

Development of Prediction models for Bond Strength of Steel Fiber Reinforced Concrete by Computational Machine Learning

Priyanka Singh¹, Chakshu Garg², Aman Namdeo¹, Krishna Mohan Agarwal³, and Rajesh Kumar Rai⁴

¹Department of Civil Engineering, Amity School of Engineering & Technology, Amity University Uttar Pradesh, Noida, India

²Department of Computer Science & Engineering, Amity School of Engineering & Technology, Amity University Uttar Pradesh, Noida, India

³Department of Mechanical Engineering, Amity School of Engineering & Technology, Amity University Uttar Pradesh, Noida, India

⁴Department of Agricultural Engineering, Udai Pratap College, Varanasi, Uttar Pradesh-221 002, India

Abstract. Sustainable construction contributed to the usage of recycled and waste materials to substitute conventional concrete. This research focuses on prediction of normalized bond strength of cement concrete substituted by large amounts of waste materials and products with strong mechanical properties and sustainability. It also emphasizes on using analytical model for the prediction of bond strength of the green concrete, so that there is a reduction in the cost of construction, conserve energy, and it will lead to a reduction of CO₂ production from cement industries within reliable limits. In this paper machine learning approach has been used to predict the normalized bond strength of green and sustainable concrete. Machine learning empowers machines to learn from their experiences and data provided. The system analyses the datasets and finds different patterns formed in the given data. Then, based on its learnings the machine can make certain predictions. In civil engineering application, a special computing technique called the Machine learning (ML) is in huge demand. ANN is a soft computing technique that learns from previous situations and adapts without constraints to a new environment. In this work, a ML network model for prediction of normalized bond strength of concrete has been illustrated. Different sets of data based upon several concrete design mixes were taken from technical literature and were fed to the model. The model is then trained for prediction, which are being influenced by several input attributes and were jotted down a linear regression analysis.

1 Introduction

Machine learning is an area of study which helps computers or systems to learn from their experiences and improve. Arthur Samuel defines machine learning as “the field of study that gives computers the ability to learn without being explicitly programmed”. This definition given by Arthur Samuel is not a very formal definition of machine learning [1,3]. Therefore a relatively new definition is given by Tom Mitchell, it says, “A computer program is said to learn from experience E with respect to some class of task T, and a performance measure P, if its performance at tasks in T, as measured by P, improves with experience E”. For example, two digits are paired, the user inputs one digit to expect the other digit from the machine. The machine then has to identify the logic between the pairs and give the other value as a prediction to the user [4, 7]. This process of finding or evaluating the logic, and learning from experiences is what machine learning is all about.

Machine learning is a technology that emphasises on learning from data. The system analyses the datasets and finds different patterns formed in the given data. Then,

based on its learnings the machine can make certain predictions [8].

There are many different approaches that can be used for machine learning. The approaches commonly used in machine learning are supervised learning, semi-supervised learning, unsupervised learning and reinforcement learning. In this work data has been collected experimentally and has been analysed in order to apply the machine learning techniques. The data comprises of both the input features and the output features. Therefore, supervised learning became handy.

2 Experimental Setup

The dataset obtained was not in the desired form so as to apply different machine learning algorithms. So proper data preprocessing was performed. Initially the dataset was normalised, thereafter outliers were removed and missing data was handled. After all the cleaning and pre-processing was done the dataset contained 361 instances. This cleaned and pre-processed dataset was then used to apply machine learning algorithms so

that the Normalised Bond Strength [NBS] of steel could be predicted [9,12].

In order to make predictions for the normalised bond strength of concrete, NBS was considered as an output parameter and strength grade, f_c cube, fibres, volume fraction, types of aggregate, admixture, admixture content (in percentage), specimen geometry, diameter, bond length, length to diameter, concrete cover, cover to diameter, type of bar, age at testing, maximum temperature, time at maximum, cooling, type of test, temperature variation were considered as input parameters [11,13]. So there were a total of 21 features, 20 being input features and 1 being the output feature that needs to be predicted as shown in Table 1.

We used different machine learning algorithms like elastic net regression, lasso regression, ridge regression, support vector regression, random forest regression, multiple regression and CART regression to train our machine. We found that multiple regression and CART regression gave the best results. In this paper we will be discussing about the results obtained by training the machine using these algorithms [14,16].

Table 1. Details of Input and Output Parameters.

Input	Parameter
X1	Strength Grade
X2	f_c , cube, 20 degree Celsius [MPa]
X3	Fibres
X4	Volume Fraction
X5	Type of Aggregate
X6	Admixture
X7	Admixture Content [%]
X8	Specimen Geometry
X9	Diameter [mm]
X10	Bond Length [mm]
X11	Length to Diameter [l/d]
X12	Concrete Cover [mm]
X13	Cover to Diameter [c/d]
X14	Type of Bar
X15	Age at Testing [days]
X16	Maximum Temperature [degree Celsius]
X17	Time at Maximum Temperature [h]
X18	Cooling
X19	Type of Test
X20	Temperature Variation
Output	Parameter
Y1	Normalised Bond Strength [NBS]

3 Multiple Linear Regression

Multiple linear regression is a technique to find a relationship between one dependent variable and two or more independent variables. In this case we have one dependent variable also known as the response variable and we have 20 independent variables that we call predictors. The coefficient of each predictor in the equation depicts its contribution in predicting the response value. Using Minitab we can derive an equation showing the relationship between these parameters so that normalised bond strength of steel could be estimated.

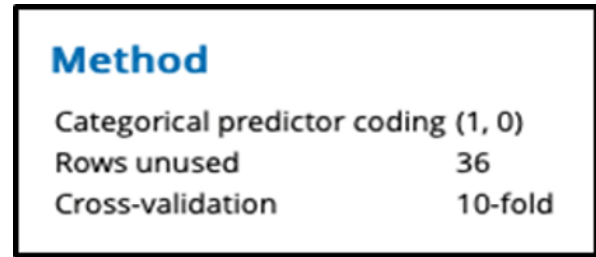


Fig. 1. Approach to apply multiple linear regression.

Figure 1 depicts the methodology used to evaluate the NBS. We used (1,0) coding for categorical predictors. We also used 10 fold cross validation to ensure that our model is neither overfitting the data nor is it under fitting the data.

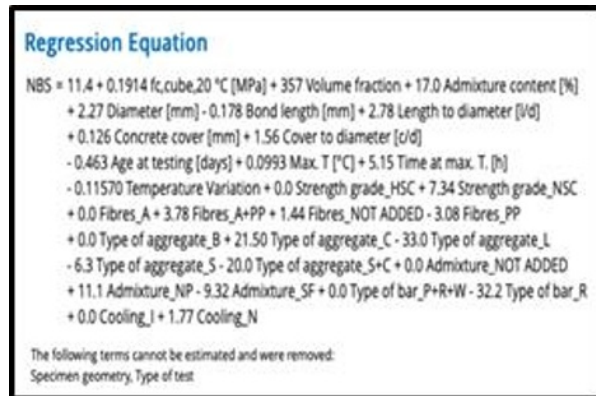


Fig. 2. Regression equation.

Figure 2 shows the equation derived for estimation of NBS using multiple linear regression.

The dataset taken in to consideration for analysis is normally distributed hence analysis of variance is done to analyse the differences in prediction from the proposed input parameters fitted in the model. Adjusted Sum of Squares measures the variation in the output and is clarified by individual unit of model. Adjusted mean squares quantifies how much variety a term clarifies, expecting that every single other term are in the model irrespective of their order of entry, it is the variance around fitted values.

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	24	224314	9346	101.88	0.000
fc,cube,20 °C [MPa]	1	937	937	10.22	0.002
Volume fraction	1	108	108	1.18	0.278
Admixture content [%]	1	4	4	0.05	0.827
Diameter [mm]	1	59	59	0.65	0.422
Bond length [mm]	1	151	151	1.64	0.201
Length to diameter [l/d]	1	560	560	6.10	0.014
Concrete cover [mm]	1	6	6	0.07	0.794
Cover to diameter [c/d]	1	5	5	0.05	0.819
Age at testing [days]	1	379	379	4.13	0.043
Max. T [°C]	1	177	177	1.92	0.166
Time at max. T. [h]	1	29	29	0.32	0.575
Temperature Variation	1	169770	169770	1850.59	0.000
Strength grade	1	437	437	4.76	0.030
Fibres	3	301	100	1.09	0.353
Type of aggregate	4	3769	942	10.27	0.000
Admixture	2	724	362	3.95	0.020
Type of bar	1	246	246	2.68	0.103
Cooling	1	118	118	1.29	0.257
Error	300	27521	92		
Lack-of-Fit	288	26810	93	1.57	0.190
Pure Error	12	711	59		
Total	324	251836			

Fig. 3. Analysis of variance.

Figure 3 shows the analysis of variance which clearly depicts that temperature variation, types of aggregate, fc cube, admixture, strength grade, length to diameter, age at testing are variables that have a significant effect on the estimation of NBS.

Model Summary					
S	R-sq	R-sq(adj)	R-sq(pred)	10-fold S	10-fold R-sq
9.57801	89.07%	88.20%	86.95%	10.1077	86.82%

Fig. 4. Model summary.

The model obtained after training the machine was around 87% accurate as shown in the model summary, figure 4.

Normal probability plot of residuals is plotted to verify the assumptions are normally distributed.

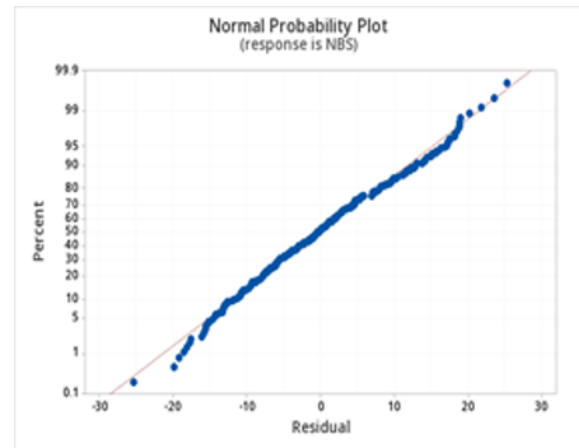


Fig. 5. Normal probability plot.

The plot obtained for our model shown in figure 5 depicts the accuracy of our model.

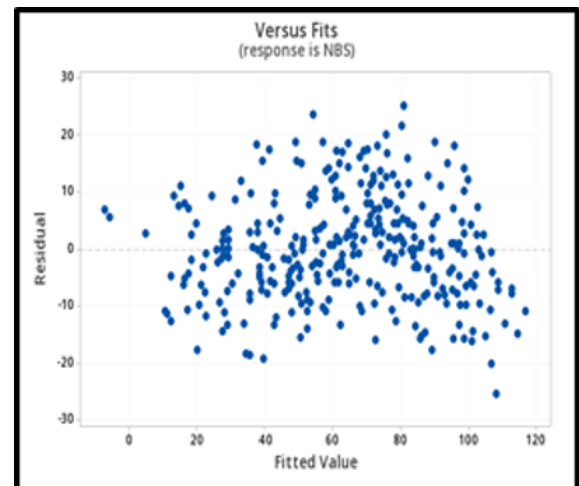


Fig. 6. Residuals versus fits plot.

Residuals versus fits plot given in figure 6 verifies assumption that the residuals have an acceptable variance.

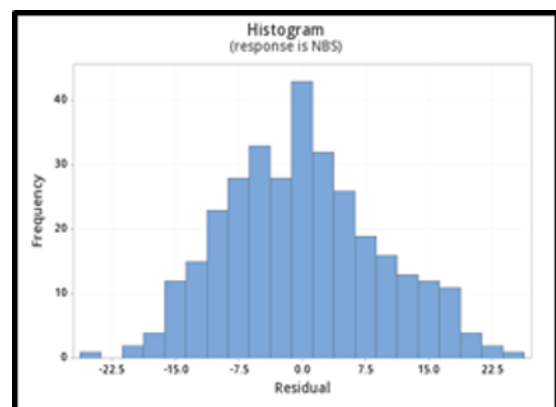


Fig. 7. Histogram of residual plot.

The plot we obtained as shown in figure 7 is the histogram of residual which is used to determine the skewness of data and outlier's existence in the data.

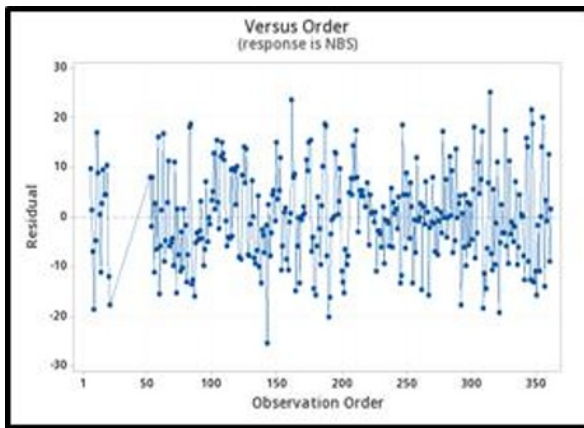


Fig. 8. Residual versus order plot.

Residual versus order of data verifies assumption that the residuals are uncorrelated with each other as shown in figure 8.

4 CART Regression

CART stands for Classification And Regression Trees. It refers to a decision tree model used for predictive modelling. The representation of this model is in the form of a binary tree. It can be used for both classification and regression. Since our output variable is continuous we used CART regression for predicting the normalised bond strength of steel.

While training the machine we used strength grade, fibres, type of aggregate, admixture, specimen geometry, type of bar, cooling and type of test as categorical parameters and the remaining inputs as continuous parameters for predicting the NBS.

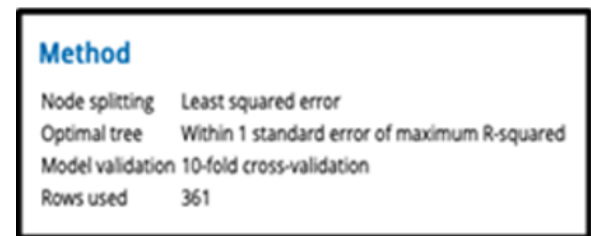


Fig. 9. Approach to apply CART regression.

The method that we applied to make predictions using CART regression is mentioned in figure 9. We used least squared error for node splitting and we applied 10 fold cross validation to avoid overfitting. For evaluating the model we used various parameters like R squared value and root mean squared error. or ideal condition, R - squared value should be 1 and RMSE value should be 0.

The tree obtained by using this approach is shown in figure 10.

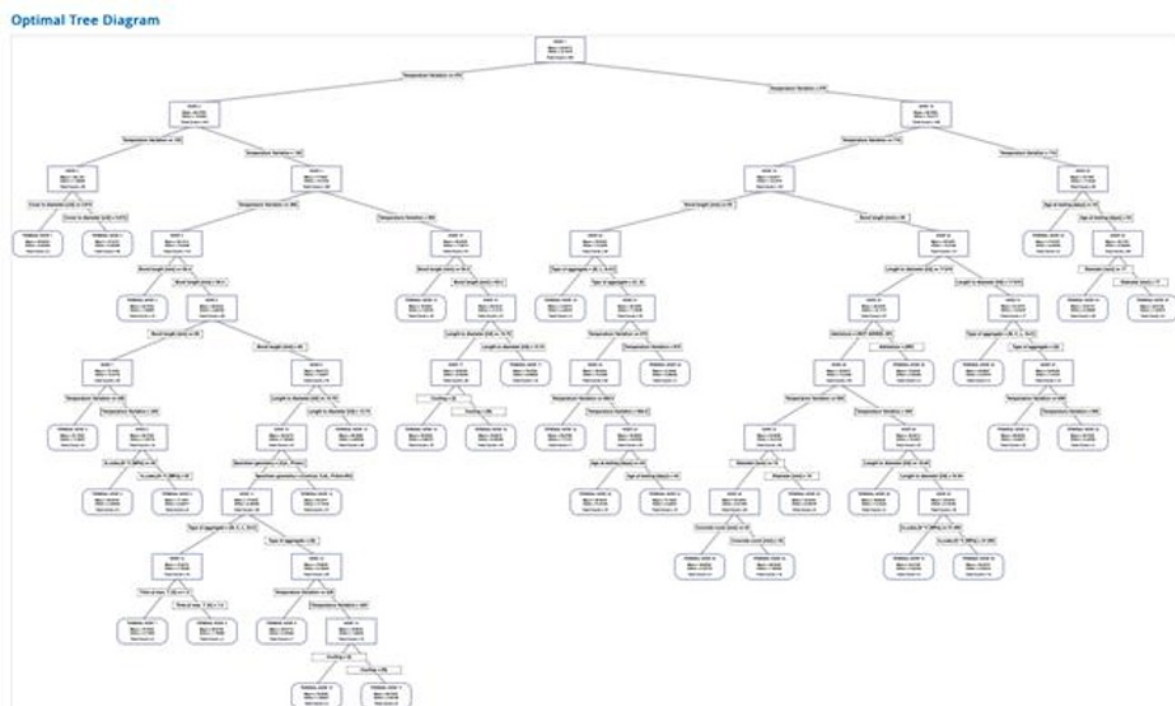


Fig. 10. Optimal Tree Diagram.

Model Summary		
Total predictors	20	
Important predictors	20	
Number of terminal nodes	35	
Minimum terminal node size	3	
Statistics	Training	Test
R-squared	0.9348	0.8622
Root mean squared error (RMSE)	7.0068	10.1873
Mean squared error (MSE)	49.0947	103.7805
Mean absolute deviation (MAD)	5.4526	7.7110
Mean absolute percent error (MAPE)	0.1482	0.2044

Fig. 11. Model summary.

The model that was applied for predicting NBS used 35 terminal nodes. It was more than 86 percent accurate as depicted by the R - squared and the RMSE values in figure 11. The figure also shows that all the input parameters are important for making predictions for normalised bond strength of steel.

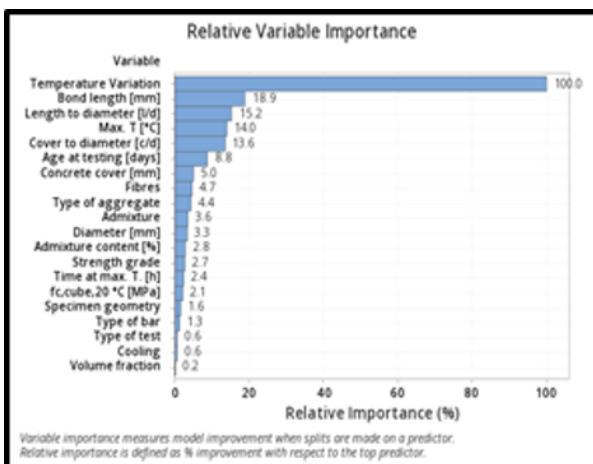


Fig. 12. Relative variable importance.

Figure 12 shows the relative importance of all the predictors.

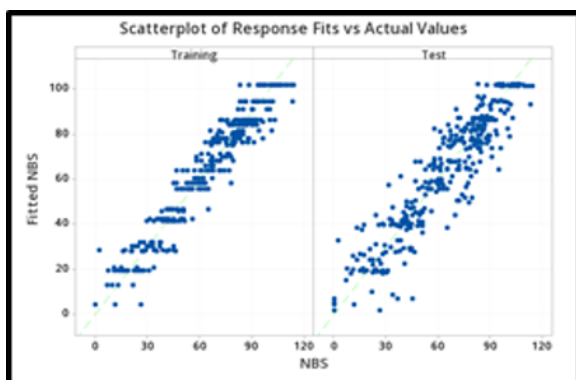


Fig. 13. Scatterplot of response fits versus actual values.

Figure 13 shows the goodness of fit of our regression model.

5 Results and Discussion

For multiple linear regression model variable importance in the decreasing order was temperature variation, type of aggregate, fc cube, length to diameter, admixture, strength grade and age at testing. The other features were not very significant while predicting NBS. While making predictions using the CART regression model the variable importance obtained in decreasing order was temperature variation, bond length, length to diameter, maximum temperature, cover to diameter, concrete cover, fibres, type of aggregate, admixture, diameter, admixture content, strength grade, time at maximum temperature, fc cube, specimen geometry, type of bar, type of test, cooling and volume fraction.

Multiple linear regression model didn't consider specimen geometry and type of test in regression equation. Whereas all of the 20 parameters were considered important for making predictions using CART regressor.

6 Conclusion

In this study we developed two machine learning models to predict the normalised bond strength of steel. The first model was multiple linear regression model and the other was CART regression model. Both the models were equally good as multiple linear regression model showed on accuracy of 86.82 percent and the CART regression model was 86.22 percent precise as depicted by their R-squared values. The models were developed utilising ANN and linear regression models from a reference dataset. The model with its respective inputs has been checked for its output with the regression equation to forecast the bond strength. The plot of residual data for bond strength is built to evaluate the fitness level after regression for the various sets of input parameters, deciding whether the ordinary least square assumptions are satisfied if presumptions are met, ordinary lowest square regression results unbiased coefficient estimates with the least variance. In machine learning, the ideal value of RMSE is 0 and that of the R squared method is 1. This model, which is about 86.82 percent accurate, can predict bond strength precisely in short time. The ANN models developed will contribute to timesaving, reducing waste material and decreasing the project's overall cost.

References

1. P. Singh and P. Khaskil, "Prediction of Compressive Strength of Green Concrete with Admixtures Using Neural Networks," 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, 2020, pp. 714-717, doi: 10.1109/GUCON48875.2020.9231230.

2. Golafshani, Emadaldin Mohammadi, Alireza Rahai, Mohammad Hassan Sebt, and Hamed Akbarpour. "Prediction of bond strength of spliced steel bars in concrete using artificial neural network and fuzzy logic." *Construction and building materials* 36 (2012): 411-418.
3. Yan, Fei, Zhibin Lin, Xingyu Wang, Fardad Azarmi, and Konstantin Sobolev. "Evaluation and prediction of bond strength of GFRP-bar reinforced concrete using artificial neural network optimized with genetic algorithm." *Composite Structures* 161 (2017): 441-452.
4. Dahou, Zohra, Z. Mehdi Sbartaï, Arnaud Castel, and Fouad Ghomari. "Artificial neural network model for steel-concrete bond prediction." *Engineering Structures* 31, no. 8 (2009): 1724-1733.
5. Sancak, Emre. "Prediction of bond strength of lightweight concretes by using artificial neural networks." *Scientific Research and Essays* 4, no. 4 (2009): 256-266.
6. Ahmad, Sohaib, Kypros Pilakoutas, Muhammad M. Rafi, and Qaiser U. Zaman. "Bond strength prediction of steel bars in low strength concrete by using ANN." *Computers and Concrete* 22, no. 2 (2018): 249-259.
7. Dantas, Adriana Trocoli Abdon, Monica Batista Leite, and Koji de Jesus Nagahama. "Prediction of compressive strength of concrete containing construction and demolition waste using artificial neural networks." *Construction and Building Materials* 38 (2013): 717-722.
8. Duan, Zhen-Hua, Shi-Cong Kou, and Chi-Sun Poon. "Prediction of compressive strength of recycled aggregate concrete using artificial neural networks." *Construction and Building Materials* 40 (2013): 1200-1206.
9. Khademi, Faezehossadat, Sayed Mohammadmehdi Jamal, Neela Deshpande, and Shreenivas Londhe. "Predicting strength of recycled aggregate concrete using artificial neural network, adaptive neuro-fuzzy inference system and multiple linear regression." *International Journal of Sustainable Built Environment* 5, no. 2 (2016): 355-369.
10. Vakhshouri, Behnam, and Shami Nejadi. "Prediction of compressive strength of self-compacting concrete by ANFIS models." *Neurocomputing* 280 (2018): 13-22.
11. Singh, Priyanka Ramkripal, Aishwarya Goel, Shailendra Thakur, and N. D. Shah. "An experimental approach to investigate effect of steel fibers on tensile and flexural strength of fly ash concrete." *Int J Sci Eng Appl Sci (IJSEAS)* 2, no. 5 (2016): 384-392.
12. Singh, Priyanka, and Niraj D. Shah. "An experimental investigation on sustainable concrete with flyash and steel fibers." *Int J Civil Eng Technoly* 9.6 (2018): 1131-1140.
13. Singh, Priyanka Ramkripal, and N. D. Shah. "Impact of coal combustion fly ash used as a binder in pavement." *Civ. Eng. Environ. Tech* 1 (2014): 57-60.
14. S. Dixit, A. Stefan, A. Musiuk, and P. Singh, "Study of enabling factors affecting the adoption of ICT in the Indian built environment sector," *Ain Shams Engineering Journal*, no. xxxx, 2020.
15. S. Dixit, A. Stefan, and A. Musiuk, "Architectural form finding in arboreal supporting structure optimisation," *Ain Shams Engineering Journal*, no. xxxx, 2020.
16. Wang, Yi, Zong Woo Geem, and Kohei Nagai. "Bond strength assessment of concrete-corroded rebar interface using artificial neural network." *Applied Sciences* 10, no. 14 (2020): 4724.