https://dx.doi.org/10.4314/dujopas.v11i1c.22
ISSN (Print): 2476-8316
ISSN (Online): 2635-3490
Dutse Journal of Pure and Applied Sciences (DUJOPAS), *Vol. 11 No. 1c March 2025*

Performance of Compressive Strength Prediction Models for Laterized Ordinary Portland Cement - Activated Meta-Kaolin Concrete

[1*]Feyidamilola Faluyi, [1]Oluwaseun Adetayo, [1]Olugbenga O. Amu

[1]Federal University
Oye-Ekiti

Email: feyidamilola.faluyi@fuoye.edu.ng

Abstract
One of the main initiatives to reduce carbon emissions and make construction more sustainable is the use of geopolymer/activated pozzolans concrete. Accurately predicting the compressive strength of activated pozzolanic concrete is essential for the use and acceptance of such a complex system of concrete. This study employed the use of eight input variables to predict the compressive strength of hybrid laterized ordinary Portland cement-activated metakaolin (OPC-AMk) concrete using four machine learning (ML) models namely: Adaptive Boosting (Adaboost), K-Nearest Neighbour (KNN), extreme gradient boosting (XGBoost), and random forest (RF). A total of 192 concrete specimens were produced to obtain data for training and testing in ratio of 80:20. Adaboost, KNN, XGBoost, and RF had respective R^2 values of 0.8895, 0.9188, 0.9300, and 0.9299; meaning that the absolute fraction of variance (R^2) for XGBoost was the highest. XGBoost also gave the lowest mean absolute error (MAE) and root mean squared error (RMSE). Thus XGBoost can be adjudged the best among the four ML considered for this research. A check on the impact of the variables on strength predictions using SHapley Additive exPlanations (SHAP) revealed that the most significant factors are curing age, Portland cement content, activator/water ratio and water/binder ratio.

Keywords: Activated pozollan, Geopolymer concrete, Laterized concrete, Strength prediction.

INTRODUCTION
Sustainability is key for the future of our world. The construction industry has a big role to play in ensuring the safety and resource security of the planet. Low energy binder such as activated pozzolans is important and will no doubt help in saving the environment from the destructive effect of high CO^2 emitting production process such as ordinary Portland cement (OPC) manufacturing. The safety of our world can only be guaranteed with a reduction in green house polluting gases.

However, the acceptability and continual use of activated pozzolans is hinged on accurate prediction of its strength parameters. Ability to correctly predict the compressive strength of activated pozzolanic concrete will not only help in ensuring safe design but will also help in faster project delivery by reducing cost and time associated with destructive testing

*Author for Correspondence

Machine learning (ML) models have come into high reckoning in concrete strength prediction due to their computational capability and accuracy (Awoyera *et al*., 2020). According to de-Prado-Gil *et al.* (2022), ML was found to be very effective in predicting the strength performance of self compacting concrete. Lyngdoh *et al.,* (2022) while predicting the strength of concrete was able to demonstrate the potency of XGBoost among various approaches to predict the strength of concrete with missing data when ensembled with k-nearest neighbors (kNN). Xu *et al.* (2021) and Mai *et al.* (2021) attested to the efficacy of random forest and xgboost regressors in concrete strength prediction judging by the low mean absolute percentage error (MAE) obtained in their use; thus resulting in reduced experimental cost.

The advantage of ML in concrete strength prediction is portrayed in their ability to handle the intricate and often non linear relationship which often existed between the dependent and independent variables of the concrete matrix be it OPC, geopolymer or activated pozzolan binder (Ghosh and Ransinchung, 2022; Pan *et al.*, 2022). AdaBoost and Random Forest models have been shown to perform well in concrete strength prediction, this can be attested to by the outcome of statistical indices such as the coefficient of determination (R^2), root mean squared error, mean absolute percentage error, and relative root mean square error (Ahmad *et al*., 2021). Kalabarige *et al.,* (2024) found out that RF, XGBoost and KNN were the best performing algorithms in their prediction of compressive strength of concrete containing different industrial byproducts.

This study explored the predictability of compressive strength of hybrid laterized ordinary Portland cement-activated metakaolin (OPC-AMk) concrete by considering eight input variables and using four machine learning (ML) models namely: Adaptive Boosting (Adaboost), K-Nearest Neighbour (KNN), extreme gradient boosting (XGBoost), and random forest (RF). Although there have been numerous works on OPC concrete strength prediction, not much is available on predicting the compressive strength of concrete when OPC is hybridized with activated metakaolin as binder. Especially with the use of ensemble machine learning approach as done in this research. The influence of each of the input (independent) variables on prediction was also checked.

METHODOLOGY

Materials

The binders for the OPC-AMk concrete are mainly ordinary Portland cement of grade 32.5R which satisfied the BS/EN 197–1:2000 standard, and activated metakaolin which was obtained by calcinatng kaolin in an electric furnace at a temperature of 600 °C. The total sum of Fe_2O_3, SiO_2 and Al_2O_3 oxides in the metakaolin after calcination was 86%, this confirmed its suitableness as a natural pozzolan based on ASTM C618-19, The metakaolin was activated using a combination of sodium hydroxide solution (NaOH) and sodium silicate (Na_2SiO_3) with combine specific gravity of 1.53; these were obtained from African Fertilizer and Chemicals, Agbara in Ogun state, Nigeria. Crushed granite of sizes 4.75 to 19 mm was used as coarse aggregate while the fine aggregates were river sand (conforming to BS 882:1992 and laterite of fineness modulus 3.98 (classification as coarse-fine according to ASTM C136/C136M).

Design Mix

Mix ratio of 1:2:4 (binder:fine aggregate:coarse aggregate) was used for both the control and laterized OPC-AMk hybrid concrete specimens. As shown in Table 1, the control mix was without laterite and activated metakaolin, while the other mixtures contained activated AMk and laterite in various proportions. The batching and mixing of materials was by weight in kilogram (kg). Ordinary Portland Cement was replaced with AMk at 10%, 20% and 30% while laterite was also used to replace river sand at 10%, 20% and 30%respectively.

Table 1. Mix proportions of OPC- AMK concrete (kg/m^3)

Group	Designation*	OPC	MK	Coarse Aggregate	Sand	Laterite	Activator	Total Water
C	Control	326.0	0	1304	652.0	0	0	238
	C1	326.0	0	1304	586.8	65.2	0	238
	C2	326.0	0	1304	521.6	130.4	0	238
	C3	326.0	0	1304	456.4	195.6	0	238
M1	M1A	293.4	32.6	1304	652.0	0	14.67	239.4
	M1B	293.4	32.6	1304	586.8	65.2	14.67	239.4
	M1C	293.4	32.6	1304	521.6	130.4	14.67	239.4
	M1D	293.4	32.6	1304	456.4	195.6	14.67	239.4
M2	M2A	260.8	65.2	1304	652.0	0	29.34	248.7
	M2B	260.8	65.2	1304	586.8	65.2	29.34	248.7
	M2C	260.8	65.2	1304	521.6	130.4	29.34	248.7
	M2D	260.8	65.2	1304	456.4	195.6	29.34	248.7
M3	M3A	228.2	97.8	1304	652.0	0	44.01	251.1
	M3B	228.2	97.8	1304	586.8	65.2	44.01	251.1
	M3C	228.2	97.8	1304	521.6	130.4	44.01	251.1
	M3D	228.2	97.8	1304	456.4	195.6	44.01	251.1

* C = concrete with 100% OPC, M1 = concrete with 10% MK, M2 = concrete with 20% MK, M3 = concrete with 30%, MK, A = 0% laterite, B = 10% laterite, C = 20% laterite, D = 30% laterite,

Specimen Description

Compressive strength which was the target variable in the ML prediction models was obtained by crushing concrete cube specimens of size 150mm x 150mm x 150 mm. The concrete samples were cured in water for 7, 28, 56 and 91 days. The compressive strength was taken as load (kN) at failure divided by the cross sectional area (mm^2) of the cube.

Machine Learning Models

Adaboost

Adaptive Boosting, which is also known as Adaboost was developed by Yoav Freund and Robert Schapire in 2003. It is an algorithm that can be used for statistical classification as well as regression. The performance of other learning algorithms can be improved by combining with Adaboost. Adaboost is adaptive because it favors instances previous classifiers/regressors wrongly classified by adjusting succeeding feeble learners. In several cases, Adaboost had been observed to be less susceptible to over fitting in comparison to many other learning algorithms. Giving the weak learners whose error is small larger weight in system will enhance the general accuracy of the strong learners (Feng *et al.*, 2020). AdaBoost is suitable for unbalanced datasets but under-performs in the existence of noise. Training AdaBoost is slower and its difficult to optimize the hyper-parameters (Misra and Li, 2020).

K-Nearest Neighbour (KNN)

K-Nearest Neighbour is one of the simplest ML algorithms based on supervised learning method. The theory of KNN algorithm assumes the presence of similar nearby objects (Amor *et al.*, 2023). KNN algorithm take on the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available class. The *k*-nearest neighbor is a non parametric algorithm, meaning that underlying data will not be subject to any assumption. In KNN, the represented class is assigned an unclassified structure by a bulk of its *k* nearest neighbors (Huang *et al.*, 2019). It is also known as a lazy learner algorithm because it does not immediately learn from the training set rather it stored the samples and works on them at the time of classification (Shi *et al*., 2022). The three basic factor in the algorithm of KNN are K, which is the number of measured instances, the metric of the distance and the classification's decision regulation. The limit of the neighborhood property, k is germane to the accuracy of KNN regression/classification model. Abnormalities is the usual when k is too small. According to Liu *et al.,* (2018). When we have a set of grouped instances $M=\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\}$, where x_i is the feature vector of the unlabeled instance, y_i is the label and $y_i=c_1,c_2,\ldots,c_K$, $i=1,2,\ldots n$ (Shi *et al*., 2022).

Therefore for a training data (x,y), the k-NN algorithm searched for the k nearest instances to x based on a given distance metric. $n_k(x)$ is the k instances neighbourhood. The calculation of test sample x label based on decision rule is thus given by (1):

$$y = argmax_{c_j} = \sum_{x_i \in n_k(x)} I(y_i = c_j), \quad i = 1,2,3,..n;\ j = 1,2,3,\ldots K \tag{1}$$

where I is the indicator function.

The advantage of KNN is its simplicity of implementation, using it in concrete strength prediction resulted in low MAE and high coefficient of determination (Beskopylny *et al.,*2022)

XGBoost

Extreme Gradient Boosting (XGBoost) ML is an efficient and scale-able execution of the gradient boosting framework by Friedman (2001). XGBoost, which is an ensemble tree based machine learning algorithm was developed by Chen and Guestrin. According to Li *et al.,*(2018) XGBoost which is a branch of boosting algorithm mainly combine various feeble classifiers to form a powerful one, a lifting tree model which merged numerous tree model into a robust classifier . XGBoost can be used for both regression and classification. Its utilization in concrete strength prediction has shown its robustness and efficiency.

Predicted parameters and results can be expressed as (2):

$$L(\phi) = \sum_i l(y_i\, \hat{y}_i) + \sum_k \Omega\,(f_k) \tag{2}$$

Where $L(\phi)$ denotes the expression in linear space, i means the number of i th samples, k is the number of k th tree, and $\hat{y}_i$ is the prediction value of i th sample in x_i.

XGBoost is good in making predictions for unorganized data, which has found its application in diverse fields including traffic accident classification, risk assessment, concrete mix design optimization and shear strength predictions to mention a few. Wang *et al*., (2024) using several ML algorithm for the prediction of concrete compressive strength found XGBoost to possess the highest accuracy among all the algorithm explored.

Random Forest

ML approach called random forest (RF) is created by a process known as bagging, which is shorthand for bootstrap aggregation. To develop trees independently throughout the entire data set in RF, a bootstrap sample is utilized. As per Shaqadan, (2016), the primary parameters to adjust when adopting RF trees are the number of trees to be grown and the attributes selected at each split. The ability of RF to manage a high number of variables with a relatively small data set and its application in determining feature importance—an evaluation of each variable's impact on the model as a whole—are two of its strong points Denil *et al*., (2014). Data are split into homogenous groups called nodes for regression purposes; this partition is based on data of splitting variables. Tracing the path from the root node based on the split can then be used to forecast the result.

The excellent accuracy of RF algorithm as confirmed by statistical error analysis showed that it is one of the best ML that can predict the compressive strength of concrete (Gupta, 2023 ; Zhu, 2023). RF can predict the compressive strength of non conventional concrete better as evident in the accuracy and goodness of fit.This can significantly lessen test requirement, time and cost; an advantage in concrete mix design [Wang *et al*., 2023; Mai *et al*., 2021; Zhang *et al*., 2023]

Data Description

To predict the compressive strength of OPC-AMk concrete, eight independent variables were used. They are; ordinary Portland cement (OPC), metakaolin (Mk), sand, laterite, slump value, water/ cement ratio, Alkaline activator / water ratio and age of concrete. The response variable was the concrete compressive strength (CS). The details of the input variables were as described on Table 2. Data was obtained from 192 concrete specimens which were splitted for training and testing in ratio 80:20. The statistical relationship between the variables were checked using Pearson correlation coefficient, a widely used means of measuring relationship between variables (Xu *et al*., 2021), Pearson correlation heat map as shown on Figure 1 was done to show the relationship among the variables. The perfect negative correlation between OPC and Mk, sand and laterite and the near perfect relationship between water-cement ratio and activator-water ratio are expected because Mk was used to replace OPC, likewise laterite was used to replace sand in the concrete mix. Also, water/cement ratio and alkali activator/water are directly related by water used in the concrete.

Table 2: The Details of The Input Data Used For Analysis.

Input Variables	Minimum	Maximum	Range
Ordinary Portland Cement (OPC) (kg/m^3)C	228.2	326	97.8
Metakaolin (MK) (kg/m^3)	0	97.8	97.8
Fine Aggregate (Sand) (kg/m^3)	456.4	652	195.6
Fine Aggregate (Laterite) (kg/m^3)	0	195.6	195.6
Slump Value (mm)	10	65	55
Water / OPC ratio	0.67	1.04	0.37
Alkaline Activator / Water ratio	0.0	0.2	0.2
Age (days)	7	91	84

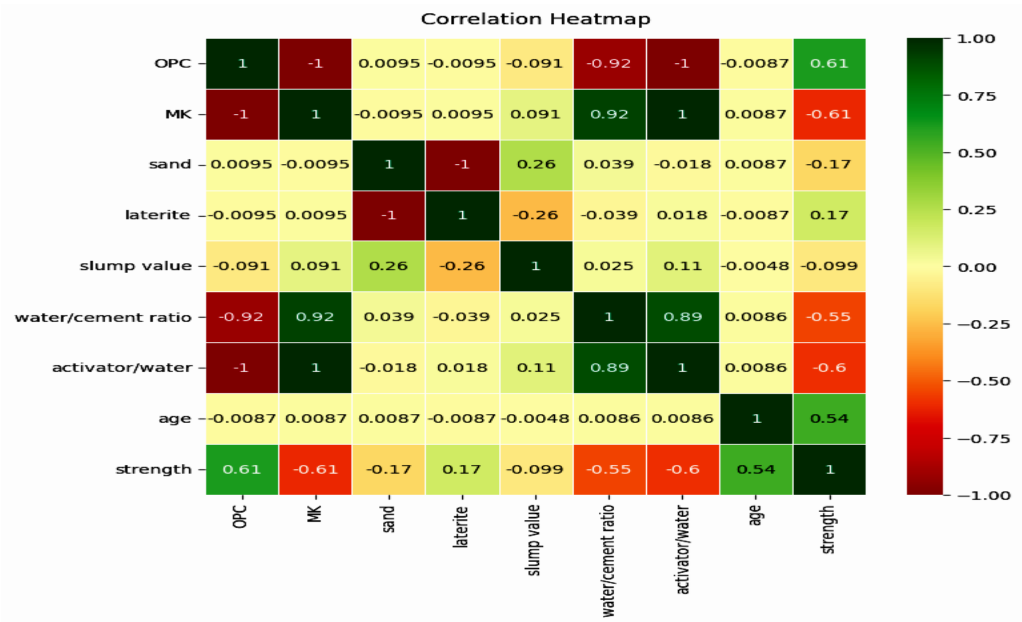


Figure 1. Correlation Coefficients for the Independent and Dependent Variables

Parameter Settings for Models

First, starting parameter values were assigned to each machine learning model. Nevertheless, it was found that the optimal tuning/setting for each ML was obtained using hyper-parameter tweaking using "GridSearchCV," which is a Python "sklearn" library feature. "GridSearchCV" functions by recommending the optimal configuration based on factors that yield the maximum R2. For the ML models, the parameters provided by "GridSearchCV" were thus employed.

Variable Importance

The influence of each of the independent variables on the target variable can be better understand through sensitivity analysis or variable importance plot. Higher sensitivity value is an indication of higher significance or impact such a variable has on the output variable. According to Shang *et al.,* (2002), the input variables have a notable effect on the prediction of the output variables. To measure the importance and impact of each input variable on the compressive strength, SHAP (SHapley Additive exPlanations) was used to assess the feature significance (measured as the mean absolute Shapley values) and feature value (impact on model output).

Model Evaluation Metrics

Evaluating a model is very important in data analysis. Three main metrics (3) to (5) were used for evaluation in these models. Namely:

i. R Square/Adjusted R Square,

$$R = \frac{n\Sigma yy\prime - (\Sigma y)(y\prime)}{\sqrt{n(\Sigma y^2) - (\Sigma y)^2}\sqrt{n(\Sigma y\prime^2) - (\Sigma y\prime)^2}} \quad (3)$$

ii. Mean Square Error (MSE)/Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y - y')^2} \quad (4)$$

iii. Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y - y'| \quad (5)$$

Where y is the measurement, y' is the corresponding prediction and n is number of data point

In R^2 the closer the value is to 1, the better the model fits. The disparities between predicted and actual values are expressed in terms of mean absolute error (MAE) and root mean squared error (RMSE). Their value should be around zero for a good fit. When squaring to amplify deviation, RMSE gives outliers greater weight than MAE. As a result, although MAE is more resilient to outliers and more accurately represents the actual condition of predicted value errors, RMSE is more susceptible to outliers and reflects the variance of error. The metrics were employed to evaluate how well the different ML techniques performed. λ was introduced to indicate the difference between RMSE and MAE in order to better illustrate the magnitude of the prediction error change (Li *et al.*, 2018).

RESULTS AND DISCUSSION

The age of concrete (curing days), OPC content and activator/water ratio were found to be the most important inputs contributing to increase in compressive strength of hybrid OPC-AMk concrete. OPC content is almost as important as age of concrete in strength formation. The variable that has the least effect on the prediction outcome was Mk content. This is shown in Figure 2a-2c.

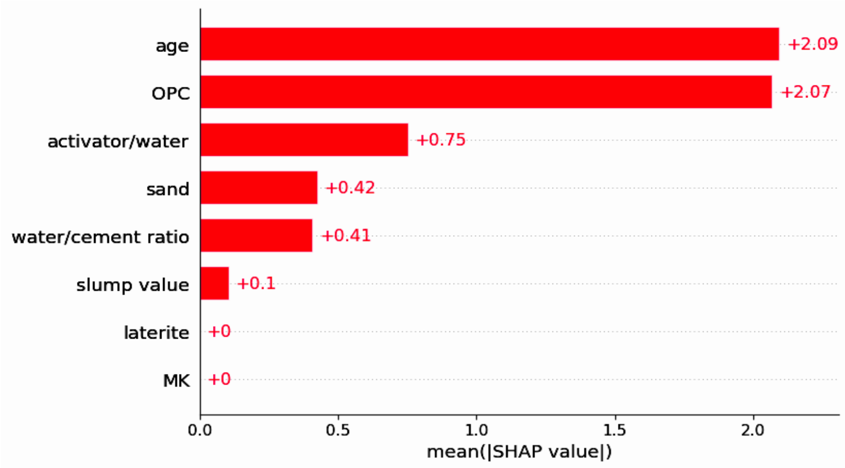


Figure 2a. SHAP feature importance measured as the mean absolute Shapley values

From the beeswarm plot in Figure 2b, it was observed that higher age caused an increase in predicted compressive strength, similarly in OPC content and slump an higher values forced an increase in predicted strength. Although activator/water ratio was the third most influential input as shown in Figure 2a (feature importance), an higher value of it only implied a reduction in the compressive strength of the concrete. The quantity of water available in the freshly mix concrete to dilute the activator can impact the strength of activated pozzolan/ geopolymer concrete. As the sand content increased, there was reduction in compressive strength of the resulting concrete. This is logical since higher sand to cement ratio lead to a weakened matrix. The same applies to the water/cement ratio. With increasing water cement ratio, compressive strength of hybrid OPC-activated pozzolanic concrete reduced (Faluyi *et al*., 2022).

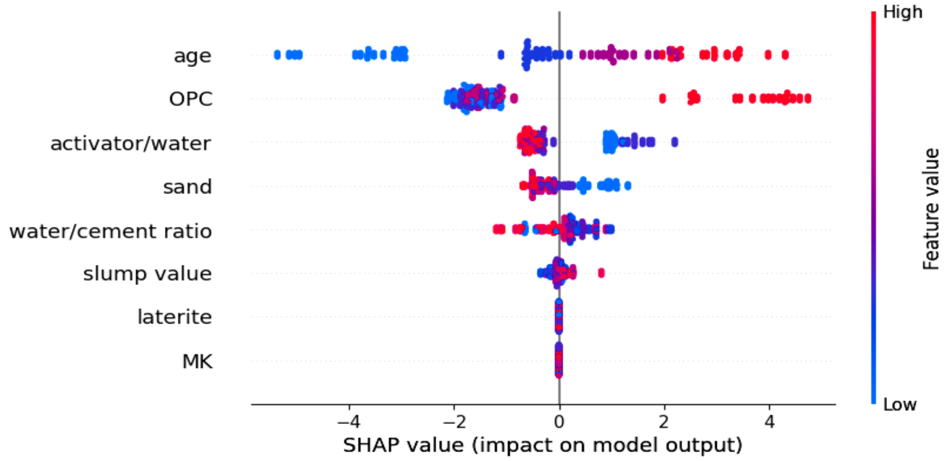


Figure 2b. SHAP Bee swarm plot

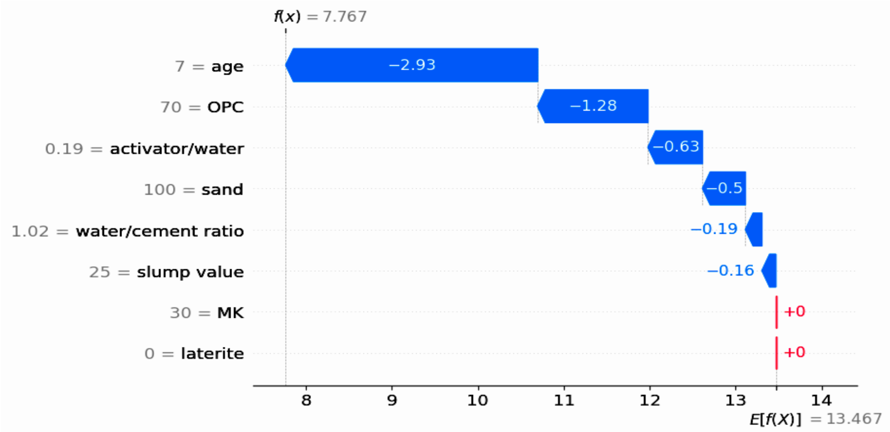


Figure 2c. SHAP feature waterfall plot

From the values obtained on the four ML models used for the prediction, it can be deduced that KNN, XGBoost, Adaboost and RF can be effectively used to predict the compressive strength of hybrid OPC-AMk concrete as seen on Table 3.

Table 3 Evaluation Metrics

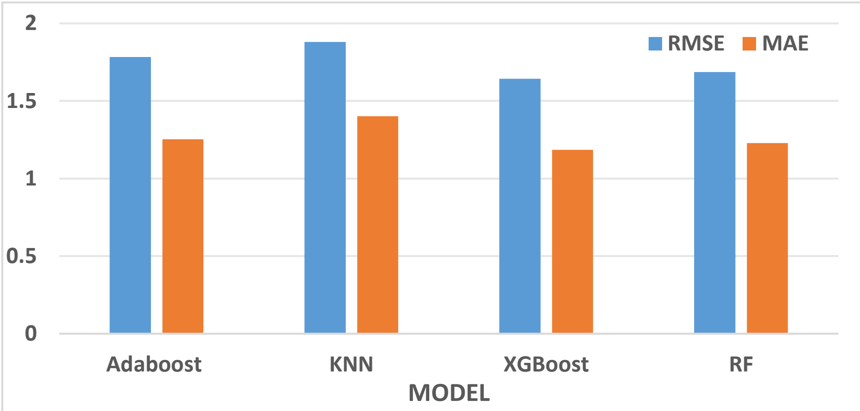
Metric	Adaboost	KNN	XGBoost	RF
R^2	0.8895	0.9188	0.9300	0.9299
RMSE	1.7830	1.8812	1.6442	1.6864
MAE	1.2528	1.4013	1.1856	1.2289
λ	0.5302	0.4799	0.4586	0.4575

R-Square

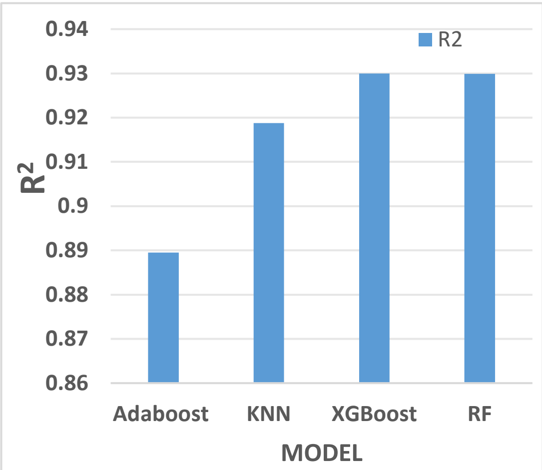
Among the four ML algorithm used for the compressive strength prediction, XGBoost has the highest R^2 as observed from Table 3 and Figure 3b. With a score of 0.93, XGBoost was the best, while Adaboost with a score of 0.8895 has the least R score among the four ML models.

RMSE and MAE

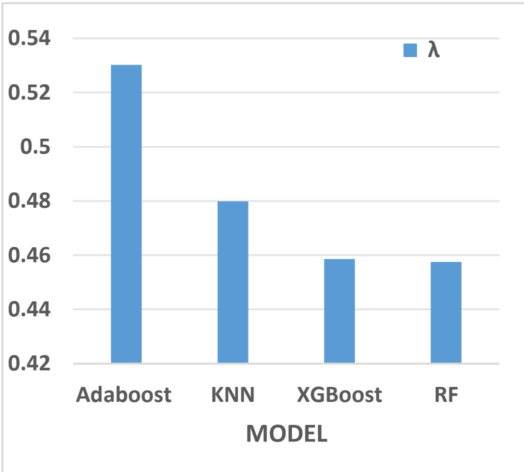
XGBoost has the lowest mean absolute error (MAE) and root mean square error (RMSE), at 1.1856 and 1.6442, respectively, as seen on Table 3 and Figure 3a. KNN has the highest RMSE and MAE values, 1.8812 and 1.4013, although its R^2 value indicated it should be a better model than Adaboost. As a measure of model stability and a better way to express the size of the prediction error, λ was introduced, which is the difference between RMSE and MAE. Out of the four machine learning models examined, RF with the lowest λ value of 0.4575 is deemed to be the most stable and effective, as demonstrated in Figure 3c. Adaboost performed the poorest, with the greatest λ value of 0.5302, whereas XGBoost was marginally better by RF.



(a)



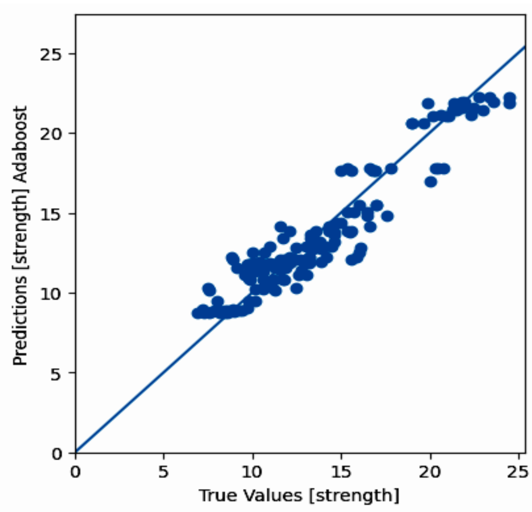
(b)



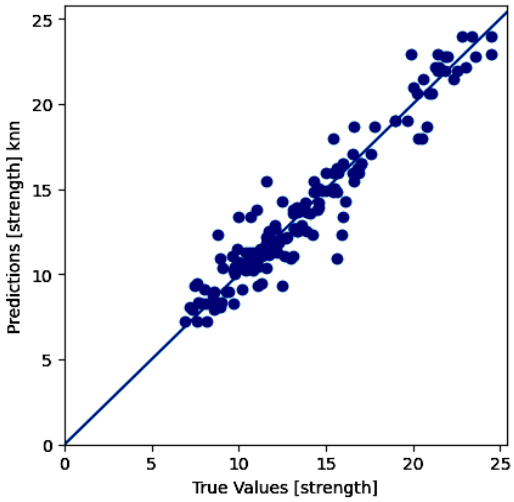
(c)

Figures 3 a-c Bar plot of the evaluation metrics

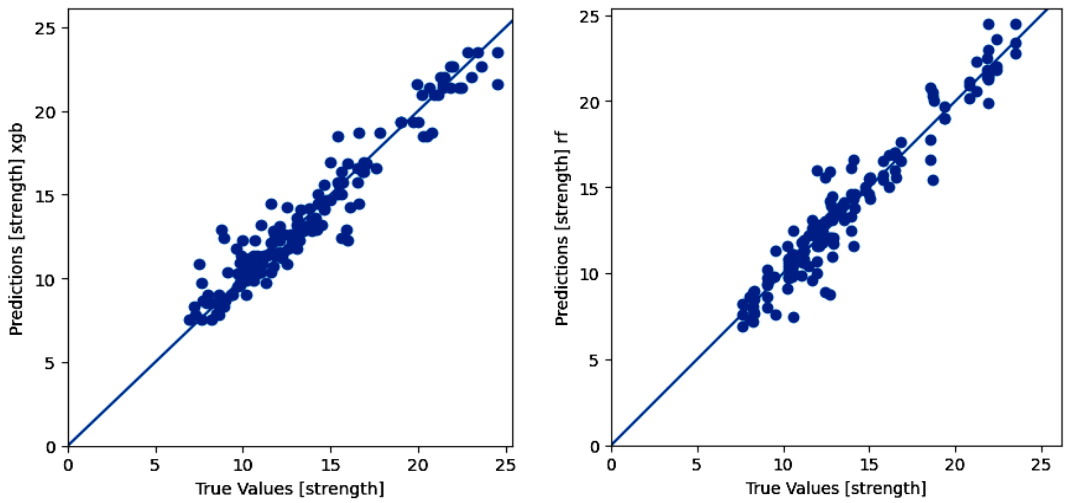
A plot of true values against the predicted values using the test data as shown In Figures 4a-d showed that KNN, XGBoost and RF models being more linear in plot performed better than Adaboost. KNN and Adaboost, having more outliers showed that they are slightly less accurate in predicting the concrete strength when compared to RF and XGBoost.



(**a**) Adaboost model



(**b**) KNN model



(**c**) XGBoost model (**d**) RF model

Figures 4 a-d Plot of true values against model predicted values using testing data

CONCLUSION

In this study, the compressive strength of OPC-AMk concrete was predicted using four machine learning models. The concrete's age (curing days) has the largest positive influence on the compressive strength, according to the influence and relevance of the input variables. Concrete gets stronger the older it gets. Portland cement content followed age on the scale of importance, showing that even though activated Mk is supposed to benefit from the hydration heat of OPC, Portland cement controls the compressive strength of hybrid OPC-Mk when water curing is used.

It was discovered that the laterite content had the least impact on the compressive strength of the matrix. When the four ML models were compared, XGBoost with an R-square value of 0.9300, outperformed RF (0.9299), KNN (0.9188), and Adaboost (0.8895). XGBoost was deemed the best model overall, despite the fact that RF had the lowest λ out of the four ML used.

REFERENCES

Ahmad, M., Hu, J., Ahmad, F., Tang, X., Amjad, M. *et. al.* (2021). Supervised Learning Methods for Modeling Concrete Compressive Strength Prediction at High Temperature. *Materials,* vol. 14, 1983. https://doi.org/10.3390/ma14081983

Amor, N., Noman, M.T., Petru, M., Sebastian, N. and Balram, D. (2023). A review on computational intelligence methods for modeling of light weight composite materials. *Applied Soft Computing*, vol.147. https://doi.org/10.1016/j.asoc.2023.110812.

ASTM International, (2014). Standard specification for coal fly ash and raw or calcinated natural pozzolan for use in concrete, ASTM C618-12a. Accessed 24 May 2022. https://www.astm.org/Standards/C618-12a.htm.

ASTM International (2019). Standard test method for sieve analysis of coarse and fine aggregates, ASTM C136/C136M-19. ASTM International, PA.

Awoyera, P.O., Kirgiz,M.S., Viloria, A. and Ovallos-Gazabong, D. (2020). Estimating strength properties of geopolymer self-compacting concrete using machine learning techniques. *Journal of Material Research and Technology*, vol.9(4), pp.9016–9028.

Beskopylny, A.N., Stel'makh, S.A., Shcherban, E.M., Mailyan, L.R., Meskhi, B. *et. al.* (2022). Concrete Strength Prediction Using Machine Learning Methods CatBoost, k-Nearest

Neighbors, Support Vector Regression. *Applied Sciences,* vol. 12(21) 10864. https://doi.org/10.3390/app122110864

British Standards Institution, (1992). Specification for aggregates from natural sources for concrete, BS 882:1992. British Standards Institution, London.

British Standards Institution, (2000). Cement, composition, specification and conformity criteria for common cement, BS EN 197-1: 2000, British Standards Institution, London.

Denil, M., Matheson, D. and Freitas, N. (2014). Narrowing the gap: random forests in theory and in practice. *Proceedings of the 31 st International Conference on Machine Learning,* Beijing, China. 2014.

de-Prado-Gil, J., Palencia, C., Jagadesh, P. and Martínez-García, R.A. (2022). Comparison of machine learning tools that model the splitting tensile strength of self-compacting recycled aggregate concrete. *Materials,* vol.15(12), pp.1–20.

Faluyi, F., Arum, C., Ikumapayi, C.M. and Alabi, S.A. (2022). A Review of the Compressive Strength Predictor Variables of Geopolymer Concrete. *FUOYE Journal of Engineering and Technology*, vol 7(3), pp. 404-414. https://doi.org/10.46792/fuoyejet.v7i3.884.

Feng, D., Liu, Z., Wang, X., Chen, Y., Chang, J. et. al.(2020). Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. *Construction and Building Materials*, vol.230. https://doi.org/10.1016/j.conbuildmat.2019.117000.

Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, vol.29(5), pp. 1189-1232.

Ghosh, A. and Ransinchung, R.N. (2022). Application of machine learning algorithm to assess the efficacy of varying industrial wastes and curing methods on strength development of geopolymer concrete. *Construction and Building Materials,* vol. 341, 127828.

Gupta, P., Gupta, N. and Saxena, K. K.(2023). Predicting compressive strength of geopolymer concrete using machine learning. *Innovation and Emerging Technologies,* vol 10, 2350003, pp. 1-9.

Huang, S., Huang, M. and Lyu,Y. (2019). A novel approach for sand liquefaction prediction via local mean-based pseudo nearest neighbor algorithm and its engineering application. *Advanced Engineering Informatics*, vol. 41. https://doi.org/10.1016/j.aei.2019.04.008.

Kalabarige, L.R., Sridhar, J., Subbaram, S., Prasath, P. and Gobinath, R. (2024). Machine Learning Modeling Integrating Experimental Analysis for Predicting Compressive Strength of Concrete Containing Different Industrial Byproducts. *Advances in Civil Engineering*, 7844854, pp. 1-11. https://doi.org/10.1155/2024/7844854

Li, Y., Zou, C., Berecibar, M., Naninimaury, E. and Cha, J.C. *et al*. (2018). Random forest regression for online capacity estimation of lithium-ion batteries. *Applied Energy,*vol.232, pp. 197–210

Liu, R., Yang, B., Zio, E. and Chen, Y. (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, vol. 108, pp. 33-47. https://doi.org/10.1016/j.ymssp.2018.02.016.

Lyngdoh, G.A., Zaki, M., Krishnan,N.M.A. and Das, S. (2022). Prediction of concrete strengths enabled by missing data imputation and interpretable machine learnin. *Cement and Concrete Composites*, vol. 128(2022). doi.org/10.1016/j.cemconcomp.2022.104414.

Mai, H.V.T., Nguyen, T.A., Ly, H.B. and Tran, V.Q. (2021). Prediction Compressive Strength of Concrete Containing GGBFS using Random Forest Model. *Advances in Civil Engineering,* vol. 2021, pp.12. https://doi.org/10.1155/2021/6671448.

Misra, S. and Li, H. (2020). Noninvasive fracture characterization based on the classification of sonic wave travel times, Machine Learning for Subsurface Characterization. *Gulf Professional Publishing*, 2020, pp. 243-287, doi.org/10.1016/B978-0-12-817736-5.00009-0.

Pan, X., Xiao, Y., Suhail, S.A., Ahmad, W., Murali, G. *et. al*. (2022). Use of artificial intelligence methods for predicting the strength of recycled aggregate concrete and the influence of raw ingredients. *Materials*, vol.15, 4194, 2022.

Shang, M., Li, H., Ahmad, A., Ahmad, W., Ostrowski, K.A., *et al.* (2002). Predicting the mechanical properties of rca-based concrete using supervised machine learning algorithms. *Materials,* vol.15(647). https://doi.org/10.3390/ma15020647

Shaqadan, A. (2016). Prediction of concrete mix strength using random forest model. *International Journal of Applied Engineering Research,* vol.11(22), pp. 11024-11029.

Shi,Y.,Yang, K.,Yang, Z. and Zhou, Y. (2022). Primer on artificial intelligence," *Mobile Edge Artificial Intelligence,* Academic Press, 2022. pp 7-36, https://doi.org/10.1016/B978-0-12-823817-2.00011-5.

Wang, W., Zhong, Y., Liao, G., Ding, Q.and Zhang, T. *et al*. (2024). Prediction of Compressive Strength of Concrete Specimens Based on Interpretable Machine Learning. *Materials,* vol. 1(15), 3661. https://doi.org/10.3390/ma17153661

Wang, K., Ren, J, Yan, J., Wu, X. and Dang, F. (2023). Research on a concrete compressive strength prediction method based on the random forest and LCSSA-improved BP neural network. *Journal of Building Engineering*, vol. 76, 107150, https://doi.org/10.1016/j.jobe.2023.107150.

Xu, J., Zhou, L., He, G., Ji, X. and Dai, Y. *et al.* (2021). Comprehensive machine learning-based model for predicting compressive strength of ready-mix concrete. *Materials*, vol. 14(5), pp.1068-1086.

Zhang, F., Fang, H., Li, P. and Qiao, Y. (2023). Strength prediction of concrete with large amount of fly ash based on improved random forest. *International Journal of Microstructure and Materials Properties*, vol.16(6), pp.467 - 481

Zhu, J., Fang, S., Yang, Z., Qin, Y. and Chen, H. (2023). Prediction of Concrete Strength Based on Random Forest and Gradient Boosting Machine. 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA), Shenyang, China, 2023, pp. 306-312, doi: 10.1109/ICPECA56706.2023.10075839