3rd International Conference on Advances in Civil and Environmental Engineering (ICACEE-2024)

University of Engineering & Technology Taxila, Pakistan

Conference dates: 21st and 22nd February 2024; ISBN: 978-969-23675-2-3

A Study on the Prediction of Compressive Strength of Geo-Polymer Concrete through Stacking Regression and Random Forest

Muhammad Waqas Akram[1,*], Faisal Shabbir[1]

[1]University of Engineering & Technology Taxila, 47050, Pakistan.

*Corresponding author: engr.waqasakram786@gmail.com

ABSTRACT

Concrete plays a pivotal role in construction, and the emergence of innovative variations such as geo-polymer concrete holds significant promise for sustainability. This research focuses on predicting compressive strength in geo-polymer concrete, utilizing both Stacking regression and random forest model. The model achieved notable R^2 values of 0.72 and 0.89, indicating their ability to make accurate prediction. Compressive strength ranged from 20 MPa to 65 MPa, with an average of 40 MPa. Moreover, model with higher R^2 values of 0.89 and 0.83 exhibited increased accuracy, showing differences between predicted and observed compressive strengths (highest: 90 MPa, lowest: 10 MPa, average: 45 MPa). These results highlight the effectiveness of machine learning technique, particularly the random forest method, in forecasting the performance of geo-polymer concrete. Furthermore, they underscore the critical importance of meticulous model selection and evaluation in engineering applications to ensure dependable predictions.

KEYWORDS: Machine learning, Concrete, Geo-polymer concrete, Compressive strength

1 INTRODUCTION

Geo-polymer concrete represents a compelling alternative to conventional Portland cement concrete, primarily owing to its diminished carbon footprint and enhanced durability characteristics. Unlike traditional concrete, which depends on Portland cement for binding, geo-polymer concrete employs aluminosilicate materials and alkaline activators to form a robust and chemically stable binder matrix. This innovative material not only contributes to environmental sustainability by curtailing CO2 emissions linked with cement manufacturing but also showcases superior mechanical attributes, including heightened compressive strength and reduced permeability[1]. The compressive strength of concrete is a critical parameter that governs its structural performance and suitability for various engineering applications[2]. Therefore, accurate prediction of compressive strength is essential for ensuring the structural integrity and longevity of geo-polymer concrete structures[3].Traditional approaches to predicting compressive strength often entail labour-intensive and expensive laboratory testing procedures. However, in recent years, machine learning model like Stacking regression and random forest have emerged as efficient alternatives for predicting material properties using input parameters[4]. Stacking regression is a statistical method employed to model the correlation between a dependent variable (e.g., compressive strength) and one or more independent variables (such as mix proportions, curing conditions, and aggregate properties). By constructing a linear equation based on the observed data, stacking regression enables the quantification of the relationship strength between the input variables and the predicted outcome, facilitating accurate predictions for unseen data[5]. On the other hand, random forest is a powerful ensemble learning algorithm that operates by constructing a multitude of weightage tree during training and outputting the mode of the classes (classification) or mean prediction (regression) of the individual tree[6]. Random forest is acknowledged for its robustness to over fitting and its capability to capture intricate nonlinear relationships between input and output variables. These qualities make it well-suited for modelling the complex behaviour of geo-polymer concrete[7]. In this study, the aim is to investigate the effectiveness of stacking regression and random forest model in predicting the compressive strength of geo-polymer concrete. The research have explore the influence of various input parameters, including binder composition, curing temperature, and curing duration, on the compressive strength of geo-polymer concrete specimens. By comparing the predictive performance of Stacking regression and random forest models, the goal is to identify the most accurate and reliable approach for predicting the compressive strength of geo-polymer concrete under different conditions. Overall, this research contributes to the advancement of predictive modelling techniques for geo-polymer concrete, providing insights into the factors affecting its compressive strength and guiding the development of more sustainable and durable construction materials.

2 RESEARCH METHODOLOGY

Python is used to train the models using a Windows 10 PC with a GeForce 980 GTX GPU and an Intel 2.20 GHz CPU for the tests. In order to accelerate calculations, Google associated with NVIDIA to provide GPUs and CPUs. Python language 3.8 was used for the study, and Google Colab had 16 gigabytes of RAM available for the trials. Moreover, Google Colab has a user-friendly interface and works well with Python, making it a useful tool for running code and doing tests. Furthermore, the platform provides a significant amount of RAM (16 gigabytes) for the experiments in order to support larger datasets as shown in figure 1 and memory-intensive tasks.

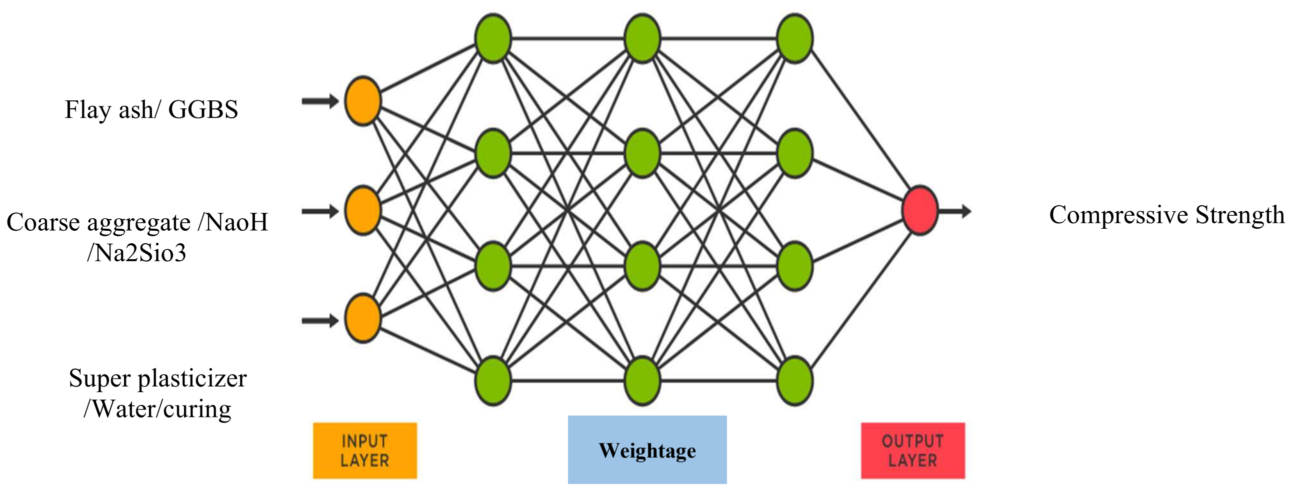

Figure 1: Network architecture

2.1 Data record and Normalization

The data was obtained from the Research paper of Elsevier and frontiers through Google scalar. Total records of 480 experiment dataset were collected and the database covers wide range of Fly ash, NAOH, Na_2Sio_3, Coarse aggregate, GGBS and Super plasticizer as shown in above figure 1. Data sets become very huge due to outliers, and they should be removed for improving efficiency even though it narrows down the error to correct the result. A box with defined variance and medium is plotted to identify outliers. No observation should be discarded until it is certain that the outlier is the result of measurement error, or its presence affects the model.

Table 1: Data Normalizations

Parameters	Curing Temperature (°C)	Curing Time (h)	Age of Specimen (days)	Fly ash (kg/m³)	Na_2SiO_3 (kg/m³)	NaOH (kg/m³)	Coarse Aggregate (kg/m³)	Super plasticizer (kg/m³)	H_20 (kg/m³)	Compr-ssive Strength (Mpa)
Count	480.0	480.0	480.0	480.0	480.0	480.0	480.0	480.0	480.0	480.0
Mean	68.937	27.4625	12.748	386.47	158.61	2.248	782.84	1.931	53.55	39.13
St-deviation	25.22	13.255	11.641	56.9	70.11	0.533	170.54	2.41063	3.827	14.32
Minimum	20.0	4.0	1.0	246.0	102.0	0.4	467.0	0.0	45.1	1.5
Maximum	120.0	96.0	56.0	527.0	347.1	4.0	1209.0	11.3	64.0	89.0
25%	60.0	24.0	7.0	343.0	120.0	2.0	626.0	0.0	49.0	30.0
50%	70.0	24.0	7.0	400.0	143.0	2.5	756.0	1.6	55.9	39.0
75%	90.0	24.0	28.0	424.8	156.0	2.5	900.0	2.0	55.9	47.0
Count	480.0	480.0	480.0	480.0	480.0	480.0	480.0	480.0	480.0	480.0
Mean	68.937	27.4625	12.748	386.47	158.61	2.248	782.84	1.931	53.55	39.13

The performance of the models on both the training and testing datasets demonstrated remarkable similarity, signifying that our models are not over fitting but rather capable of successfully applying their learned knowledge to the trained ranges as shown in table1. This robustness and uniformity in performance give us confidence that our models can be effectively utilized in practical applications within the specified data ranges, and they are well-equipped to handle real-world scenarios with reliability and accuracy.

3 RESULTS AND DISCUSSION

3.1.1 Random forest

The random forest model achieved an R-squared (R^2) value of 0.72, indicating that approximately 72% of the variance in compressive strength can be explained by the model. Additionally, the high R^2 value of 0.89 suggests that the model provides accurate predictions. The compressive strength predictions from the random forest model revealed significant insights. The highest predicted compressive strength was 65 MPa, while the lowest predicted value was 20 MPa. These results demonstrate the model's ability to capture the range of compressive strength values observed in geo-polymer concrete specimens as shown in figure 2. Moreover, the average predicted compressive strength of 40 MPa align closely with the average observed compressive strength, indicating the model's reliability in estimating typical concrete behaviour.

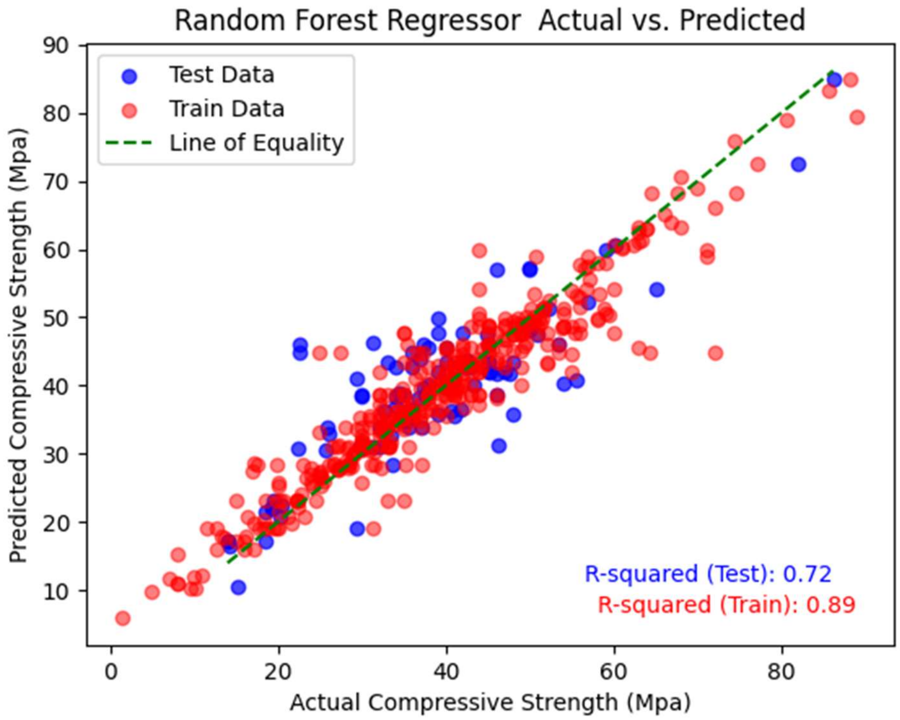

Figure 2 : Compressive strength Prediction through Random Forest

The random forest model stand out in capturing nonlinear relationships between input parameters and compressive strength, making it a suitable choice for modeling the intricate behavior of geo-polymer concrete. By utilizing an ensemble of Random forest, the model effectively integrates the influence of various factors, including binder composition and curing conditions, on concrete strength. The high R^2 value indicates the model's robustness and its capacity to generalize well to unseen data. However, it's important to acknowledge potential limitations, such as the necessity for a sufficiently diverse and representative dataset to accurately train the model. Additionally, further optimization and fine-tuning of hyper parameters may improve the model's predictive performance.

3.1.1.1 Stacking Regression

The highest compressive strength value of 90 MPa indicates the material outstanding capacity to endure compressive forces. This suggests that the material may possess superior qualities stemming from factors such as its composition, manufacturing methods, or curing procedures. In real-world scenarios, materials boasting high compressive strength are highly sought after due to their ability to offer enhanced structural stability and longevity. The R-squared (R^2) value serves as a statistical metric revealing the percentage of the variance in the dependent variable (compressive strength) that the independent variable (predictors) can predict in a regression model. A higher R^2 value, nearing 1, signifies that the model elucidates a substantial portion of the data's variability. Here in figure 3 reveals that, with R^2 values of 0.89 and 0.83, the model showcases robust predictive prowess by encapsulating around 89% and 83% of the variance in compressive strength, respectively.

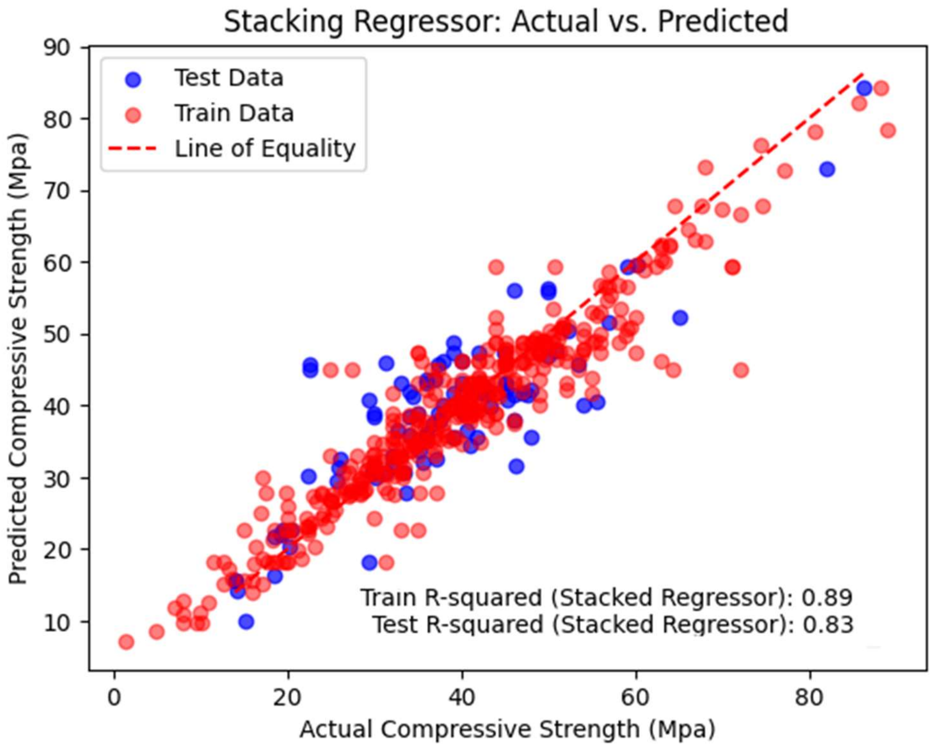

Figure 3: Compressive strength prediction through stacking regressor

The evaluation of the predictive model performance in conjunction with the supplied data on compressive strength underscores the critical significance of precise strength forecasting in engineering and construction domains. The model's notable R^2 values underscore its efficiency in predicting compressive strength, offering invaluable insights for material selection, design refinement, and structural assessment. Grasping the variability and spread of compressive strength figures assists in making well-informed choices concerning material application, thereby guaranteeing the safety and durability of built structures.

4 CONCLUSIONS

In this work, the efficiency of two different machine learning model applied for prediction of compressive strength. It has been found that stacking regression model performed well for Compressive strength prediction with R-squared of 0.89.Its Scatter diagram indicated the least scattering from the actual values.

1. With lower R^2 values (0.72, 0.89), the model demonstrates less trained random forest in predicting geo-polymer concrete compressive strength.
2. Consistent with observed values (20-65 MPa, average 40 MPa), the model reliably estimates compressive strength, validating its effectiveness for engineering applications.
3. The high R^2 values (0.89, 0.83) indicate best predictive accuracy of the model for geo-polymer concrete compressive strength with stacking regression model.
4. Inconsistencies between predicted and observed values (20-90 MPa, average 45 MPa) highlight the model high accuracy, necessitating further refinement for reliable predictions.

5 ACKNOWLEDGEMENTS

All praise and thanks are directed toward the All-Mighty Allah. I am grateful to my supervisor for providing both guidance and assistance during this study effort. I'd want to express my gratitude to my supervisor, Prof. Dr. Faisal Shabbir, instilled in me the values of perseverance and efficiency in my work. I would want to offer my most sincere gratitude to him since he allowed me to work on this project, which I am hoping would be beneficial to me in my future professional endeavours as well conclusion.

REFERENCES

1. Turkey, F.A., et al., *Concrete Compressive Strength Prediction Using Machine Learning Algorithms(Asian journal of civil engineering).* 2022.
2. Kamath, M.V., et al., *Machine-Learning-Algorithm to predict the High-Performance concrete compressive strength using multiple data(frontiers in material).* 2022.
3. Zhou, J., et al., *Decision tree models for the estimation of geo-polymer concrete compressive strength (Mathematical Biosciences and Engineering).* 2024. **21**(1): p. 1413-1444.
4. Upreti, K., et al., *Prediction of mechanical strength by using an artificial neural network and random forest algorithm (Journal of nano material).* 2022. **2022**.
5. Kashyap, V., A. Poddar, and P. Sihag, *Prediction of the Compressive Strength of Jute fibre Reinforced Concrete: A Comparative Study of ANFIS, ANN, RF and RT models(Material and design).* 2023.
6. Khan, M.A., et al., *Compressive strength of fly-ash-based geopolymer concrete by gene expression programming and random forest(Advance in civil engineering).* 2021. **2021**: p. 1-17.
7. Uddin, M.N., et al., *Prediction of compressive strength and tensile strain of engineered cementitious composite using machine learning(Mechanics and material in design).* 2024: p. 1-46.