Machine Learning Research
2019; 4(2): 27-32
http://www.sciencepublishinggroup.com/j/mlr
doi: 10.11648/j.mlr.20190402.11
ISSN: 2637-5672 (Print); ISSN: 2637-5680 (Online)

Learning Algorithms Using BPNN & SFS for Prediction of Compressive Strength of Ultra-High Performance Concrete

Deepak Choudhary

Electronics & Communication Department, ABES Engineering College, Ghaziabad, India

Email address:
engg_deepak@yahoo.com

To cite this article:
Deepak Choudhary. Learning Algorithms Using BPNN & SFS for Prediction of Compressive Strength of Ultra-High Performance Concrete. *Machine Learning Research*. Vol. 4, No. 2, 2019, pp. 27-32. doi: 10.11648/j.mlr.20190402.11

Received: April 4, 2019; **Accepted**: May 23, 2019; **Published**: June 25, 2019

Abstract: This paper presents machine learning algorithms based on back-propagation neural network (BPNN) that employs sequential feature selection (SFS) for predicting the compressive strength of Ultra-High Performance Concrete (UHPC). A database, containing 110 points and eight material constituents, was collected from the literature for the development of models using machine learning techniques. The BPNN and SFS were used interchangeably to identify the relevant features that contributed with the response variable. As a result, the BPNN with the selected features was able to interpret more accurate results (r^2 = 0.991) than the model with all the features (r^2 = 0.816). The utilization of ANN modelling made its way into the prediction of fresh and hardened properties of concrete based on given experimental input parameters, whereby several authors developed AI models to predict the compressive strength of normal weight, light weight and recycled concrete. The steps that were are followed in developing a robust and accurate numerical model using SFS include (1) design and validation of ANN model by manipulating the number of neurons and hidden layers; (2) execution of SFS using ANN as a wrapper; and (3) analysis of selected features using both ANN and nonlinear regression. It is concluded that the usage of ANN with SFS provided an improvement to the prediction model's accuracy, making it a viable tool for machine learning approaches in civil engineering case studies.

Keywords: ANN, SFS, UHPC, Compressive Strength, Constituents

1. Introduction

Several types of machine learning algorithms such as Artificial Neural Network (ANN) have been used in different fields for the development of models that predict response parameters (experimental dataset) using certain independent input parameters. However, an experiment could have a large number of independent parameters most of which are redundant and have negligible effects on the response parameters. Therefore, an artificially intelligent (AI) selection algorithm is required to overcome this shortcoming and identify the underlying parameters that improve the model's accuracy and simplify the computational complexity. The need for soft computing tools and models for the prediction of behavioural properties of engineering components, systems and materials is continuously rising. ANN emerged as one of soft computing paradigms that have been successfully applied in several engineering fields [1]. Specifically, ANN has been used to solve a wide variety of civil engineering problems [2–4]. Mainly, ANN was utilized to model the nonlinear behaviour of fatigue and creep of Reinforced Concrete (RC) members [5–8]. Recently, research interest has revolved around the development of ANN models to interpret the behaviour of structural materials such as steel, concrete, and composites [9– 14]. The utilization of ANN modelling made its way into the prediction of fresh and hardened properties of concrete based on given experimental input parameters, whereby several authors developed AI models to predict the compressive strength of normal weight, light weight and recycled concrete [14–17]. Afterwards, several authors began developing ANN models for the prediction of compressive strength of high performance concrete [18–21]. In this study ANN is employed with other machine learning techniques to identify the parameters that capture the compressive strength of UHPC using data collected from the literature.

2. UHPC and ANN Background

The evolution of UHPC has lead structural engineers to improve the compressive strength, ductility, and durability of heavy loaded reinforced concrete (RC) structures. Several researchers have been investigating the mechanical behaviour of UHPC and its applications over the last four decade, where it was founded that UHPC exhibits a compressive strength that would range from 150 to 810 MPa [22, 23]. The underlying material constituents that enable such a superior mix are cement (up to 800 kg/m3), water/binder ratio that is lower than 0.20, high-range water-reducing (HRWR) admixture, very fine powders (crushed quartzite and silica fume), and steel fibers.

[24]. Other researchers proposed different mixtures by adding fly ash and sand to reduce the amount of cement and silica fume, and acquire an optimum mix that is both economical and sustainable [25, 26]. However, most of the aforementioned mixtures result in exhausting a large amount of resources and performing tests on many batches, while barely predicting the strength of UHPC [19]. Therefore, researchers began conducting investigations on the utilization of machine learning techniques for the development of prediction models that could assist engineers and researchers to produce appropriate UHPC mixes.

Ghafari et al. [19] used the back-propagation neural network (BPNN) and statistical mixture design (SMD) in predicting the required performance of UHPC. The objective of the study was to develop an ANN and SMD model to predict both the compressive strength and the consistency of UHPC with two different types of curing systems (steam curing and wet curing). As a result, BPNN proved to be more accurate than SMD in the prediction of compressive strength and slump flow of UHPC. Despite the statistical advantages of ANN, it has been long regarded as a black box that evaluates functions using input covariates and yielding outputs. Meaning, the model does not produce any analytical model with a mathematical structure that can be studied. Therefore, ANN should be utilized in detecting the dominant input parameters that have direct association with the ANN model. This will reduce the amount of parameters in the model, which will improve the computation complexity of the ANN model and simplify the derivation strategies of a mathematical model used to predict the compressive strength of UHPC. In addition prediction of compressive strength of high strength and high performance concrete was addressed by other researchers [20, 21].

There are several machine learning techniques, in the literature, that assist researchers in identifying the underlying covariates impacting the prediction model. Sequential feature selection (SFS) is a machine learning tool that sequentially selects features and inputs them into a fitting model (i.e. ANN) until the model's error function increases. This technique makes use of ANN's complex computation and allows the SFS tool to select and remove the influential and redundant parameters, respectively. The reduction in the covariate domain improves the accuracy of the fitting model, decreases its computation time, and facilitates a better understanding of the data processing [27]. There are two types of SFS classes – mainly filter method and wrapper method [28], where Zhou et al. [29] used the Markov Blanket with a wrapper method to select the most relevant features of human motion recognition. Four sets of open human motion data and two types of machine learning algorithms were used. The total number of features was reduced rapidly, where this reduction helped the algorithm demonstrate better recognition accuracy than traditional methods. Moreover, Rodriguez-Galiano et al. [30] used SFS when tackling ground water quality problems, where 20 datasets of parameters were extracted from a GIS database. Four types of machine learning algorithms were used as wrappers for the SFS. As a result, the Rain Forest machine learning algorithm used with SFS showed promising results, where only three features were sufficient enough in predicting the most accurate results.

Table 1. *Experimental Program.*

Variable	Symbols	Minimum	Maximum
Cement (kg/m^3)	C	383	1600
Silica fume (kg/m^3)	SI	0	367.95
Fly ash (kg/m^3)	FA	0	448
Sand (kg/m^3)	S	0	1898
Steel Fiber (kg/m^3)	SF	0	470
Quartz Powder (kg/m^3)	QP	0	750
Water (kg/m^3)	W	109	334.5
Admixture (kg/m^3)	A	0	185
f'c (MPa)	fc	95	240

3. Methodology of Modeling

The steps that were are followed in developing a robust and accurate numerical model using SFS include (1) design and validation of ANN model by manipulating the number of neurons and hidden layers; (2) execution of SFS using ANN as a wrapper; and (3) analysis of selected features using both ANN and nonlinear regression. Table 1 presents the initial input variables together with their range (maximum and minimum values) and symbols for identifying them in this experimental program.

3.1. Artificial Neural Network

Artificial neural network (ANN) is a machine learning tool that imitates the learning functions of a human brain by providing a robust technique in classifying and predicting certain outcomes based on the model's objective. There are two types of ANN models: (1) feed forward; and (2) feed backward. In this study, the feed backward ANN is used, where it is composed of input neurons, hidden neurons, bias units, wires containing randomly generated weights, and output neurons. The input neurons are responsible for containing the independent parameter presented by the user, the wires represent the randomly generated matrices called weights that manipulate the function's slope or steepness, the hidden neurons map the weights variables using an activation function, and the bias units control the output function's shift,

either upward or downward. Equation (1) shows the linear combination of mapping weights from each input neuron, via wires, to the hidden neurons.

$$O_i = g\left(\sum_{i=0}^{n}\sum_{i=0}^{m}\left(X_i \times \theta_{ij}^{T}\right)\right) \quad (1)$$

Where X_i represents the first input parameter of size $R\times1$ (R is the number of data points), θ_{ij} is the weight of size $R\times(n+1)$, O_i is the value of the output neuron or prediction function $h_\theta(X)$, and $g(x)$ is the activation function. The bias unit is simulated by creating a column vector of size $R\times1$ and assigning it with values of ones, where X_0 and θ_{i0} contain the bias values.

3.2. Sequential Feature Selection

SFS reduces the dimensionality of data by selecting only a subset of measured features to create a prediction model. SFS is composed of two components: the objective function, which is the criteria the algorithm follows when selecting the features (i.e. the NMSE), and the search algorithm, which are the methods of how the machine add/removes the features from the subset. There are two types of search algorithms: sequential forward selection and sequential backward selection. In this study, the previously verified ANN model was used as the objective function and the forward selection was used in selecting the relevant features. Figure 1 shows the algorithm SFS uses when performing forward selection.

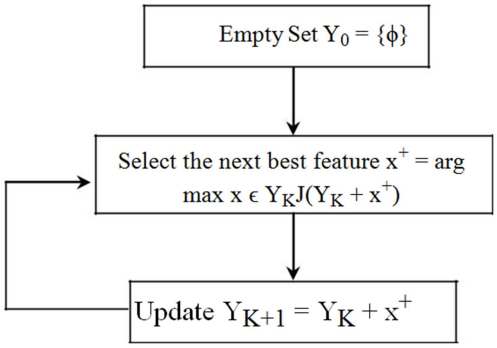


Figure 1. *Foward Selection Algorithm.*

4. Ai Modeling for Prediction of Compressive Strength of UHPC

4.1. Verification of ANN

The ANN numerical solver, Levenberg-Marquardt, was verified by testing different number of neurons using a basis like the normalized mean square error (NMSE) to measure the error. The increment started from one neuron and ended with 15 neurons, where the model was analyzed 10 times, for each increment, because the Levenberg-Marquardt algorithm locates the local, and not the global, minimum of a function. Hence, for each neuron tested, ten NMSE values will be stored in a column vector, where each column vector will be averaged and plotted against its corresponding number of neuron(s). Figure 2 shows the plot of all the scenarios with the minimum point circled at 11 neurons. Therefore, 11 neurons is, approximately, the number of neurons that is sufficient enough for BPNN to facilitate an accurate ANN model for the collected dataset.

4.2. Execution of SFS

The SFS algorithm was run 200 times to capture all possible combinations of independent features when using ANN. Table 2 tabulates the percentage of features that were used during the 200 trials. Based on the results of these trials, the most abundant combination during the SFS analysis, within a 20% threshold, was selected as the important parameters that contribute mostly in the model. In this study, four variables (Cement, Sillica Fume, Flyash, and Water) were selected as the most relevant features for the prediction model. Figure 3(a) and (b) present the architecture of both ANN models, before and after selection.

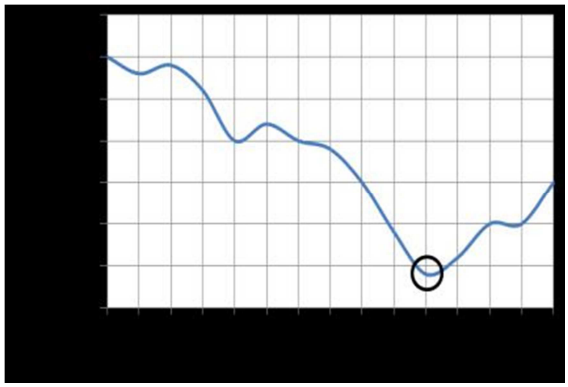


Figure 2. *Verification Curve for ANN.*

Table 2. *SFS Results.*

Materials	Percentage Used
C	24%
SI	33%
FA	24%
S	14%
SF	6%
QP	5%
W	31%
A	13%

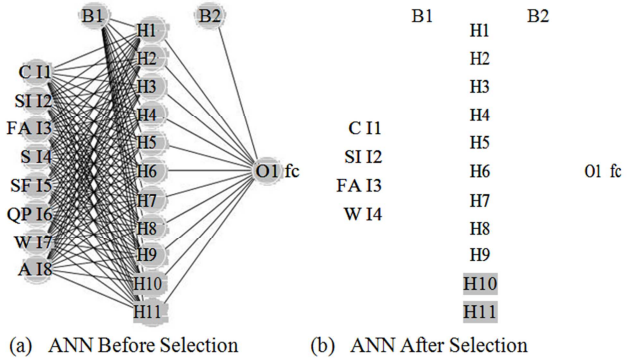


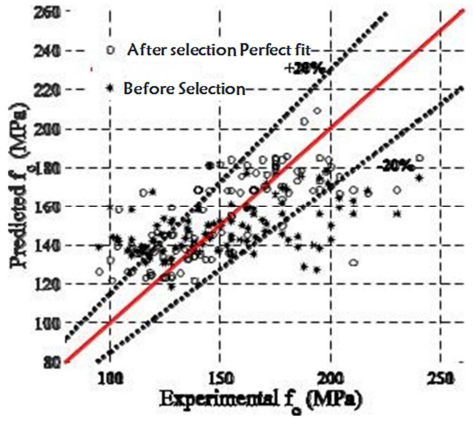
Figure 3. *(a & b): ANN Before and After Selection.*

5. Results and Discussions

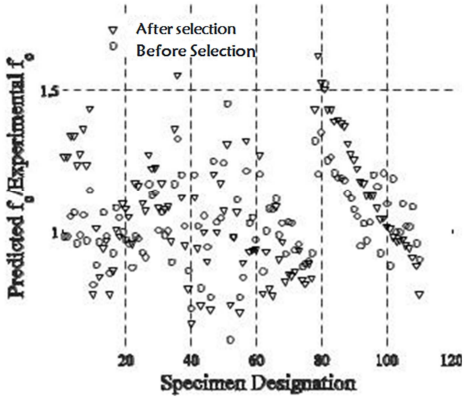
5.1. ANN Results

The selected features, using SFS, were analyzed by the previous BPNN model. As a result, the model that used the selected features showed stronger agreement with the experimental results in contrast with that prior to the selection. Table 3 shows the statistical measurements calculated for both cases. It was observed that the r^2 and NMSE before and after selection yielded 81.6% and 99.1%, respectively, and 0.0594 and 0.026, respectively.

The correlation plots between the predicted and experimental results for the ANN models, with and without selected features using SFS, are summarized in Figure 4(a) presents the percent deviation, where an arbitrary percent deviation was plotted above and below the perfect fit line with a deviation value of ±20%. As a result, the ANN model with the relevant features was capable of predicting 89.6% of its values within the aforementioned boundaries, as opposed to the ANN model with all the features which predicted 58.7% of its values within the boundaries. Figure 4(b) summarizes the compressive strength ratio between both ANN models, where the model with the selected features demonstrated a higher percentage of values (90.6%) ranging between 0.8 and 1.2 than the other model (63.6 %).



(a) Percent Deviation of Predicted versus Experimental Values.



(b) Ratios of Experimental and Predicted Values for Both ANN Models

Figure 4. *Comparison between Experimental and Predicted Values.*

Table 3. *Statistical Measurements Before and After SFS.*

Statistics Measurements	Before Selection	After Selection
r^2	0.816	0.991
NMSE	0.0594	0.026

5.2. Linear Square Regression Analysis

A linear regression model was developed using the Least Square Regression (LSR) method, where the analytical model consisted of the previously selected features. Table 4 shows the coefficient values, with their corresponding symbols, for each UHPC constituent with the statistical measurements of the LSR model. The LSR model is a linear function and its form is shown in (2).

$$fc = \theta 1C + \theta 2SI + \theta 3FA + \theta 4W \quad (2)$$

Table 4. *LSR Coefficients.*

UHPC Constituents	Coefficient Symbols	Coefficients Values	r^2	NMSE
C	θ_1	115.56	0.938	0.0830
SI	θ_2	23.97		
FA	θ_3	5.49		
W	θ_4	27.85		

Figure 5 summarizes the performance of the proposed model, where the model achieved an r2 and NMSE of 0.938 and 0.0830, respectively.

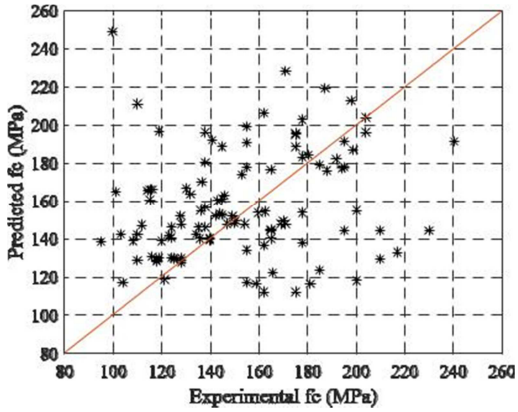


Figure 5. *Summary of the Proposed Model's Performance.*

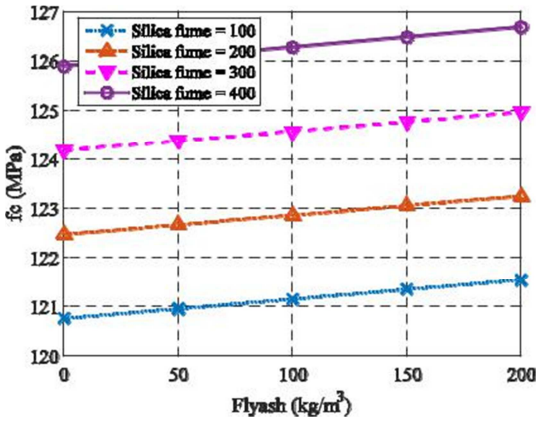


Figure 6. *Variation of compressive strength as a function of Fly Ash and Silica Fume.*

5.3. Parametric Study

Since the developed LSR model is capable of accurately predicting the experimentally measured compressive strength, a parametric study was conducted, using this model, to study the effect of Fly Ash and Silica Fume on the compressive strength of UHPC. Using Fly Ash quantities that range between 0-200 kg/m^3 and Silica Fume quantities that range between 40-160 kg/m^3 while fixing the quantity of cement at 1400 kg/m^3 and water at 175 kg/m^3, several plots showing the variation of strength of UHPC were generated as shown in Figure 5. It is observed from Figure 6 that there is noticeable increase in the compressive strength of UHPC with the increase in Flay Ash and more noticeable with the increase in Silica Fume.

6. Conclusion

This study was conducted to detect the correlation between the material constituents of UHPC and its compressive strength. BPNN was used and three major steps were executed: (1) verification of ANN; (2) application of both SFS; and NID, and (3) analysis of selected features using ANN and LSR. The SFS tool was used to select the relevant constituent that impacted have the most impact on the compressive strength of UHPC which are mainly Cement, Sillica Fume, Flyash, and Water. It can be concluded from this study that:

1) The use of ANN with SFS reduced the number of input parameters needed to accurately predict the compressive strength of UHPC mix for the prediction of compressive strength, making it less computationally expensive.
2) The use of ANN with selected input parameters improved the accuracy of prediction of compressive strength of UHPC and reduced the computational effort. The correlation coefficient (r^2) before and after the use of SFS improved from 81.6% to 99.1% while the NMSE improved from 0.0594 to 0.026, respectively.
3) The ANN model with the selected relevant input parameters also showed a lower deviation (89.6 %) than the ANN model with all the features (58.7%).
4) LSR was implemented using the selected input parameters to develop an analytical model that can be used to accurately predict the compressive strength of UHPC.

References

[1] S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed. Upper Saddle River: Prentice Hall PTR, 1998.

[2] I. Flood and N. Kartam, Artificial Neural Networks for Civil Engineers: Advanced Features and Applications. New York: American Society of Civil Engineers, 1998.

[3] J. A. Abdalla, M. F. Attom, and R. Hawileh, "Prediction of minimum factor of safety against slope failure in clayey soils using artificial neural network," Environ. Earth Sci., vol. 73, no. 9, pp. 5463–5477, 2015.

[4] J. A. Abdalla, M. Attom, R. Hawileh. "Artificial neural network prediction of factor of safety of slope stability of soils," The 14th International Conference on Computing in Civil and Building Engineering, Available at Moscow, Russia, 27-29, 2012.

[5] L. Bal and F. Buyle-Bodin, "Artificial neural network for predicting creep of concrete," Neural Comput. Appl., vol. 25, no. 6, pp. 1359– 1367, 2014.

[6] J. A. Abdalla and R. Hawileh, "Modeling and simulation of low-cycle fatigue life of steel reinforcing bars using artificial neural network," J. Franklin Inst., vol. 348, no. 7, pp. 1393–1403, 2011.

[7] J. A. Abdalla, RA Hawileh. "Artificial neural network predictions of fatigue life of steel bars based on hysteretic energy." Journal of Computing in Civil Engineering Vol. 27, no. 5, pp. 489-496, 2013.

[8] J. Abdalla and R. Hawileh, "Energy-based predictions of number of reversals to fatigue failure of steel bars using artificial neural network," in The 13th International Conference on Computing in Civil and Building Engineering, 2010.

[9] J. A. Abdalla, A. Elsanosi, and A. Abdelwahab, "Modeling and simulation of shear resistance of R/C beams using artificial neural network," J. Franklin Inst., vol. 344, no. 5, pp. 741–756, 2007.

[10] J. A. Abdalla, A. Abdelwahab, A. Sanousi. "Prediction of shear strength of reinforced concrete beams using artificial neural network," The First International Conference on Modeling, Simulation and Applied Optimization, Sharjah, U. A. E., 2005.

[11] H. M. Tanarslan, M. Secer, and A. Kumanlioglu, "An approach for estimating the capacity of RC beams strengthened in shear with FRP reinforcements using artificial neural networks," Constr. Build. Mater., vol. 30, pp. 556–568, 2012.

[12] J. A. Abdalla, R. Hawileh, A. Al-Tamimi. "Prediction of FRP-concrete ultimate bond strength using Artificial Neural Network," International Conference on Modeling, Simulation and Applied Optimization (ICMSAO), Kuala Lumpur, Malaysia, 2011.

[13] J. A. Abdalla, E. I. Saqan, R. A. Hawileh. "Optimum seismic design of unbonded post-tensioned precast concrete walls using ANN," Computers and Concrete, Vol. 13, no. 4, pp. 547-567, 2014.

[14] F. Deng, Y. He, S. Zhou, Y. Yu, H. Cheng, and X. Wu, "Compressive strength prediction of recycled concrete based on deep learning," Constr. Build. Mater., vol. 175, pp. 562–569, 2018.

[15] H. Naderpour, A. H. Rafiean, and P. Fakharian, "Compressive strength prediction of environmentally friendly concrete using artificial neural networks," J. Build. Eng., vol. 16, no. January, pp. 213–219, 2018.

[16] A. Heidari, M. Hashempour, and D. Tavakoli, "Using of Backpropagation Neural Network in Estimating of Compressive Strength of Waste Concrete," vol. 1, pp. 54–64, 2017.

[17] R. Gupta, M. A. Kewalramani, and A. Goel, "Prediction of Concrete Strength Using Neural-Expert System," J. Mater. Civ. Eng., vol. 18, no. 3, pp. 462–466, 2006.

[18] S. C. Lee. Prediction of concrete strength using artificial neural networks. Engineering structures; Vo. 25, No. 7:849-857, 2003.

[19] E. Ghafari, M. Bandarabadi, H. Costa, and E. Júlio, "Prediction of Fresh and Hardened State Properties of UHPC : Comparative Study of Statistical Mixture Design and an Artificial Neural Network Model," vol. 27, no. 11, pp. 1–11, 2015.

[20] G. Tayfur, T. K. Erdem, and K. Önder, "Strength Prediction of High-Strength Concrete by Fuzzy Logic and Artificial Neural Networks," J. Mater. Civ. Eng., vol. 26, no. 11, pp. 1–7, 2014.

[21] A. Torre, F. Garcia, I. Moromi, P. Espinoza, L. Acuña. Prediction of compression strength of high performance concrete using artificial neural networks. VII International Congress of Engineering Physics, Journal of Physics: Conference Series 582, 2015.

[22] A. E. N. Kay Wille and Gustavo J. Parra-Montesinos, "Ultra-High Performance Concrete with Compressive Strength Exceeding 150 MPa (22 ksi): A Simpler Way," Mater. J., vol. 108, no. 1.

[23] Y. W. W. Zheng, B. Luo, "Compressive and tensile properties of reactive powder concrete with steel fibres at elevated temperatures,"Constr. Build. Mater., vol. 41, pp. 844–851, 2013.

[24] K.-Q. Yu, J.-T. Yu, J.-G. Dai, Z.-D. Lu, and S. P. Shah, "Development of ultra-high performance engineered cementitious composites using polyethylene (PE) fibers," Constr. Build. Mater., vol. 158, pp. 217–227, 2018.

[25] S. Abbas, M. L. Nehdi, and M. A. Saleem, "Ultra-High Performance Concrete: Mechanical Performance, Durability, Sustainability and Implementation Challenges," Int. J. Concr. Struct. Mater., vol. 10, no. 3, pp. 271–295, 2016.

[26] N. Soliman and A. Tagnit-Hamou, Using glass sand as an alternative for quartz sand in UHPC, vol. 145. 2017.

[27] N. Gu, M. Fan, L. Du, and D. Ren, "Efficient sequential feature selection based on adaptive eigenspace model," Neurocomputing, vol. 161, pp. 199–209, 2015.

[28] R. Kohavi and G. H. John, "Wrappers for feature subset selection," Artif. Intell., vol. 97, no. 1–2, pp. 273–324, 1997.

[29] H. Zhou, M. You, L. Liu, and C. Zhuang, "Sequential data feature selection for human motion recognition via Markov blanket," Pattern Recognit. Lett., vol. 86, pp. 18–25, 2017.

[30] V. F. Rodriguez-Galiano, J. A. Luque-Espinar, M. Chica-Olmo, and M. P. Mendes, "Feature selection approaches for predictive modelling of groundwater nitrate pollution: An evaluation of filters, embedded and wrapper methods," Sci. Total Environ., vol. 624, pp. 661–672, 2018. S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed. Upper Saddle River: Prentice Hall PTR, 1998.