Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 3): pp.4093-4101
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.03.421

# An Improved Machine Learning-Based Compressive Strength Prediction For Concrete In Civil Structures

**Sanjeev Kumar** Department of Civil Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India, 248002 er.sanjeevkr@gmail.com

**Bhasker Pant** Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India, 248002 pantbhasker2@gmail.com

**Anoop Bahuguna** Department of Civil Engineering, Graphic Era Hill University, Uttarakhand India abahuguna@gehu.ac.in

## ABSTRACT

An innovative method for estimating the compressible strength of concrete is suggested in this study and is dependent on the machine-learning technology. This method uses the dynamic boosted technique to combine numerous weak classifiers into a strong classifier that can identify the connection between input and values. The total reliability of the strong classifier will be improved since the weak classifier with the tiny forecasting mistake will be given more value in the network. To train and assess the learners, a maximum of 1030 frames of tangible compression tests have been gathered, where the input signal are the elements of the cementitious material (such as coarse/fine agglomerates, concretes, moisture, thickeners, etc.) and the healing time, and the data flow are the fracture toughness values. The suggested technique produces an average reliability of above 95% in notion of coefficient of determination ( $r^2$) and is verified using a ten-fold cross validation procedure. In order to show the suggested mode's generalizability, a fresh collection of 103 trials for compressive strength in concrete is also employed. Artificial neural network (ANN) and support vector machine (SVM) are two additional separate machine learning techniques that are currently being used in this sector, and the suggested methodology outperforms existing approaches in every way. The impact of various important adaptive boosting technique components, including as the quantity of training examples, the selection of base learners, and the impact of the sensitivities and variety of input characteristics, is also examined in the last section. It is demonstrated that the logistic regression is the best option for the generalization error in the boost architecture when utilising 80% of the complete data as training data. Additionally, based on the findings of the parameter estimation, the significance of various input factors is determined.

## INTRODUCTION

The biomechanical features of concrete must be thoroughly studied in order to provide appropriate techniques. Compressive strength is perhaps the most vital part of the different cement property index values because it significantly influences the safeness of the frameworks and is necessary for determining the achievement of the frameworks throughout their entire lifespan, from innovative architectural engineering to old detailed inspection. But as everyone is aware, cement is manufactured up of many ingredients, such as coarse/fine aggregate, cementing pastes, extra combinations, etc., and these ingredients are dispersed at randomly across the whole cement structure. Correctly predicting the compression strength of a concrete stuff is quite difficult due to the complexity of the mechanism [1,2]. Physical investigations are often the most straightforward technique to determine the concrete's compressive strength. Typically, the cube or cylinders exhibits were made using a certain planned suitable method and then allowed to cure for the necessary period of time. The compressive strength may then be readily acquired using the ductility test device, as shown in [3-5]. However, this strategy is expensive in both terms of money and energy, therefore it won't function very well. In addition to the classic experimental approaches, [6–8] proposes several inductive logistic regressions to forecast the concrete's compression strength given the specified mix proportion of the multiple elements. However, there is a weakly nonlinear relationship between the cementitious material and strength properties, making it challenging to obtain an adequate linear relationship for this issue. Computer modelling is the third method of encapsulating concrete behaviour (see [9,10], etc.). As previously indicated, it is difficult to faithfully mimic the actual behaviour since randomness and nonlinearity are coupled [1,11]. However, since artificial intelligence (AI) has advanced recently, there is a tendency to apply machine learning (ML) methods to forecast the compressive strength of concrete. AI's machine learning (ML) division may be used for a variety of tasks, including categorization, regression, clustering, etc. One use of the regression model of ML is to predict the compressive strength of concrete. A clear benefit of machine learning (ML) over other classic regression approaches is that it uses specific algorithms that may adapt from either the data input itself and provide very precise findings for the data output [12].

## LITERATURE REVIEW

The most popular ML techniques for predicting concrete compressive strength are artificial neural networks (ANN) and support vector machines. (SVM). Siddique et al[13] .'s use of ANN to forecast compressive strength of bottom ash-containing self-compacting concrete is only one example. In order to estimate the strength properties of self-compacting concrete after exposure to high temperatures, Uysal and Tanyildizi [14] used ANN. For the purpose of predicting the strength of concrete including construction debris or recycled aggregate concrete, Dantas et al. [15] and Duan [16] utilised ANN. In order to estimate the compressive strength of concrete, Chou et al. [17,18] looked into a number of ML approaches, including ANN and SVM. Least square SVM, a more sophisticated variation of SVM, was presented to this subject by Aiyer et al. [19]. With regard to a more challenging issue, namely the unrestrained compressive and flexural of cockle shell "cement" sand mixes, Motamedi et al. [20] expanded the SVM-based prediction technique. When predicting the compression strength of advanced concrete, Pham et al. [21]

significantly enhanced the least square SVM utilising a metaheuristic optimization method. In order to forecast the strength properties of environmentally friendly concrete, Omran et al. [22] examined the precision of several data mining algorithms. In order to forecast the concrete's compressive strength containing copper slag and nano silica, Chithra et al. [23] examined the use of ANN. Naderpour [24] also employed ANN to foretell the durability of green concrete. In order to forecast the strength and acoustic pulse velocity of fibre reinforced concrete, Ashrafian [25] employed heuristic regression techniques. In order to forecast the compressive strength, Zhang et al. [26] employed random forest (RF) and spoke about the significance of each input variable. The majority of the algorithms listed above, however, are for individual learning, as opposed to the next series of ensembles ml algorithms [27], which are more precise, reliable, and potent. Ensemble learning models' fundamental premise is to build a strong learner by first integrating a number of weak learners into it. The personalized learning algorithms, such as ANN or SVM, form the foundation for the weak learners. As a result, collective ml algorithms (specifically, the robust learner) will certainly have better predictability and robustness. The review study [28] details the distinctions between the three main classes of techniques for supervised methods, namely bagging, boosting, and stacking. Adaptive boosting (AdaBoost) [29] is probably the most popular method for ensemble learning among the several techniques available. In this research, we suggested a novel intelligent method for forecasting the concrete's compressive strength depending on the AdaBoost algorithm. The mixture components and setting period are specified as the required information and the compressive is specified as the data output for a maximum of 1030 pairs of aggregate compression tests.

The introduction of AdaBoost's essential mathematical foundation and the associated implementation follows. The AdaBoost-based model may be trained using the gathered data to produce a powerful learner that can be used to forecast the compressive strength value. The created model is validated using ten-fold cross validation approach, and the generalizability of the model is also shown using a fresh data set of 103 samples. The AdaBoost model's performance is also contrasted with that of earlier individual machine learning models, such as ANN-based and SVM-based, demonstrating the supremacy of the ensembles learning approach. The impacts of various important AdaBoost technique components, including as the quantity of training data, the kind of weak learner, and the sensitivities and variety of input variables, are also examined in this last section.

## PROPOSED SYSTEM

The AdaBoost algorithm, or more precisely, the augmenting series of ensembles learning algorithms, is the most common and commonly used algorithm in the ensemble learning approach. AdaBoost's unique feature is that it creates a weak learner from the first training data and then modifies the dispersion of the data sets for the subsequent weak learner training round in response to the predicted performance. Keep in mind that in the next stage, greater attention will be paid to the training instances with poor prior prediction accuracy. Finally, the weaker learners are combined with varying weights to create a powerful learner. The following will be a thorough mathematical basis and implementation method. Regression and classification tasks may both be accomplished using the AdaBoost. In this study, we use AdaBoost regression to forecast the compressive strength of concrete. The train data set H may be represented as follows when considering a generic regression problem:

$$\theta = \{(X_1, Y_1), (X_2, Y_2)\} \qquad (1)$$

For simplicity, the linear loss function is adopted, namely,

$$E_I = \left|\frac{Y_i - G(X_i)}{E}\right| \qquad (2)$$

AdaBoost can be implemented quickly and simply. Generally speaking, it consists of four stages: (1) gathering experimental data; (2) developing a strong learner; (3) testing or validating the learner; and (4) applying the learner to engineering challenges. The second level is undoubtedly the foundation of AdaBoost. It truly comprises two stages, as was previously said, namely a framework to combine weak learners with strong ones and a regressive classification algorithm to create the weak learners using training data. The DT method, or more precisely the classification and regression tree (CART) [33], is utilised to produce the weak learner, and the weak learners are integrated by the median of the weighted weak learners. Fig. 1 depicts the workflow for this process.

The AdaBoost parameters also comprise two levels: first layer refers to the AdaBoost architecture, and the other level is for the weak learner's algorithm, or CART. There are just two parameters in the framework: the regularisation constant (or integral gain) m and the number of weak learners (or maximum number of iterations, N). The CART has a number of parameters and is somewhat more difficult. The most important variables for the compressive strength of concrete prediction issue with small features are the actual measured blow roots, the minimum split samples, the minimum leaf node samples, and the minimal impurities for further split. Here, an optimization approach is used to choose the aforementioned parameters in order to attain the optimum model performance. The initial ranges of values for these parameters are established using data from other studies in related fields and the literature, such as [18,34]. Next, a grid search approach is used to locate the precise values; for example, a parameter grid is defined from the starting value ranges, and the model is repeatedly trained and tested utilizing cross validation to identify the parameter set that performs the best.

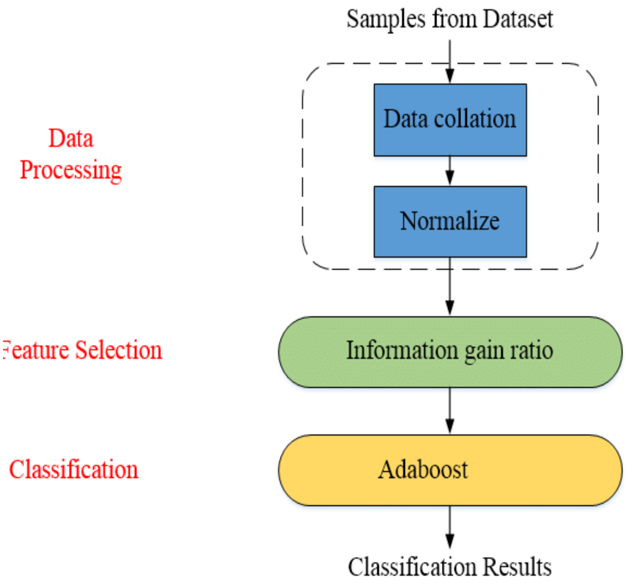


**Figure 1. Flow chart of the proposed system**

## RESULTS AND DISCUSSION

A dataset from the UCI Repository will form the foundation of this research. The dataset has 1030 observations organised into 9 characteristics. Eight quantitative inputs and one quantitative output make up the characteristics. There are no blank values in the dataset. A concrete's compressive strength is the main topic of the dataset. The characteristics include elements that have an impact on concrete's strength, such as fly ash, water, coarse and fine aggregate, and cement. This project's goal is to forecast concrete compressive strength using significant predictors. The purpose of the research is to determine how various elements, such as fly ash, cement, water, age, and possibly additives, affect the final product. Through visualisation or a correlation matrix, we will assess the factors that have a strong association to concrete compressive strength as well as those that are less important and may be overlooked. In this work, we will use several machine learning methods to forecast the compressive strength of concrete. For the prediction, many modelling approaches will be applied. Multi-linear regression, decision trees, and random forests, among other modelling techniques, will be used. The most accurate model for our forecast will be found via a comparison study. For the building of bridges and homes, the finest modelling will assist civil engineers in selecting the proper concrete.

| Cement | blast_furnace_slag | fly_ash | water | superplasticizer | coarse_aggregate | fine_aggregate | Age | concrete_compressive_strength |
|---|---|---|---|---|---|---|---|---|
| 0 | 540.0 | 0.0 | 0.0 | 162.0 | 2.5 | 1040.0 | 28 | 79.99 |
| 1 | 540.0 | 0.0 | 0.0 | 162.0 | 2.5 | 1055.0 | 28 | 61.89 |
| 2 | 332.5 | 142.5 | 0.0 | 228.0 | 0.0 | 932.0 | 270 | 270 |
| 3 | 198.6 | 132.4 | 0.0 | 192.0 | 0.0 | 978.4 | 365 | 360 |

Table 1. Prediction values of different materials

Different true positive values and true negative values are plotted in the confusion matrix as shown in figure 2. Figure 3 and Figure 4 distribution and scatter plot between various compoenents.

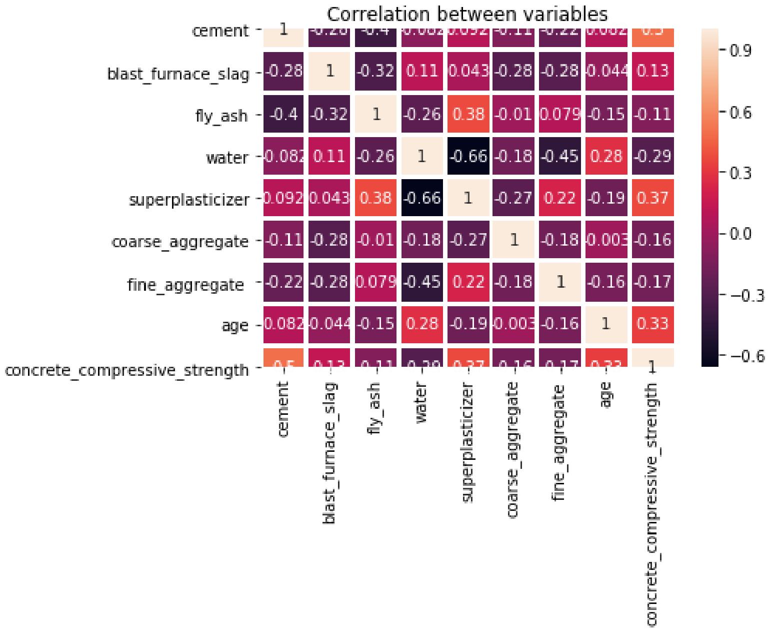


Figure 2. Confusion Matrix

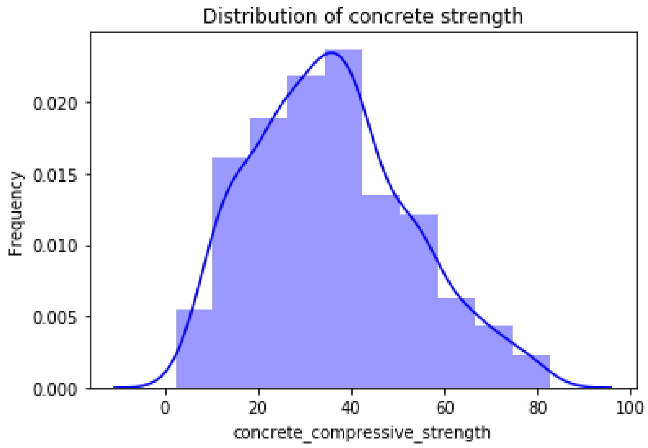


Figure 3. Distribution of components of concrete

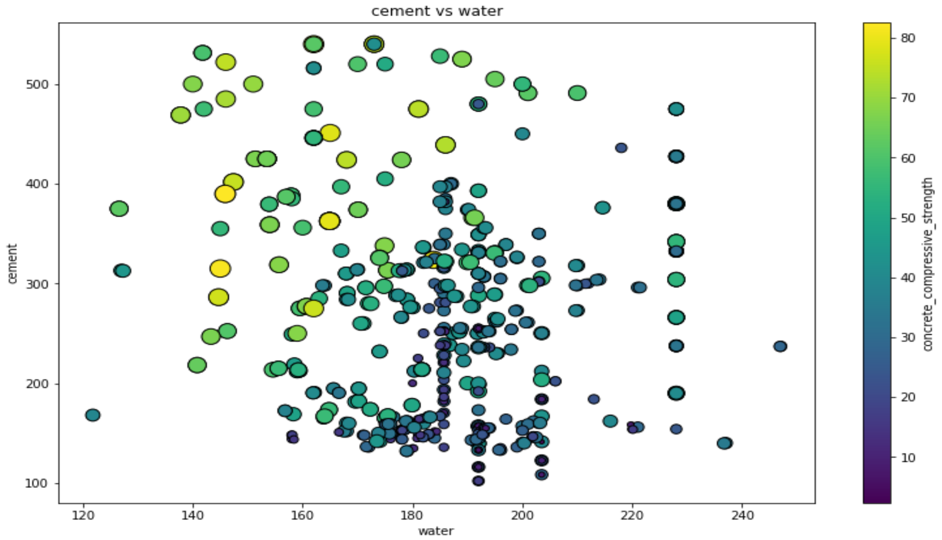

Figure 4. Scatter plot of different components

## Conclusion

With the AdaBoost model, we can easily ascertain the cement's compressive strength at distinct periods and with different control mixes. If the anticipated mix fraction can satisfy the goal structural requirements, it may be utilized to make that determination. Alternately, it can calculate the fracture toughness population increase at different ages, helping experts with duties like safety studies during the service stage and clearance assessments during the construction phase. It is also possible to use this model as a starting point for developing more complex models that forecast many concrete attributes simultaneously, such as strength, slump, and so on. In order to produce a new mix percentage of cement with improved qualities, the method may be used to optimise the design mix for target strength and droop constraints. According to the findings in this study, these elements will soon be researched.

## References

[1] D.-C. Feng, J. Li, Stochastic nonlinear behavior of reinforced concrete frames. ii: numerical simulation, J. Struct. Eng. 142 (3) (2015) 04015163.
[2] D.-C. Feng, X.-D. Ren, J. Li, Stochastic damage hysteretic model for concrete based on micromechanical approach, Int. J. Non-Linear Mech. 83 (2016) 15–25.
[3] P.H. Bischoff, S. Perry, Compressive behaviour of concrete at high strain rates, Mater. Struct. 24 (6) (1991) 425–450.
[4] M. Lessard, O. Challal, P.-C. Aticin, Testing high-strength concrete compressive strength, ACI Mater. J. 90 (4) (1993) 303–307.
[5] H.-S. Shi, B.-W. Xu, X.-C. Zhou, Influence of mineral admixtures on compressive strength, gas permeability and carbonation of high performance concrete, Constr. Build. Mater. 23 (5) (2009) 1980–1985.
[6] B. Bharatkumar, R. Narayanan, B. Raghuprasad, D. Ramachandramurthy, Mix proportioning of high performance concrete, Cem. Concr. Compos. 23 (1) (2001) 71–80.

[7] S. Bhanja, B. Sengupta, Investigations on the compressive strength of silica fume concrete using statistical methods, Cem. Concr. Res. 32 (9) (2002) 1391– 1394.
[8] M.F.M. Zain, S.M. Abd, Multiple regression model for compressive strength prediction of high performance concrete, J. Appl. Sci. 9 (1) (2009) 155–160.
[9] D.C. Feng, X.D. Ren, J. Li, Softened damage-plasticity model for analysis of cracked reinforced concrete structures, J. Struct. Eng. 144 (6) (2018) 04018044, https://doi.org/10.1061/(ASCE)ST.1943-541X.0002015.
[10] D.-C. Feng, Z. Wang, G. Wu, Progressive collapse performance analysis of precast reinforced concrete structures, Struct. Des. Tall Special Build. 28 (5) (2019) e1588.
[11] D.-C. Feng, S.-C. Xie, W.-N. Deng, Z.-D. Ding, Probabilistic failure analysis of reinforced concrete beam-column sub-assemblage under column removal scenario, Eng. Fail. Anal. 100 (2019) 381–392.
[12] H. Salehi, R. Burgueño, Emerging artificial intelligence methods in structural engineering, Eng. Struct. 171 (2018) 170–189.
[13] R. Siddique, P. Aggarwal, Y. Aggarwal, Prediction of compressive strength of self-compacting concrete containing bottom ash using artificial neural networks, Adv. Eng. Softw. 42 (10) (2011) 780–786.
[14] M. Uysal, H. Tanyildizi, Estimation of compressive strength of self compacting concrete containing polypropylene fiber and mineral additives exposed to high temperature using artificial neural network, Constr. Build. Mater. 27 (1) (2012) 404–414.
[15] A.T.A. Dantas, M.B. Leite, K. de Jesus Nagahama, Prediction of compressive strength of concrete containing construction and demolition waste using artificial neural networks, Constr. Build. Mater. 38 (2013) 717–722.
[16] Z.-H. Duan, S.-C. Kou, C.-S. Poon, Prediction of compressive strength of recycled aggregate concrete using artificial neural networks, Constr. Build. Mater. 40 (2013) 1200–1206.
[17] J.-S. Chou, C.-K. Chiu, M. Farfoura, I. Al-Taharwa, Optimizing the prediction accuracy of concrete compressive strength based on a comparison of datamining techniques, J. Comput. Civil Eng. 25 (3) (2010) 242–253.
[18] J.-S. Chou, C.-F. Tsai, A.-D. Pham, Y.-H. Lu, Machine learning in concrete strength simulations: multi-nation data analytics, Constr. Build. Mater. 73 (2014) 771–780.
[19] B.G. Aiyer, D. Kim, N. Karingattikkal, P. Samui, P.R. Rao, Prediction of compressive strength of self-compacting concrete using least square support vector machine and relevance vector machine, KSCE J. Civil Eng. 18 (6) (2014) 1753–1758.
[20] S. Motamedi, S. Shamshirband, R. Hashim, D. Petkovic´, C. Roy, Estimating unconfined compressive strength of cockle shell-cement-sand mixtures using soft computing methodologies, Eng. Struct. 98 (2015) 49–58.
[21] A.-D. Pham, N.-D. Hoang, Q.-T. Nguyen, Predicting compressive strength of high-performance concrete using metaheuristic-optimized least squares support vector regression, J. Comput. Civil Eng. 30 (3) (2015) 06015002.
[22] B.A. Omran, Q. Chen, R. Jin, Comparison of data mining techniques for predicting compressive strength of environmentally friendly concrete, J. Comput. Civil Eng. 30 (6) (2016) 04016029.

[23] S. Chithra, S.S. Kumar, K. Chinnaraju, F.A. Ashmita, A comparative study on the compressive strength prediction models for high performance concrete containing nano silica and copper slag using regression analysis and artificial neural networks, Constr. Build. Mater. 114 (2016) 528–535.
[24] H. Naderpour, A.H. Rafiean, P. Fakharian, Compressive strength prediction of environmentally friendly concrete using artificial neural networks, J. Build. Eng. 16 (2018) 213–219.
[25] A. Ashrafian, M.J.T. Amiri, M. Rezaie-Balf, T. Ozbakkaloglu, O. Lotfi-Omran, Prediction of compressive strength and ultrasonic pulse velocity of fiber reinforced concrete incorporating nano silica using heuristic regression methods, Constr. Build. Mater. 190 (2018) 479–494.
[26] J. Zhang, G. Ma, Y. Huang, F. Aslani, B. Nener, et al., Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression, Constr. Build. Mater. 210 (2019) 713–719.
[27] Z.-H. Zhou, Ensemble Methods: Foundations and Algorithms, Chapman and Hall/CRC, 2012.
[28] Z.-H. Zhou, Ensemble learning, Encyclopedia of biometrics (2015) 411–416.
[29] X. Wu, V. Kumar, J.R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G.J. McLachlan, A. Ng, B. Liu, S.Y. Philip, et al., Top 10 algorithms in data mining, Knowl. Inf. Syst. 14 (1) (2008) 1–37.
[30] I.-C. Yeh, Modeling of strength of high-performance concrete using artificial neural networks, Cem. Concr. Res. 28 (12) (1998) 1797–1808.
[31] I.-C. Yeh, Modeling slump of concrete with fly ash and superplasticizer, Comput. Concr. 5 (6) (2008) 559–572.
[32] S.R. Safavian, D. Landgrebe, A survey of decision tree classifier methodology, IEEE Trans. Syst., Man, Cybern. 21 (3) (1991) 660–674.
[33] R.A. Berk, Classification and regression trees (cart), in: Statistical Learning From a Regression Perspective, Springer, 2008, pp. 1–65.
[34] H.I. Erdal, O. Karakurt, E. Namli, High performance concrete compressive strength forecasting using ensemble models based on discrete wavelet transform, Eng. Appl. Artif. Intell. 26 (4) (2013) 1246–1254.