Estimation of concrete compression using regression models

Tsehay Admassu Assegie[1], Ayodeji Olalekan Salau[2,4], Tayo Uthman Badrudeen[3]

[1]Department of Computer Science, College of Natural and Computationall Science, Injibara University, Injibara, Ethiopia
[2]Department of Electrical/Electronics and Computer Engineering, Afe Babalola University, Ado-Ekiti, Nigeria
[3]Department of Electronic and Electrical Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria
[4]Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

Article Info

ABSTRACT

The objective of this study is to evaluate the effectiveness of different regression models in concrete compressive strength estimation. A concrete compressive strength dataset is employed for the estimation of the regressor models. Regression models such as linear regressor, ridge regressor, k-neighbors regressor, decision tree regressor, random forest regressor, gradient boosting regressor, AdaBoost regressor, and support vector regressor are used for developing the model that predicts the concrete strength. Cross-validation techniques and grid search are used to tune the parameters for better model performance. Python 3.8 programming language is used to conduct the experiment. The Performance evaluation result reveals that the gradient boosting regressor has better performance as compared to other models using root mean square error (RMSE).

Corresponding Author:

Ayodeji Olalekan Salau
Department of Electrical/Electronics and Computer Engineering, Afe Babalola University
Ado-Ekiti, Nigeria
Email: ayodejisalau98@gmail.com

1. INTRODUCTION

Concrete is a widely used material for construction in the world [1]. Different components such as cement, age, coarse and fine aggregate, and water should be combined to make the concrete [2], [3]. These components can affect the compressive strength of the concrete. To obtain the real compressive strength of concrete (target labels in the dataset), an engineer needs to break the cylinder samples under the compression-testing machine [4], [5]. The failure load is divided by the cylinder's cross-section to obtain the compressive strength. Engineers use different kinds of concretes for different building purposes. For example, the strength of concrete used for residential buildings should not be lower than 2500 pounds per square inch (psi) (17.2 megapascals (MPa) [6]. Concrete is a material with high strength in compression, but low strength in tension. That is why engineers use reinforced concrete (usually with steel rebars) to build structures.

The compressive strength of concrete is one of the significant parameters in structural engineering [7]. Because today's construction work requires high-strength concrete for higher durability. Determining the strength of concrete involves time, planning, and financial resources because the commonly used compressive strength factor is obtained on the 28th day [8]. For this reason, concrete strength is estimated before the concrete is used for building construction. The concrete strength is estimated by conducting laboratory tests. But, laboratory analysis of concrete sample strength requires significant experimentationtime and costs.

Machine learning has been widely used for concrete strength prediction [9]. Moreover, there is a clear opportunity for an automated model to reduce the wait time for estimating concrete strength using traditional laboratory tests. Thus, with the concrete dataset acquired, it is possible to develop a model that learns about the relations between variables and develop a predictive model. However, there is no previous work that has

conducted extensive experiments on the better regression model that is recommended for concrete strength prediction. Thus, this research aims to evaluate the effectiveness of different regression models for concrete strength prediction. The objective of this work is to study the effectiveness of various supervised regression models on the prediction of concrete strength depending on its compositions. This study investigates the variability of different regression models in the prediction of concrete strength. (1) To develop different regression models for predicting the compressive strength of concrete, (2) To compare the models using visualizations and measures of accuracy, and (3) To analyze the importance of concrete contrastive features of each model. The rest of this work is organized as follows: in section 2, the state-of-the-art concretestrength prediction models are reviewed, and in section 3, the dataset and the regression models used in the simulation are presented. Section 4 presents the result found and finally section 5 concludes the findings of the work.

2.   RELATED WORK

There have been significant efforts in the utilization of supervised machine learning algorithms to tackle construction problems. Supervised algorithms and prediction models have been widely employed in the estimation of concrete strength that is used in construction. For instance, in [10] backpropagation (BP) neural network (NN) is applied to the concrete compressive strength dataset to automate concrete strength analysis. Chen *et al.* [11] employed a bagging classifier and developed an automated concrete compressive strength prediction. The evaluation of the predictive accuracy of the developed model reveals that the artificial neural network outperforms compared decision tree and bagging classifier for concrete strength prediction.

The performance of K-nearest neighbor (KNN), random forest, and decision tree algorithms are compared for concrete strength prediction [10]. The comparative result for concrete strength prediction using the (KNN), random forest, and decision tree shows that random forest outperforms compared to KNN and decision tree model. The random forest model performed with an accuracy of 91.26% on concrete strength prediction. Similarly, in [12], long short-term memory is applied to the concrete strength prediction problem. The authors employed a support vector regression algorithm to develop a model for strength prediction. As shown in the result analysis, the support vector regression model achieved a root mean square error (RMSE) of 0.508 and an R-Squared error of 0.997. Moreover, the authors compare the conventional support vector regression model with developed long short-term memory and the result shows that long short-term memory outperformed compared to the support vecotor regressor.

Advanced machine learning techniques are applied to develop a model that predicts concrete strength [13]. The study compared different machine learning models such as decision tree, AdaBoost regressor, and bagging regressor. The result reveals that the bagging regressor or random forest regressor achieved higher accuracy as compared to the Adaboost and decision tree for concrete strength prediction. Another study [14] applied hybrid machine learning models to predict concrete strength. The researchers applied an artificial neural network (ANN) to develop the model that predicts concrete strength. The model is evaluated on a concrete test set and the result shows that the ANN model achieves 97% accuracy on concrete strength prediction. The performance of the support vector machine (SVM) and ANN is evaluated for concrete strength prediction [15]. The comparative result reveals that SVM outperforms as compared to the ANN model. Although different models have been developed for concrete strength prediction, the model in the literature has scope for improvement in terms of accuracy.

3.   METHOD

The concrete strength prediction dataset is obtained from the University of California Irvine (UCI) machine learning data repository. The dataset consists of 1,030 samples of different concrete compositions and the concrete strength value. The UCI concrete compressive strength dataset is widely used to develop a machine learning model to predict concrete strength [16]-[18]. To develop the model, different regression algorithms are employed. The features of cement compressive are demonstrated in Table 1.

3.1. Performance measures

To evaluate the performance of different regression models employed to predict concrete compressive strength, the root means square error (RMSE) and R-squared are employed, and accuracy as a performance measure or metric. The root means the square error is determined by the following formula [19]-[22]. The RMSE is employed due to its wider applicability in regression model evaluation [23]-[25].

$$RMSE = \sqrt{\sum (p_i - O_i)^2}\Big/N \tag{1}$$

where: Σ is the sum of the difference between the predicted and observed values for the ith observation in the dataset, O_i is the observed value for the ith observation in the dataset and N is the sample size.

Table 1. The performance of regression models for concrete strength prediction.

No.	Model	RMSE	R Squared
1	Linear regressor	0.628799	0.599578
2	Ridge regressor	0.628678	0.599819
3	K-Neighbors regressor	0.550102	0.694892
4	Decision tree regressor	0.458048	0.786737
5	Random forest regressor	0.340006	0.886869
6	Gradient boosting regressor	0.330691	0.890700
7	AdaBoost regressor	0.471189	0.775838
8	Support vector regressor	0.428961	0.814650

The regression models are trained on 7 input values of concrete strength composition parameters such as water, cement, coarse and fine aggregate, and other features. The effect of each concrete composition parameter on the concrete strength is examined. A correlation matrix between input parameters is calculated by employing the Person correlation coefficient for each pair of concrete composition parameters. The correlation matrix used to investigate the effect of different concrete composition parameters is demonstrated in Figure 1. The correlation matrix shows the relationship between each concrete compression strength feature.

As shown in Figure 1, the concrete strength is highly affected by the cement parameter having a correlation value of 0.50. Thus, a strong positive correlation is found between composite strength and the amount of cement in the composition of the concrete. Moreover, there is a strong positive correlation between concrete composition namely, superplasticizer and water, and similarly a positive correlation between superplasticizer and fly ash. Moreover, the 3-dimensional plot of the three most important features, namely cement and superplasticizer against age and compressive strength is demonstrated in Figure 2. As demonstrated in Figure 2, cement and superplasticizer have a high impact on the concrete compressive strength.
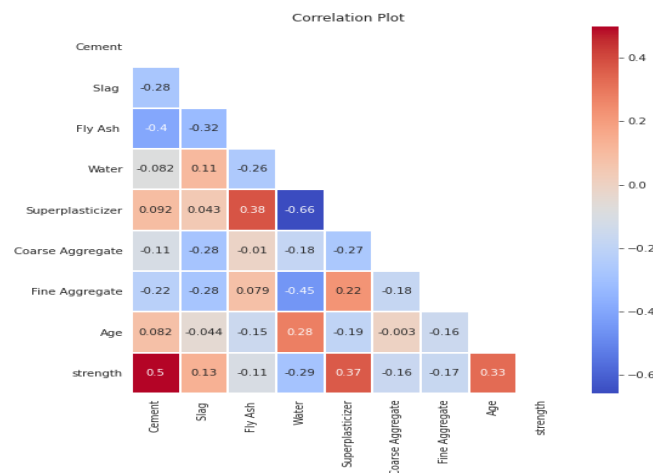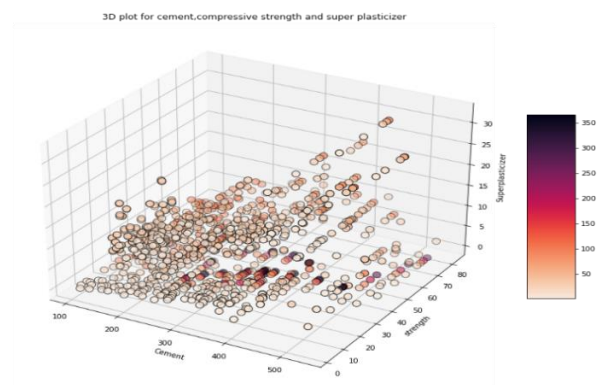

Figure 1. Correlation matrix.

Figure 2. Three-dimensional plot of the most important features

4. RESULTS AND DISCUSSION

The variability among different regression models such as linear regression, decision tree regression, random forest regression, support vector regression, K-neighbors regression, gradient, and AdaBoost regression is analyzed on the concrete compressive strength dataset. The results are visualized using tables and graphs using RMSE and R squared error as performance evaluation measures. Table 2 demonstrates the performance of the regression model on concrete strength prediction. In Table 2, the performance of different regression models is demonstrated. As shown in Table 2, the gradient boosting regressor model has better performance as compared to other regression models. Moreover, the performance of the regression models on correct strength prediction is shown in Figure 3.

Table 2. The performance of regression models on concrete strength estimation

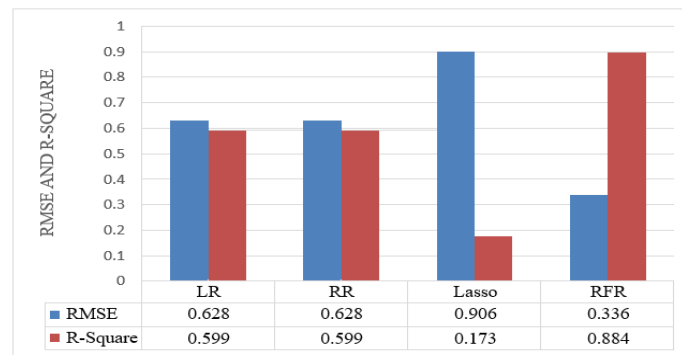| No.	Model	Training score in (%)	Test score in (%)	95% confidence interval
1	Linear regression	93.2	86.3	-57<->41.47
2	Lasso regression	74.2	67.9	-0.1<->0.92
3	Adaboost regression	83.5	77.8	0.44<->0.68
4	Random forest regression	83.2	74.4	0.30<->0.78
5	Gradient boost regression	97.2	90.2	0.54<->1.00

Figure 3. RMSE and R Square comparison of regression model on concrete strength prediction.

4.1. Performance of the regression models for concrete strength estimation

The regression models are compared using training and test accuracy for concrete strength estimation. The comparative results for each regression model on concrete strength estimation are demonstrated in Table 2. As demonstrated in Table 2, The Gradient boosting regression is the better model as it provides better complexity with a 95% confidence interval between 88% and 95%. The test score for each regression model on concrete strength estimation is demonstrated in Figure 4.

As shown in Table 2, the gradient boosting regressor model outperforms compared to other regression models for concrete strength estimation. The confidence interval for gradient boosting regressor is between 0.54 and 1.00. Moreover, the gradient boosting regressor model has an accuracy of 90.2% on test data with a confidence of 95%. Figure 4 demonstrates the accuracy of different regression models on concrete strength estimation.
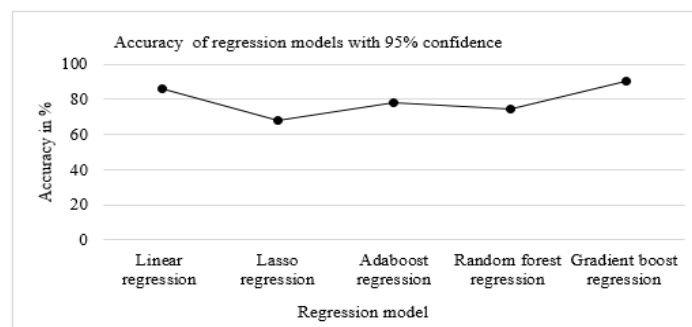

Figure 4. Accuracy of regression models for concrete strength estimation with 95% confidence

5. CONCLUSION

This study evaluated the performance of regression models for concrete strength estimation. Experimental simulation shows that the gradient boosting regressor gives a good accuracy score as compared to other regression models for concrete strength estimation. The experiment has also proved that with validation better result is obtained with gradient boosting regressor. Thus, a gradient boosting regressor is better for modeling the strength of high-performance concrete. The regression models are significant to improve the concrete strength prediction and reduce the number of experimental tests required when checking concrete composition. Finally, the study reveals that regression models are capable of mapping input features to target concrete compressive strength.

REFERENCES

[1] Y. Yu, W. Lio, J. Li, and T. N. Nguyen, "A novel optimized self-learning method for compressive strength prediction of high-performance concrete," *Construction and Building Materials.*, vol. 184, pp. 229-247, 2018, doi: 10.1016/j.conbuildmat.2018.06.219.

[2] I. U. Ekanayake, D. P. P. Meddage, and U. Rathnayake., "A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP)s," *Case Studies in Construction Materials*, vol. 16, pp. e01059, 2022, doi: https://doi.org/10.1016/j.cscm.2022.e01059.

[3] M. T. Cihan., "Prediction of Concrete Compressive Strength and Slump by Machine Learning Methods," *Hindawi Advances in Civil Engineering*, vol. 2019, pp. 1–12, doi: https://doi.org/10.1155/2019/3069046.

[4] V. H. Gaidhane, N. Kumar, R. K. Mittal, and J. Rajevenceltha "An efficient approach for cement strength prediction," *International Journal of Computers and Applications*, 2019, doi: 10.1080/1206212X.2019.1673288.

[5] J. M. P. Ojeda, S. R. Bocanegra, and L. Q. Huatangari, "Determination of the Compressive Strength of Concrete Using Artificial Neural Network," *International Journal of Engineering and Technology Innovation*, vol. 11, no. 3, pp. 204-215, 2021, doi: https://doi.org/10.46604/ijeti.2021.7479.

[6] N. K. Ngawang and S. V. Deo, "Estimating the Concrete Compressive Strength Using Hard Clustering and Fuzzy Clustering Based Regression Techniques," Hindawi *Publishing Corporation the Scientific World Journal.*, vol. 2014, no. 10, pp. 381549, doi: http://dx.doi.org/10.1155/2014/381549.

[7] K. Maina, A. O. Gwaya and D. O. Koteng, "Concrete Strength Prediction using Multi-Linear Regression Model: A case study of Nairobi Metropolitan," *International Journal of Soft Computing and Engineering*, vol. 8, no. 5, pp. 11–20, 2019.

[8] S. Ma, X. Shi, C. Yu, W. Wang, Y. Ren and X. Tian, "Research on Improved BP Neural Network Gangue Powder Concrete Compressive Strength Prediction Model," *2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, 2020, pp. 419-423, doi: 10.1109/ITAIC49862.2020.9339112.

[9] M. K. Keleş, A. E. Keleş And Ü. Kiliç, "Prediction Of Concrete Strength With Data Mining Methods Using Artificial Bee Colony As Feature Selector," *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, 2018, pp. 1-4, doi: 10.1109/IDAP.2018.8620905.

[10] J. D. P. Gil, C. Palencia, N. S. Monteiro, and R. M. Garcia, "To predict the compressive strength of self-compacting concrete with recycled aggregates utilizing ensemble machine learning models," *Case Studies in Construction Materials*, vol. 16, p. r01046, , doi: 10.1016/j.cscm.2022.e01046.

[11] H. Chen, X. Li, Y. Wu, L. Zuo, M. Lu and Y. Zhou, "Compressive Strength Prediction of High-Strength Concrete Using Long Short-Term Memory and Machine Learning Algorithms," *Buildings*, vol. 12, no. 302, pp. 1–2, 2022, doi: 10.3390/buildings12030302.

[12] A. Ahmad, W. Ahmad, F. Aslam, and P. Joyklad, "Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques," *Case Studies in Construction Materials*, volo. 16, p. e00840, 2021, doi: 10.1016/j.cscm.2021.e00840.

[13] X. Wang, Y. Liu, and H. Xin, "Bond strength prediction of concrete-encased steel structures using hybrid machine learning method," *Structures*, vol. 32, pp. 2279–2292, 2021, doi: 10.1016/j.istruc.2021.04.018.

[14] T. Q. Nguyen, "Using viscous resistance system based on machine learning in engineering concrete structures," *Journal of Low-Frequency Noise, Vibration and Active Control*, 2022, doi: 10.1177/14613484211072366.

[15] K. O. Akande, T. Owolabi, A. Twaha, and S. O. Olatunji, "Performance Comparison of SVM and ANN in Predicting Compressive Strength of Concrete," *IOSR Journal of Computer Engineering*, vol. 16, no. 5, pp. 88-94, 2014.

[16] A. O. Salau and S. Jain, "Feature Extraction: A Survey of the Types, Techniques, Applications," *2019 International Conference on Signal Processing and Communication (ICSC)*, 2019, pp. 158-164, doi: 10.1109/ICSC45622.2019.8938371.

[17] M. Y. Cheng, P. M. Firdausi, and D. Prayogo, "High-performance concrete compressive strength prediction using Genetic Weighted Pyramid Operation Tree (GWPOT)," *Applications of Artificial Intelligence.*, vol. 29, pp. 104-113, 2014, doi: 10.1016/j.engappai.2013.11.014.

[18] X. Y. Huang *et al.*, "Compressive Strength Prediction of Rubber Concrete Based on Artificial Neural Network Model with Hybrid Particle Swarm Optimization Algorithm," *Materials*, vol. 3934, no.15, pp. 1–20, 2021, doi: 10.3390/ma15113934.

[19] M. Ahmad *et al.*, "Supervised Learning Methods for Modeling Concrete Compressive Strength Prediction at High Temperature," *Materials*, vol.19832, no. 14, pp. 1-19–305, 2021, doi: 10.3390/ma14081983.

[20] Y. Liu *et al.*, "High-Performance Concrete Strength Prediction Based on Machine Learning," *Hindawi Computational Intelligence and Neuroscience*, vol. 2022, pp. 23816–23833, 2022, doi: 10.1155/2022/5802217.

[21] P. Chopra and R. K. Sh, "Comparison of Machine Learning Techniques for the Prediction of Compressive Strength of Concrete," *Hindawi Advances in Civil Engineering*, vol. 2018, co. 3, pp. 1–9, 2018, doi: 10.1155/2018/5481705.

[22] D. Zhang, F. Sun, and T. Liu, "Prediction of Compressive Strength of Geopolymer Concrete Based on Support Vector Machine and Modified Cuckoo Algorithm," *Hindawi Advances in Materials Science and Engineering*, vol. 2021, pp. 1-14, doi: 10.1155/2021/4286810.

[23] B. A. Salami, T. Olayiwola, T. A. Oyehan, and I. A. Raji, "Data-driven model for ternary-blend concrete compressive strength prediction using machine learning approach, Construction, and Building Materials," *Hindawi Advances in Materials Science and Engineering*, vol. 301, pp: 124-152 doi: 10.1016/j.conbuildmat.2021.124152.

[24] Q. Han, C. Gui, J. Xu, and G. Lacidogna "A generalized method to predict the compressive strength of high-performance concrete by improved random forest algorithm," *Construction and Building Materials*, vol. 226, pp. 734-742, 2019, doi:

       10.1016/j.conbuildmat.2019.07.315.
[25]   M. S. Barkhordari, D. J. Armaghani, A. S. Mohammed, and D. V. Ulrikh, "Data-Driven Compressive Strength Prediction of Fly
       Ash Concrete Using Ensemble Learner Algorithms," *Buildings*, vol. 12, no. 132, pp. 1-16, 2022, doi: 10.3390/ buildings12020132.

BIOGRAPHIES OF AUTHORS

Mr. Tsehay Admassu 🆔 🔍 SC P obtained a Master of Science degree in Computer Science from Andhra University Faculty of Science, India 2016. He received B.Sc. (Computer Science) from Dilla University, Ethiopia in 2013. His research interest includes machine learning, data mining, bioinformatics, network security, and software-defined networking. He has published over 35 journal articles in international journals and conferences. Tsehay is a member of the International Association of Engineers (IAENG), with membership number: 254711. He can be contacted at email: tsehayadmassu2006@gmail.com.

Dr. Ayodeji Olalekan Salau 🆔 🔍 SC P received the B.Eng. in Electrical/Computer Engineering from the Federal University of Technology, Minna, Nigeria. He received the M.Sc. and PhD degrees from the Obafemi Awolowo University, Ile-Ife, Nigeria. His research interests include research in the fields of computer vision, image processing, signal processing, machine learning, control systems engineering and power systems technology. Dr. Salau serves as a reviewer for several reputable international journals. His research has been published in many reputable international conferences, books, and major international journals. He is a registered Engineer with the Council for the Regulation of Engineering in Nigeria (COREN), a member of the International Association of Engineers (IAENG), and a recipient of the Quarterly Franklin Membership with ID number CR32878 given by the Editorial Board of London Journals Press in 2020 for top quality research output. More recently, Dr. Salau's research paper was awarded the best paper of the year 2019 in Cogent Engineering. In addition, he is the recipient of the International Research Award on New Science Inventions (NESIN) under the category of "Best Researcher Award" given by Science Father with ID number 9249, 2020. Currently, Dr. Salau works at Afe Babalola University in the Department of Electrical/Electronics and Computer Engineering. He can be contacted at email: ayodejisalau98@gmai.com.

Mr. Tayo Uthman Badrudeen 🆔 🔍 SC P received his Bachelor of Engineering degree in Electrical and Electronics Engineering from the University of Ilorin, Nigeria, in 2012, and masters of Science degree in Electronic and Electrical Engineering from Obafemi Awolowo University, Ile-Ife, Nigeria, in 2017. Currently, He is a PhD Student in Power System Engineering, Department of Electronic and Electrical Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria. His research interests include control strategies for power system, renewable energy and power electronics. He can be contacted at email: tiyoh4energie@gmail.com.