms: The number of pairs of information contained in the Excel files for drawing curves had to be consistent across documents. This requirement has now not been met in most cases. In order to standardize the statistics, a total of 1,000 pairs of statistics were brought together for general research. All Excel files that contained the experimental results of the articles were converted into documents with 1,000 pairs of data using OriginPro 2024 software.

6- Methodology

The modeling methodology used in this study consists of six main steps:

Step 1. Database collection. This step involves collecting the entire data set (input features and output target).

Step 2. Manage and optimize data sets. In this stage, optimizing and preprocessing the data set includes data cleansing and removing unnecessary, incomplete information and empty rows, converting categorical variables to numerical variables, and using data correlation charts (Fig. B1) to identify relationships between features in the data, which can help make better decisions about feature selection and model design for machine learning. The data was converted into two sets of training and testing to provide a data set of sufficient and consistent quality for subsequent training and validation of ML models.

Step 3. Build ML models. In this stage, models for machine learning are constructed, and the models are trained using the training data. 4 models were used in this study, each of which is explained below in terms of the architecture used.

CNN1D: The architecture used in this model includes a Conv1D layer with 32 filters and a kernel size of 3, which uses the RELU activation function for feature extraction. A BatchNormalization layer is then added to improve stability and training speed. Another Conv1D layer with 64 filters and RELU activation is then used to extract more complex features. The output of this layer is converted to a one-dimensional vector using a flattened layer, and a dropout level at a rate of 0.25 is applied to reduce the risk of overfitting. Next, a dense layer of 128 neurons and ReLU activation is added, followed by another dense layer of 2000 neurons and a linear activation function for output prediction.

CNN2D: The architecture used in this model includes two 2D convolutional layers with 32 and 64 filters, using the ReLU activation function for feature extraction. After each convolution layer, a MaxPooling layer is added to reduce dimensionality, along with BatchNormalization layers to improve training speed and stability. A flattened layer is used to convert the multidimensional output of the convolutional layers into a 1D vector that can be entered into fully connected layers. A dropout layer is included to prevent overfitting. Finally, four dense layers of 128, 256, 512, and 2000 neurons are added, using various activation functions (ReLU, Sigmoid, Linear) for final prediction tasks.

LSTM: The architecture of this model includes a 32-unit LSTM level that receives inputs in the specified form and, by setting return_sequences=True, forwards the output sequence to the next level. A MaxPooling1d level is then added to reduce dimensionality, and a dropout level is added at a rate of 0.2 to reduce the risk of over-adjustment. Another LSTM level with 64 units followed by a BatchNormalization level and another dropout level at a rate of 0.2. Next, a final 64-unit LSTM level without return_sequences is added to produce the final output. Finally, a dense layer with 2000 units and a linear activation function is placed for the output forecast.

XGBoost: Parameters defined in this model include "reg: squarederror" for regression tasks, where the squared error is calculated, and eval_metric as the metric for evaluating performance during the training, which is set to "rmse". Training and test data are prepared as dMatrix structures, which are optimized for XGBoost. The model is trained with a maximum of 1,000 boosting rounds and uses an early stop after 5 rounds without any improvement in the validation set (this allows effective training of the model and avoids overfitting).

Step 4. Prediction and Validation. In this step, the validation metrics R^2 (coefficient of determination) and MAE (mean absolute error) are calculated for the test data from all four models (Table 3).

R^2: The coefficient of determination shows the degree of fit of the data to a statistical model. This value varies between 0 and 1. The closer this value is to 1, the more the variance of the dependent variables is explained by the independent variables. Their formula is as follows:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{1}$$

Where SS_{res} is the sum of squares of the remainder and SS_{tot} is the sum of squares of the whole.

MAE: This measure calculates the average of the errors regardless of the measurement direction and enables a direct interpretation of the error measurement. Their formula is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

Table 3. R^2 and MAE for CNND, CNN2D, LSTM, and XGBoost model.

	R^2	MAE
CNN1D	0.63	9.11
CNN2D	0.58	8.80
LSTM	0.59	9.67
XGBoost	0.71	8.38

Where y_i represents the actual values, $\hat{y}_i$ represents the predicted values, and n is the number of observations.

Next, using 4 models for a sample of data (Table 4), its curve is predicted from the test data set, and the curve predicted by each model is plotted with its actual curve for comparison (Fig. 7(a)- s7(d)).

Step 5. Select the best model. The XGBoost model is the best model with the highest R^2 of 0.71 and the lowest MAE of 8.38. The XGBoost model also made the best prediction when it came to comparing the predicted curves with the actual curve and comparing the predicted curves of the models with each other.

Step 6. Feature Importance Graph. This graph, drawn using the best-selected model (XGBoost), is great for interpreting the effect of input features on the curve predicted by the model. The feature importance chart shows the effect of each feature on the model's predictions. Two types of feature importance graphs were extracted from the model. In the first graph (Fig. 8), The horizontal axis of the graph shows how often each feature was used in the decision trees. If a feature is used frequently in various decision trees, it is likely to provide useful information about the effect of the feature, and special attention should be paid to that feature.

In Fig. 8, the diagram shows how often a characteristic is used in different decision trees and how much it contributes to the structure of the model trees. Therefore, features such as "embedded length" and "twisted number in 10 mm," which have the highest values, indicate that these input features have a significant effect on curve prediction. Features such as Cement Type (II), "No Hook" and "Yes Hook" that do not appear in this diagram indicate that these features are ineffective.

In the second graph (Fig. 9), the horizontal axis shows the percentage importance of each feature based on how much each feature contributes to reducing the error. It is calculated as the average gain of all splits that use the feature. Features with a higher percentage of importance (gain) are considered more important because they significantly contribute to reducing prediction errors. According to this graph, the features "Silica fume" with 56.7% and "pitch of spirals" with 18.04% have the greatest influence on the model's predictions.

Table 4. Sample data input features.

feature	value
embedded Length (mm)	20
Angle of inclination	30
Material	Steel, Spirally Deformed
End-Type	No Hook
pitch of spirals(mm)	4.44
Twisted Number in 10 mm	0
Length (mm)	40
Diameter (mm)	0.4
Aspect ratio (lf/df)	100
Loading rate (mm/s)	0.025
cement kg/m3	580
Type Cement	II
binder kg/m3	667
Silica fume(microsilica) kg/m3	0
Sand kg/m3	792
Gravel kg/m3	0
Ground quartz kg/m3	0
Water kg/m3	250
SP(Superplasticizer) kg/m3	0
W/C	0.43
W/B	0.2
Curing age(week)	1
$f\,c'$ (MPa)	30.8

7- Discussion

It was not possible to compare this study with a similar sample, as no research has been carried out to date in the field of predicting fiber pullout curves. Therefore, the results of the study were discussed in three ways:

7- 1- Compare curves

As can be seen in Fig. 7(d), the curve predicted by the XGBoost model almost correctly predicted the behavior and properties of the experimental curve, and this indicates a relatively good and acceptable accuracy of the algorithm, as previous studies (Fig. 10) with similar comparisons confirmed the accuracy of the predicted curve and its precision.

7- 2- Compare validation metrics

Letif et al. [43] state that the model is acceptable for $0.2< R^2 <0.8$. In another study, which predicted stress-crack width curves, the CNN model was used, and an R^2 value of 0.75 was obtained [44]. Muzaffer et al. [45] also stated that an R^2 above 0.67 indicates a strong model. In another article, deep learning models were used as an artificial intelligence method to predict the compressive strength of concrete. In particular, five Deep Neural Network (DNN) models named MLM1, MLM2, MLM3, MLM4, and MLM5 were examined, each with a different level of computing complexity. The MLM5 model with $R^2 = 0.64$ was the best [46]. According to previous research, the XGBoost model with the R^2 criterion (0.71) in this study therefore, has a relatively good and acceptable accuracy in predicting the fiber pullout curve and can be regarded as a reliable result.

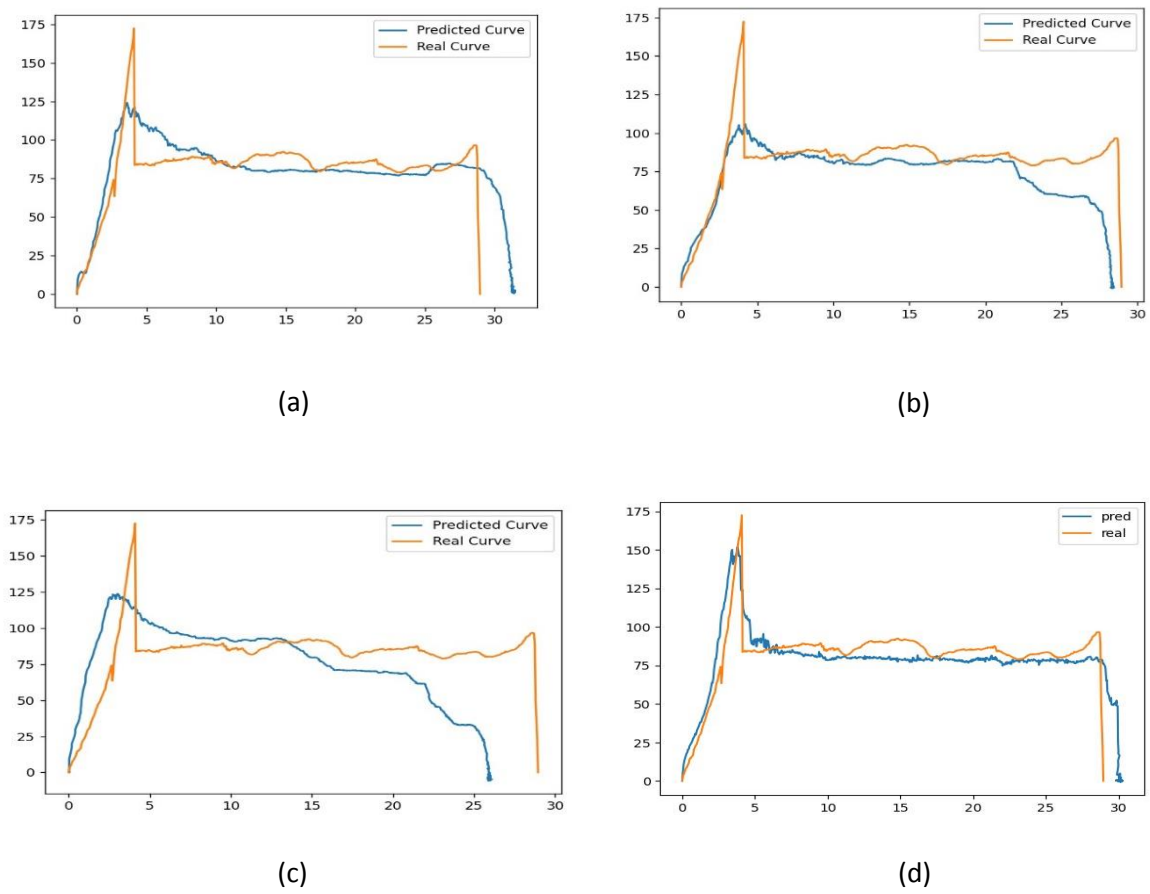


Fig. 7. Comparison of the actual curve and the curve predicted by a) CNN1D model, b) CNN2D model, c) LSTM model, and d) XGBoost model.

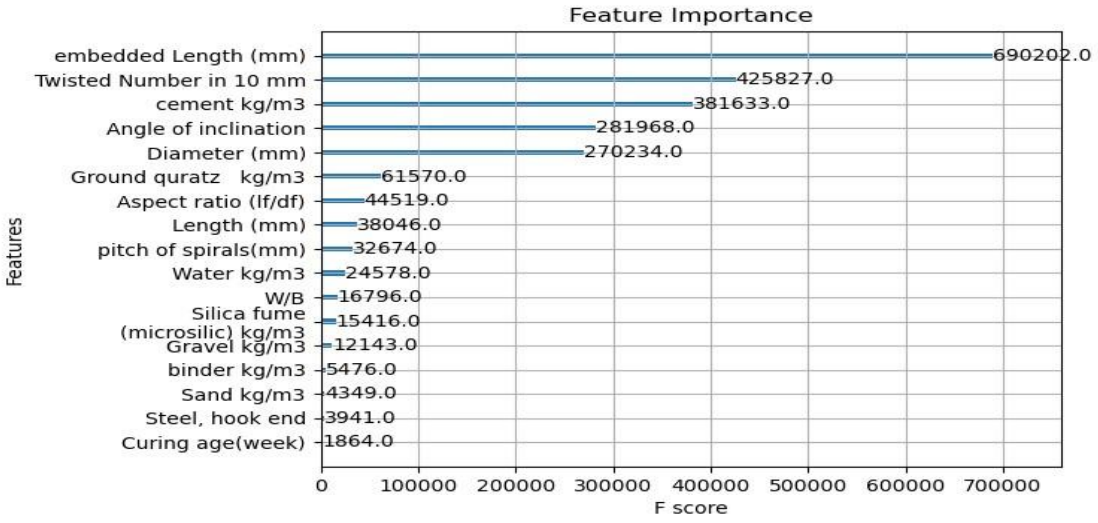


Fig. 8. Feature Importance Graph (frequency).

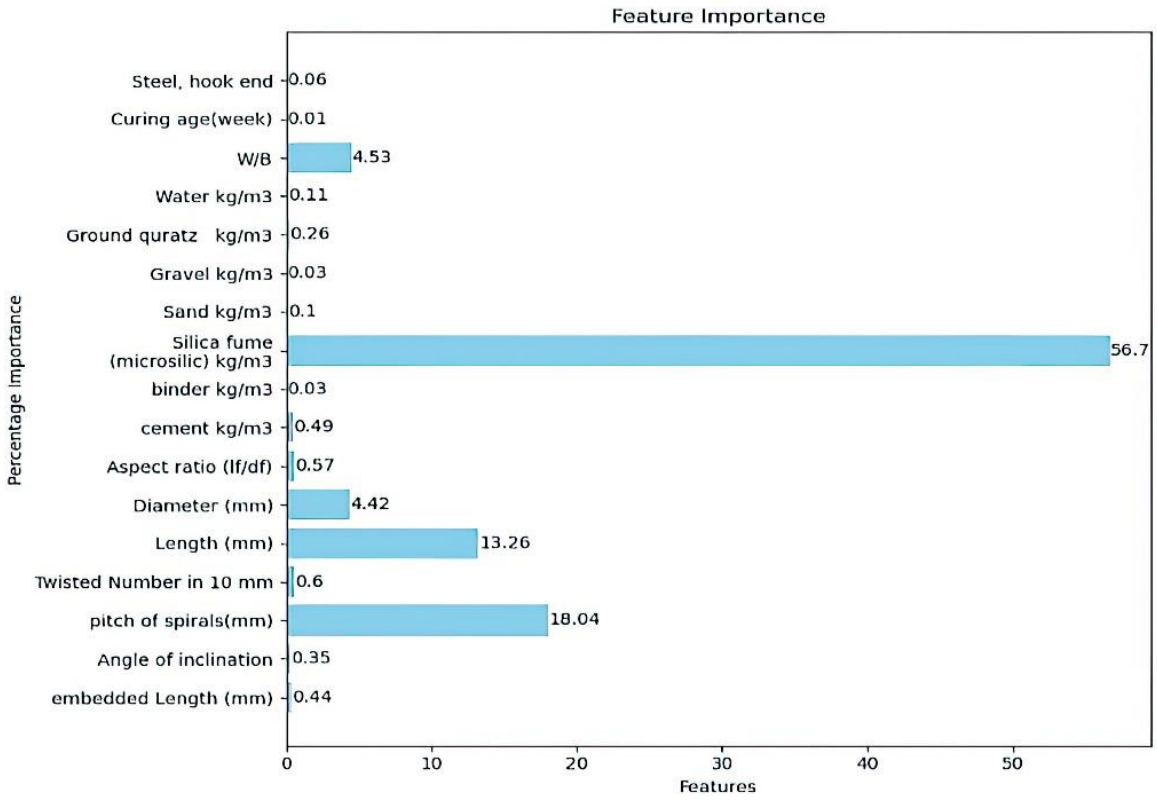


Fig. 9. Feature Importance Graph (gain).

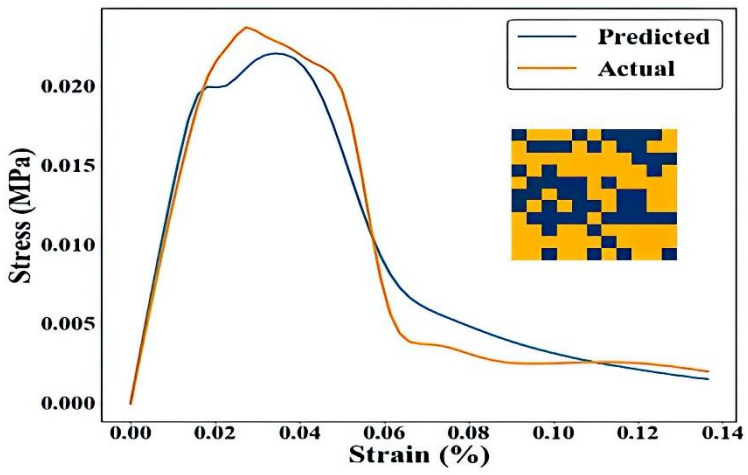


Fig. 10. Comparison of the experimental stress-strain curve with CNN predicted curve [42].

7- 3- Cross Validation

Cross-validation is an important machine-learning technique used to assess model performance. It helps us ensure that machine learning models have good generalization and can perform well on new data. ShuffleSplit is a flexible method for cross-validation. It randomly divides the data into training and test sets and repeats this process several times. In this study, cross-validation using the ShuffleSplit method was used to validate the results of the best-selected model (XGBoost) by specifying the following parameters: the number of times the data was split 5 times, 10% of the data was used for testing, and 90% of the data was used for training. The validation criteria obtained from the 5-fold are given in the table below.

According to Table 5, the highest value for R^2 is 0.82, the lowest value is 0.66, and the average is 0.73, which indicates the accurate performance and generalizability of the model.

Table 5. validation metrics obtained from cross-validation.

Fold	R^2	MAE
1	0.82	7.78
2	0.70	8.20
3	0.71	8.08
4	0.78	7.74
5	0.66	11.33

8- Conclusions

In this study, machine learning models were used to fill a significant gap in the research literature related to predicting fiber pullout curves from cementitious matrices. Four machine learning models, named CNN1D, CNN2D, LSTM, and XGBoost, were used to predict the pullout curve with 23 input features based on a total of 502 laboratory sample data. These four trained models were compared in two ways: by comparing the predicted curves and evaluation criteria. The results of the best model were then confirmed by a comparison with previous research and cross-validation. In addition, the effect of the input features on the predicted curve was investigated by creating a graph of the importance of features for the best model. Based on this study, the following results were obtained:

• Among the evaluated models, XGBoost provided the most accurate pullout curve predictions, demonstrating its effectiveness for modeling fiber-reinforced cementitious composites.

• The results of the feature importance graphs showed that based on the most frequent use in the XGBoost model decision trees, the parameters “embedded length” and “number of twists in 10 mm” have the greatest importance in the prediction, and based on increasing accuracy and reducing error, the parameters “Silica fume” and “pitch of spirals” have the most positive influence on the predicted curve.

• It is suggested that future research should expand the number of databases as well as the quantitative diversity of input features. In particular, laboratory data with different fiber types (only steel fibers were modeled in this study) should be used to model and predict the curve.

Author Contributions

Ali Hossein Pour: Data curation, Investigation, Software, Validation, Writing – original draft. Meysam Jalali: Conceptualization, Methodology, Project administration, Supervision, Formal analysis, Writing – review and editing. Hosein Naderpour: Writing – review and editing.

Acknowledgments

The authors express their gratitude for the technical assistance received from Shahrood University of Technology.

References

[1] E. Zeighami, F. Jandaghi Alaee, M. Jamee, M. Soltani Mohammadi, Experimental investigation of pull-out behavior of inclined fiber from cementitious matrix, Modares Civil Engineering journal, 16(2) (2016) 173-186. http://mcej.modares.ac.ir/article-16-6742-en.html

[2] F. Isla, G. Ruano, B. Luccioni, Analysis of steel fibers pull-out. Experimental study, Construction and Building Materials, 100 (2015) 183-193. https://doi.org/10.1016/j.conbuildmat.2015.09.034

[3] E.M. Zanjani, S. Barnett, D. Begg, Pullout behaviour of hooked end steel fibres embedded in concrete with various cement replacement materials, in: Proc., 9th RILEM Int. Symp. of Fiber Reinforced Concrete. Vancouver, Canada, 2016.

[4] Z. Wu, K.H. Khayat, C. Shi, How do fiber shape and matrix composition affect fiber pullout behavior and flexural properties of UHPC?, Cement and Concrete Composites, 90 (2018) 193-201. https://doi.org/10.1016/j.cemconcomp.2018.03.021

[5] A.C. McSwain, K.A. Berube, G. Cusatis, E.N. Landis, Confinement effects on fiber pullout forces for ultra-high-performance concrete, Cement and Concrete Composites, 91 (2018) 53-58. https://doi.org/10.1016/j.cemconcomp.2018.04.011

[6] F. Deng, X. Ding, Y. Chi, L. Xu, L. Wang, The pull-out behavior of straight and hooked-end steel fiber from hybrid fiber reinforced cementitious composite: Experimental study and analytical modelling, Composite Structures, 206 (2018) 693-712. https://doi.org/10.1016/j.compstruct.2018.08.066

[7] Y. Cao, Q. Yu, H. Brouwers, W. Chen, Predicting the rate effects on hooked-end fiber pullout performance from Ultra-High Performance Concrete (UHPC), Cement

and Concrete Research, 120 (2019) 164-175. https://doi.org/10.1016/j.cemconres.2019.03.022

[8] H. Feng, M.N. Sheikh, M.N. Hadi, L. Feng, D. Gao, J. Zhao, Pullout behaviour of different types of steel fibres embedded in magnesium phosphate cementitious matrix, International Journal of Concrete Structures and Materials, 13(1) (2019) 1-17. https://doi.org/10.1186/s40069-019-0344-1

[9] B. Chun, D.-Y. Yoo, N. Banthia, Achieving slip-hardening behavior of sanded straight steel fibers in ultra-high-performance concrete, Cement and Concrete Composites, 113 (2020) 103669. https://doi.org/10.1016/j.cemconcomp.2020.103669

[10] H. Zhang, R.C. Yu, Inclined fiber pullout from a cementitious matrix: A numerical study, Materials, 9(10) (2016) 800. https://doi.org/10.3390/ma9100800

[11] L. Hoppe, Numerical simulation of fiber-matrix debonding in single fiber pull-out tests, GAMM Archive for Students (GAMMAS), 2(1) (2020) 21-35. https://doi.org/10.14464/gammas.v2i1.437

[12] A. Hemmatian, M. Jalali, H. Naderpour, M.L. Nehdi, Machine learning prediction of fiber pull-out and bond-slip in fiber-reinforced cementitious composites, Journal of Building Engineering, 63 (2023) 105474. https://doi.org/10.1016/j.jobe.2022.105474

[13] J.-X. Huang, X.-Z. Shi, N. Zhang, Y.-Q. Hu, J.-Q. Wang, Prediction of Bond Strength between Fibers and the Matrix in UHPC Utilizing Machine Learning and Experimental Data, Materials Today Communications, (2024) 111136. https://doi.org/10.1016/j.mtcomm.2024.111136

[14] L. Huang, M. Yuan, B. Wei, D. Yan, Y. Liu, Experimental investigation on sing fiber pullout behaviour on steel fiber-matrix of reactive powder concrete (RPC), Construction and Building Materials, 318 (2022) 125899. https://doi.org/10.1016/j.conbuildmat.2021.125899

[15] S. Sharda, M. Singh, S. Singh, A review on Properties of Fiber Reinforced Cement-based materials, Iosr. J. Mech. Civ. Eng.(IOSR-JMCE), 13 (2016) 104-112. https://doi.org/10.9790/1684-130501104112

[16] G. Pachideh, M. Gholhaki, A. Moshtagh, Performance of concrete containing recycled springs in post-fire conditions, Proceedings of the Institution of Civil Engineers - Structures and Buildings, 173(1) (2018) 3-16. https://doi.org/10.1680/jstbu.18.00042

[17] G. Pachideh, M. Gholhaki, Using steel and polypropylene fibres to improve the performance of concrete sleepers, Proceedings of the Institution of Civil Engineers-Structures and Buildings, 173(9) (2020) 690-702.https:/doi.org/10.1680/jstbu.18.00154

[18] G. Pachideh, V. Toufigh, Strength of SCLC recycled springs and fibers concrete subject to high temperatures, Structural Concrete, 23(1) (2022) 285-299. https://doi.org/10.1002/suco.202100183

[19] M. Khalily, V. Saberi, H. Saberi, V. Mansouri, A. Sadeghi, G. Pachideh, An Experimental Study on the Effect of High Temperatures on Performance of the Plastic Lightweight Concrete Containing Steel, Polypropylene and Glass Fibers, Journal of Structural and Construction Engineering, 8(12) (2022) 284-307. https://dx.doi.org/10.22065/jsce.2021.254752.2277

[20] Y.M. Abbas, M. Iqbal Khan, Fiber–Matrix Interactions in Fiber-Reinforced Concrete: A Review, Arabian Journal for Science and Engineering, 41(4) (2016) 1183-1198. https://doi.org/10.1007/s13369-016-2099-1

[21] L.F. Friedrich, C. Wang, Continuous modeling technique of fiber pullout from a cement matrix with different interface mechanical properties using finite element program, Latin American Journal of Solids and Structures, 13(10) (2016) 1937-1953. https://doi.org/10.1016/j.conbuildmat.2021.125899

[22] E. Wölfel, H. Brünig, I. Curosu, V. Mechtcherine, C. Scheffler, Dynamic Single-Fiber Pull-Out of Polypropylene Fibers Produced with Different Mechanical and Surface Properties for Concrete Reinforcement, Materials (Basel), 14(4) (2021). https://doi.org/10.3390/ma14040722

[23] M.S. Barkhordari, S. Ghavaminejad, M. Tehranizadeh, Predicting Autogenous Shrinkage of Concrete Including Superabsorbent Polymers and Other Cementitious Ingredients Using Convolution-based Algorithms, Periodica Polytechnica Civil Engineering, (2024). https://doi.org/10.3311/PPci.23568

[24] D. Griffiths, J. Boehm, Rapid object detection systems, utilising deep learning and unmanned aerial systems (uas) for civil engineering applications, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42 (2018) 391-398. https://doi.org/10.5194/isprs-archives-XLII-2-391-2018

[25] F. Şermet, I. Pacal, Deep learning approaches for autonomous crack detection in concrete wall, brick deck and pavement, Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 15(2) (2024) 503-513. https://doi.org/10.24012/dumf.1450640

[26] D. Yang, Deep Learning Based Image Recognition Technology for Civil Engineering Applications, Applied Mathematics and Nonlinear Sciences, 9(1) (2024). https://doi.org/10.2478/amns-2024-0183

[27] M. Padsumbiya, V. Brahmbhatt, S.P. Thakkar, Automatic crack detection using convolutional neural network, Journal of Soft Computing in Civil Engineering, 6(3) (2022) 1-17. https://doi.org/10.22115/scce.2022.325596.1397

[28] G. Van Houdt, C. Mosquera, G. Nápoles, A review on the long short-term memory model, Artificial Intelligence Review, 53(8) (2020) 5929-5955. https://doi.org/10.1007/s10462-020-09838-1

[29] J. Madiniyeti, Y. Chao, T. Li, H. Qi, F. Wang, Concrete dam deformation prediction model research based on SSA–LSTM, Applied Sciences, 13(13) (2023) 7375.

https://doi.org/10.3390/app13137375

[30] A. Gogineni, M.K.D. Rout, K. Shubham, Evaluating machine learning algorithms for predicting compressive strength of concrete with mineral admixture using long short-term memory (LSTM) Technique, Asian Journal of Civil Engineering, (2023) 1-13. https://doi.org/10.1007/s42107-023-00885-x

[31] M. Impraimakis, Deep recurrent-convolutional neural network learning and physics Kalman filtering comparison in dynamic load identification, Structural Health Monitoring, (2024). https://doi.org/10.1177/14759217241262972

[32] D.-K. Thai, D.-N. Le, Q.H. Doan, T.-H. Pham, D.-N. Nguyen, A hybrid model for classifying the impact damage modes of fiber reinforced concrete panels based on XGBoost and Horse Herd Optimization algorithm, Structures, 60 (2024) 105872. https://doi.org/10.1016/j.istruc.2024.105872

[33] J. Wang, S. Zhou, Particle swarm optimization-XGBoost-based modeling of radio-frequency power amplifier under different temperatures, International Journal of Numerical Modelling: Electronic Networks, Devices and Fields, 37(2) (2024) e3168. https://doi.org/10.1002/jnm.3168

[34] A. Pal, K.S. Ahmed, S. Mangalathu, Data-driven machine learning approaches for predicting slump of fiber-reinforced concrete containing waste rubber and recycled aggregate, Construction and Building Materials, 417 (2024) 135369. https://doi.org/10.1016/j.conbuildmat.2024.135369

[35] Z. Shen, A.F. Deifalla, P. Kamiński, A. Dyczko, Compressive strength evaluation of ultra-high-strength concrete by machine learning, Materials, 15(10) (2022) 3523. https://doi.org/10.3390/ma15103523

[36] G. Airlangga, Comparison of Predictive Modeling Concrete Compressive Strength with Machine Learning Approaches, UKaRsT, 8(1) (2024) 28-41. https://doi.org/10.30737/ukarst.v8i1.5532

[37] D. Sathyan, D. Govind, C. Rajesh, K. Gopikrishnan, G.A. Kannan, J. Mahadevan, Modelling the shear flow behaviour of cement paste using machine learning–XGBoost, in: Journal of Physics: Conference Series, IOP Publishing, 2020, pp. 012026. https://doi.org/10.1088/1742-6596/1451/1/012026

[38] Y. Wei, R. Ji, Q. Li, Z. Song, Mechanical Performance Prediction Model of Steel Bridge Deck Pavement System Based on XGBoost, Applied Sciences, 13(21) (2023) 12048. https://doi.org/10.3390/app132112048

[39] N.M. Shahani, Q. Xiaowei, X. Wei, L. Jun, T. Aizitiliwumaier, M. Xiaohu, Q. Shigui, C. Weikang, L. Longhe, Hybrid PSO with tree-based models for predicting uniaxial compressive strength and elastic modulus of rock samples, Frontiers in Earth Science, 12 (2024) 1337823. https://doi.org/10.3389/feart.2024.1337823

[40] S. Ataee, M. Jalali, M.L. Nehdi, Pull-out behavior of twin-twisted steel fibers from various strength cement-based matrices, Construction and Building Materials, 445 (2024) 137855. https://doi.org/10.1016/j.conbuildmat.2024.137855

[41] A.H. Peyvandi, M. Jalali, M. Hajsadeghi, S. Das, Experimental investigation on the performance of engineered spiral fiber: Fiber pull-out and direct tension tests, Construction and Building Materials, 347 (2022) 128569. https://doi.org/10.1016/j.conbuildmat.2022.128569

[42] C. Yang, Y. Kim, S. Ryu, G.X. Gu, Prediction of composite microstructure stress-strain curves using convolutional neural networks, Materials & Design, 189 (2020) 108509. https://doi.org/10.1016/j.matdes.2020.108509

[43] M. Letif, R. Bahar, N. Mezouar, The Use of machine learning models and SHAP interaction values to predict the soil swelling index, Periodica Polytechnica Civil Engineering, 69(1) (2025) 239-250. https://doi.org/10.3311/PPci.36880

[44] Z. Chang, Z. Wan, Y. Xu, E. Schlangen, B. Šavija, Convolutional neural network for predicting crack pattern and stress-crack width curve of air-void structure in 3D printed concrete, Engineering Fracture Mechanics, 271 (2022) 108624. https://doi.org/10.1016/j.engfracmech.2022.108624

[45] S.A. Muzafar, K.N. Ali, M.A. Kassem, M.A. Khoiry, Civil engineering standard measurement method adoption using a structural equation modelling approach, Buildings, 13(4) (2023) 963. https://doi.org/10.3390/buildings13040963

[46] P. Ziolkowski, Computational Complexity and Its Influence on Predictive Capabilities of Machine Learning Models for Concrete Mix Design, Materials, 16(17) (2023) 5956. https://doi.org/10.3390/ma16175956

Appendix A

The figures show in Table A1 the comparison of the prediction curve with the actual curve, with 4 models for 4 data samples, with the characteristics of Table A2.

Table A1. Comparison of the prediction curve with the actual curve with 4 models for 4 data samples

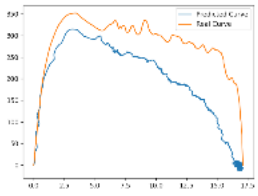
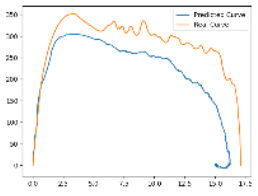
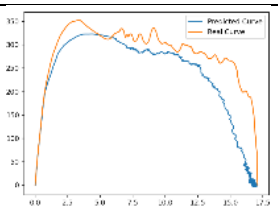
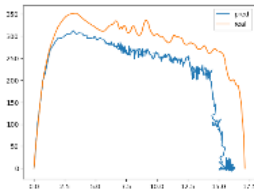
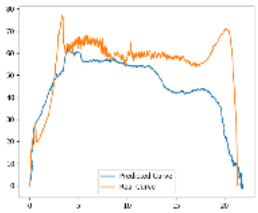
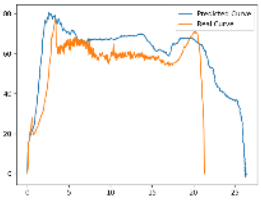
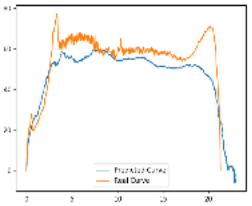
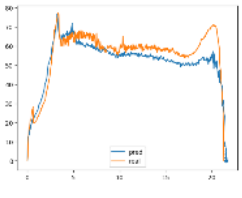
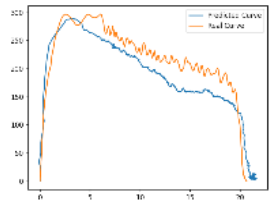
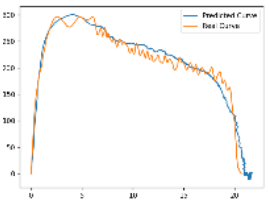
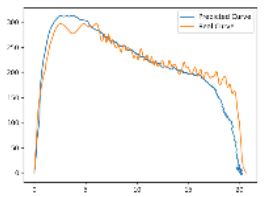
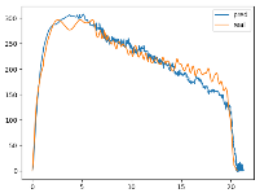
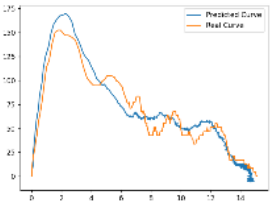
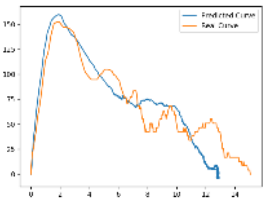
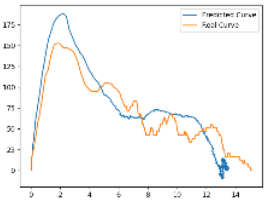
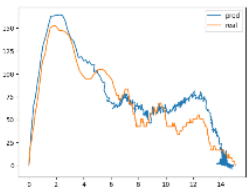
	CNN	CNN(2D)	LSTM	XGBoost
Sample 1				
Sample 2				
Sample 3				
Sample 4				

Table A2. Input features of 4 sample data

feature	value			
	Sample 1	Sample 2	Sample 3	Sample 4
embedded Length (mm)	15	15	20	15
Angle of inclination	0	45	0	0
Material	Steel, Twin-twisted straight	Steel, Spirally Deformed	Steel, Twin-twisted straight	Steel, hook end
End-Type	No hook	No hook	No hook	Yes hook
pitch of spirals(mm)	0	4.44	0	0
Twisted Number in 10 mm	4	0	2	0
Length (mm)	40	40	40	35
Diameter (mm)	0.537	0.4	0.537	0.5
Aspect ratio (lf/df)	74.48	100	74.48	70
Loading rate (mm/s)	0.13	0.025	0.13	0.13
cement kg/m3	811	350	620	400
Type Cement	II	II	II	II
binder kg/m3	0	880	0	0
Silica fume(microsilica) kg/m3	203	0	93	0
Sand kg/m3	0	988	0	813.13
Gravel kg/m3	0	0	0	665.29
Ground quartz kg/m3	1283	0	1560.9	0
Water kg/m3	181	188	148	168
SP(Superplasticizer) kg/m3	12.17	0	7.13	0
W/C	0.42	0.53	0.24	0.42
W/B	0.17	0.15	0.20	0.42
Curing age(week)	12	1	12	20
$f\,c'$ (MPa)	109.82	20.6	79.29	47.7

Appendix B

Correlation charts to identify relationships between features in the data.

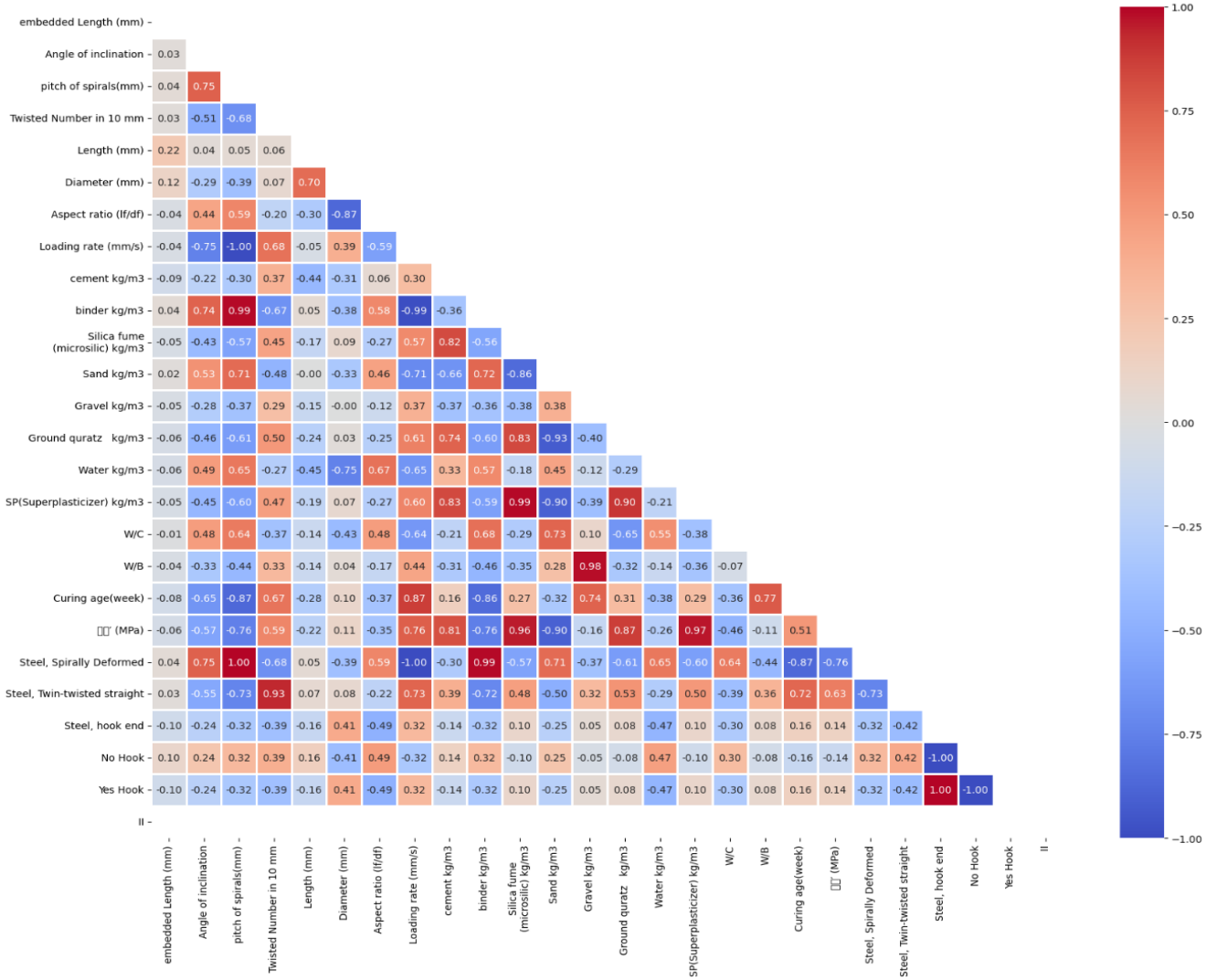


Fig. B1. correlation chart

HOW TO CITE THIS ARTICLE
A. Hossein pour, M. Jalali, H. Naderpour, Estimation of Fiber Pullout Curve from Cement Base Matrix Using Machine Learning, AUT J. Civil Eng., 9(2) (2025) 143-158.

DOI: 10.22060/ajce.2025.24055.5914