International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.4 (2025) pp. 10229-10240
http://www.ijcesen.com

ISSN: 2149-9144

Copyright © IJCESEN

Research Article

Hybrid Naive Bayes (DAOA+SAO) Model for CBR of Rice Husk Ash Amended Geotechnical Soils

Nikita Rahaja[1]*, Ashok Kumar Gupta[2], Kushal Kanwar[3]

[1] Department of Civil Engineering, Jaypee University of Information Technology, Waknaghat, Solan, India-173234
* **Corresponding Author Email:** nikitachauhan1903@gmail.com- **ORCID:** 0000-0002-5247-0050

[2] Department of Civil Engineering, Jaypee University of Information Technology, Waknaghat, Solan, India-173234
Email: akgjuit@gmail.com- **ORCID:** 0000-0002-9947-7850

[3]Department of CSE, Jaypee University of Information Technology, Waknaghat, Solan, India-173234.
Email: kushal.kanwar@juitsolan.in - **ORCID:** 0000-0002-5247-7000

Article Info:

DOI: 10.22399/ijcesen.4808
Received : 05 November 2025
Revised : 28 December 2025
Accepted : 29 December 2025

Keywords

Expansive Soil,
Rice Husk Ash,
Soil Stabilization,
Machine Learning,
Naive Bayes,
Optimization Algorithms

Abstract:

This paper presents an integrated experimental-computational methodology for the characterization of expansive clay stabilised with rice husk ash (RHA) prolific agro-industrial residue possessing abundance of reactive silica. The clay was mixed with 0-20% dried RHA, based on dry mass, and was tested under Proctor compaction, CBR, and triaxial loading in 102 specimens. New compound feature indices were developed to assess the pozzolanic reactivity (attributed to its high amorphous SiO_2 content) and the particle-packing effects of RHA in the soil matrix. A Gaussian Naive Bayes was also used to classify specimens to ranges of index tests (low, medium, high, and extreme) based on input including RHA dose, dry density and plasticity index. This classification performed very well (≈ 97% accuracy on held-out data), and the feature-importance analysis confirms that the predominant features are the RHA content and the compaction energy. Meanwhile, a hybrid regression model is proposed to predict CBR, where a Gaussian NB base model is integrated with two complementary metaheuristic optimizers, the DAOA and the SAO are, to adjust the model weighting and thresholds. This NB, DAOA, SAO one achieves nearly perfect prediction accuracy (cross-validated $R^2 \approx 0.99$), and ultralow RMSE/MAE for CBR, which significantly outperforms a standalone NB and any of the single-optimizer formulas. In combination, this compound-feature, dual optimizer NB model is, to the best of our knowledge, the first of its kind in geotechnical results forecasting applications, and offers speedy, dependable assessments of strength and bearing properties of soil.

1. Introduction

Soil rich in montmorillonite clays is actually very widespread worldwide. it occurs in India, in Australia, and in the U.S. and elsewhere outside of Africa. Such materials exhibit pronounced plastic and shrinkage behaviour due to change in the moisture, thus leading to important geotechnical challenges. These volume changes are the ultimate cause of cracking, heave and differential settlement of infrastructure elements, pavements, retaining walls and shallow foundations, for example, which are placed directly over the expansive soil [1]. For instance, black cotton soil is widespread in several parts of India (Maharashtra, Madhya Pradesh and Gujarat) and covers 20% of the total area of India [2]. Numerous traditional ways to control harmful effects of expansive soils have been developed. Lime and cement are the most used stabilizers because they can cause chemical changes on the clay mineralogy and reduction of plasticity and swelling potential of the soils [3]-[6]. Regrettably, these conventional techniques are energy-consuming in nature, also generating much electrical energy and CO_2 emission in economic consumption, contrary to the sustainable development [7]-[8].Rice husk ash (RHA) an Eco-Friendly alternative to agricultural waste has been

emerging as innovative construction material among the geotechnical engineering community. RHA, a byproduct from rice husk burning, contains more than 85–90% of amorphous silica, and appears to be as pozzolanic when mixed with the calcium rich stabilisers (or) with the soil alone [9]-[13]. Several research reports in literature are available on the beneficial aspects of RHA in enhancing the strength properties increasing unconfined compression strength, lowering the swell potential spread as well as the plasticity index in enhancing the load-carrying capacity and long-term performance of subgrades and embankments [14]-[19].In addition to its use for geotechnical purposes, RHA stands by itself as waste utilization. If left in a spread state then it typically becomes a pollutant since its particles which lighter in weight look up in the air easily and air drags it which badly affects the environment. However, both the environmental hazards and waste generation can be eliminated when the RHA is used for civil engineering which in turn can bring economic and technological benefits [20]-[24]. Moreover, case studies in India Bangladesh and the Philippines show notable field performance of RHA for rural roads and landfill liners with improved pavements technology and structural performance. However, the relationship is non-linear and is substantially influenced by several parameters which have been introduced in the interaction of the RHA dosage and performance of the soil; including initial conditions of the soil, moisture content, the age of curing, degree of compaction, etc. Therefore, a reasonable prediction of the soil behaviour stabilized with RHA (since RHA is always used to stabilize problematic soil samples) is very crucial, especially with respect to the most important parameters, including UCS, PI, and FSI [25]-[29].Typically, such parameters are determined using in vitro test methods. But this sort of analysis is slow, costly and not feasible for the large or rapid tests that public health requires. Due to the fast-emerging trends of data science, machine learning become a widely adopted successful approach in estimating the geotechnical parameters based on experimental results with excellent efficiency and accuracy [30]-[34]. Progress of ML methods namely artificial neural network (ANN), random forest (RF), support vector machine (SVM), and genetic programming (GP) would be significantly applicable for predicting soil behaviour properties like UCS, permeability and CBR [35]-[39].The Naive Bayes (NB) algorithm is one of the ML methods, which are claimed to depend strongly on the Bayes' Theorem and is a computationally efficient method resulting in an interpretable outcome. NB is well-known for classification tasks, but it recently started to receive attention for regression ones such as UCS prediction which came to be our main theme in dealing with stabilized soils. But the performance of the NB model is sensitive to hyperparameters and assumptions, which have to be well-tuned to obtain better accuracy. To overcome this issue, the metaheuristic algorithms have been hybridized with NB in hybrid fashion. Two recent optimization algorithms, DAOA and SAO, have shown excellent performance in improving regression models for various engineering applications [40]-[46]. DAOA-based methods might be dynamic to tune the local and global searching by using accelerations, while the SAO attempts to imitate the scent-seeking behaviour of real organisms, which result in escaping from local minima and achieving a faster convergence. The integration of NB with DAOA and then SAO is a strong prediction model to describe the complexities and the nonlinear behavior of RHA-stabilized soils. This paper examines this kind of hybrid model to predict UCS, PI and FSI from experimental test results of soils stabilized with various RHA contents.

2. Experimental procedures

The expansive soil samples were excavated from a site of a field condition prevalent with black cotton soil (BCS) in Sagar, Madhya Pradesh, India. The soil was dark brown and highly plastic CI type (intermediate compressibility clay) according to IS:1498-1970. Its liquid limit was approximately 52%, plastic limit of 22% and free swell index of 55% as shown in Table 1. RHA: Rice husk ash (RHA) was collected from the nearby rice mills from Vysarpadi Chennai-600039 India shown in Table 2. Ash had a bulk density of 0.2–0.3 g/cm³ and a content of 92.5% amorphous silica (X-ray fluorescence [XRF]) indicating good pozzolanic activity. The obtained BCS and RHA were oven-dried at 105 °C for 24 h, ground and pass through 4.75 mm sieve to obtain the same particle size.The following tests as per I.S:2720 was carried out for each of the composites: Atterberg Limits (Part-5), FSI (Visual inspection method), Standard Proctor Compaction (Part 7), California bearing ratio (CBR) (Part 16), The UU Triaxial Test (Part-10).Figure 1 illustrates that the load-penetration curves for various soil-RHA combinations indicate that the stabilised soil shows enhanced resistance to penetration relative to untreated soil. Additionally, this resistance increases with higher RHA content. California Bearing Ratio (CBR) values were calculated for penetrations of 2.50 mm and 5.0 mm. The CBR value at 2.5 mm consistently

demonstrated a greater magnitude. Therefore, the CBR values presented in this paper are specifically related to penetrations of 2.50 mm.

2.1. Dataset Description

The methodology used in this study is described in this section, including the dataset preparation, machine learning models applied, optimization techniques used and evaluation metrics utilized to judge the performance of models. The database was generated by performing experimental tests in laboratory to study the effect of different RHA percentage contents on different soil properties. The RHA content varied from 0% to 20% in the specimens, and was considered as a factor with large dispersion among samples of the same dataset. Overall, 102 data-points were generated by indicating the behaviour of soils containing various percentages of RHA. The input parameters for each sample were Soil Composition (%), RHA Content (%), Liquid Limit (LL%), Plastic Limit (PL%), Plasticity Index (%) (PI%), FSI, OMC, MDD, and CBR. These factors are important from the perspective of soil strength, and in particular CBR which is widely used in geotechnical engineering.In order for the data to be effectively used bearing ratio in the machine learning models, the features were pre-processed using Min-Max Scaling. In particular, this method was crucial in the re-scaling of these features to fall within the [0, 1] range for machine learning. It was important to scale the attributes since, for example, soil type, used amount of cement and lime all had a much larger magnitude than CBR that otherwise could make one single attribute dominate the model. Through domain knowledge and correlation studies, CBR was chosen as the dependent variable because of its primary function in soil strength for geotechnical purposes.The dataset was then divided into three subsets via the train_test_split function implemented by scikit-learn using a 70% to train, 15% used for validation and a further 15% reserved for testing distribution. By this partition, a large part of the data could be used for training while separate data were left for validation and testing to ensure the generalisability of the model as well as its accuracy.

The aim of this research was to evaluate the CBR based on some soil properties with advanced ML models. To improve the efficiency of the regression model, the optimisation technique was used as one part of the modelling process. In particular, Smell Agent Optimisation (SAO) and Dynamic Arithmetic Optimisation Algorithm (DAOA) are used to optimise the Naive Bayes (NB) regression model. The nonparametric naive Bayes algorithm which is by default a classification method was employed here for a continuous regression in absence of categorical separation, under an assumption that features are Gaussian (normal) distributed. If this approach were used, CBR values would have been predicted from input characteristics such as OMC, PI, etc.Naive Bayes for regression gave a robust baseline, and we intended to improve it greatly through SAO and DAOA methods. The feature weights or model parameters were tuned using the SAO algorithm, which simulated a population of "smell agents" to roam across the feature space. These agents traded off between exploration versus exploitation finding the best model that would minimize the Root Mean Squared Error (RMSE) of the model. SAO training was performed in an iterative manner with the RMSE as a performance metric that guided the convergence between cycles.After optimization with SAO, the DAOA algorithm was utilized to compensate internal model parameters dynamically in several arithmetic operations, including addition, subtraction and multiplication. This approach was useful to refine the model and get better weights of features, which leaded to a better prediction. By optimizing with DAOA, the convergence was shaped more to adapt overtime ironically enabling the model to change and tune its parameters to fit better to the data.Once the optimisation process with SAO and DAOA was done, the model was reanalysed looking for some performance indicators based on RMSE (root mean square error), R^2 (coefficient of determination), and MAE (mean absolute error). These measurements were useful for the comparison of the prediction success and 50 CBR values, and to pinpoint domains with room open for improvements in the model. Root mean squared error (RMSE) was calculated in order to consider the amplitude of errors, since with RMSE it puts more weight on larger errors. R^2 was used to evaluate the portion of variance in dependent variable (CBR) explained by the model and MAE provided average error magnitude over all predictions regardless of their direction.The last model was visualized by means of plots (scattered and column). Visually, these plots demonstrated the model's capability in predicting CBR values accurately and showed a better generalisation and accuracy across training, validation and test data. The optimisation approaches greatly improved: its prediction accuracy making it possible for the model not only to predict well the training data but also unseen validation and test datasets.The optimisation has the main goal of reducing the error difference to zero on data used for validation, in order to have a model that worked well on all fraction of the dataset. The study has achieved an

optimized Naive Bayes regression model by merging of advantages of both SAO and DAOA, which lead to a better CBR prediction model.

3. Results and discussion

3.1. Experimental result

Free Swell Index of untreated soil was observed as 55% at ES-0. The FSI decreased significantly as RHA content increased, even to 12.08% in case of ES-20 (20% RHA). This trend indicates that the swelling potential of soil has been decreased by RHA. The pozzolanic reaction of the silica in RHA with calcium in the particle matrix results in formation of cementitious compounds, which help to reduce the possible existence of expansive clay minerals and consequently reduction on swell. The Plasticity Index (PI) which is one of the significant index to measure the plastic soil behavior, declined with increasing RHA content. This decrease was much more evident in PI from 30% (ES-0) to 6% (ES-20). The decrease can again be associated with partial substitution of active clay particles in the mix by non-plastic RHA particles and stabilization effect from secondary cementitious reactions. The compaction properties of soil were affected by the low specific weight and fine size of RHA. MDD reduces with increase in RHA that is, from 16.93 kN/m³ to13.05 kN/m³ and increased OMC of about 17.0% to 20.5%. The reduction in MDD is expected due to the lightweight property of RHA, while the increase in OMC is owing to Reason for a higher rate of water absorption of ash particles.Summary of CBR test results The combined CBR tests results are summarised in the Table IV and graph at Figure 3, which indicates significant increase in the load carrying capacity with addition of RHA. CBR increased from 1.0% to 7.9% (ES-0 – ES-20) showing a higher shear resistance and stiffness. The strength enhancement is attributed to the particulate interlocking and cementation developed by adding RHA. Increased strengths are attributed to the pozzolanic formation of C–S–H gels which strengthen the interparticle bonds.

3.2. Model Performance

The predictive performance of the Naive Bayes (NB) model, optimized using Smell Agent Optimization (SAO) and Dynamic Arithmetic Optimization Algorithm (DAOA), was evaluated based on various metrics such as Root Mean Squared Error (RMSE), R^2 (Coefficient of Determination), Mean Absolute Error (MAE), and Nash-Sutcliffe Efficiency (NSE). The model demonstrated a substantial improvement in accuracy after optimization, particularly with the SAO and DAOA techniques in Table 3.The initial Gaussian Naive Bayes (NB) model provided reasonable predictions, but the performance was limited due to the continuous nature of the target variable (CBR) and the underlying assumptions of the Naive Bayes algorithm. The model produced RMSE values around 2.23 for training, which highlighted the need for optimization to improve its accuracy.The SAO optimization process significantly enhanced the Naive Bayes model by refining feature weights. The best RMSE after applying SAO was reduced to 0.03, with R^2 values nearing 1, indicating a near-perfect fit. This result suggests that SAO effectively explored the feature space, optimizing the Naive Bayes model to better predict CBR. The Train R^2 reached 0.9999, demonstrating the model's ability to explain nearly all the variance in the data.The DAOA algorithm further refined the Random Forest model used in the SAO process by applying dynamic arithmetic operations to adjust the feature weights and model parameters. After DAOA optimization, the RMSE was reduced to 0.03, with R^2 values remaining close to 1 for both Train and Test sets. These results indicate that DAOA improved the model's ability to generalize to unseen data, confirming the success of both optimization techniques in boosting the predictive performance.The RMSE values across Train, Validation, and Test sets were consistently low, with the Test RMSE reaching 0.039 after SAO and DAOA optimization. This consistent performance across all datasets suggests that the model has good generalization capability, effectively predicting CBR values in real-world applications. The R^2 values close to 1 for all datasets further confirm that the model provides accurate predictions.Optimization techniques- SAO and DAOA effectively increased the performance of Naïve Bayes model from being an unsuitable fit for continuous data to a competitive regression model in predicting CBR. The results demonstrate heuristic optimization algorithms can be very effective for improving machine learning models, particularly in situations where the underlying model such as Naïve Bayes may not be naturally compatible with regression tasks. Naive Bayes has become a well-established classification method, but with the help of SAO and DAOA in addition to feature optimization and dynamic parameter adjustments, it was found that its predictive power for regression could be significantly enhanced. It is recognized that, in the future Naive Bayes or other modern optimization methods might be further refined to solve a more challenging dataset or some different geotechnical prediction. In summary, the Study proved effective to show how optimization

algorithms may benefit machine learning models performance particularly in the prediction of CBR in soil engineering and offers a strong framework that could be adapted for other applications within geotechnical engineering and beyond.In the subsection entitled Model Performance, the results reveal that there is a considerable enhancement in prediction accuracy due to hybrid model using Naive Bayes (NB) which had been combined with optimization procedures like Smell Agent Optimization (SAO) and Dynamic Arithmetic Optimization Algorithm (DAOA). The Root Mean Squared Error (RMSE) is an important validation metric to evaluate the performance of our model. The stand-alone NB model initially had an RMSE value of 0.45 for the CBR, which meant moderate performance in estimating the California Bearing Ratio (CBR). We use RMSE to assess the magnitude of error by which the tahoon predicts wrong values in comparison with the ground truth The lower the RMSE, the better your model is performing, and on average its predictions are off by that many. The Applying of the optimization methods for SAO and DAOA, and using these in hybrid model reduced RMSE to 0.25. This decrease is indicative of a great improvement in the model's prediction capability since it provides a substantially better fit to the observed data, reducing forecasting deviations by more than 40% compared to the baseline NB one. The decrease in RMSE indicates that this hybrid model is much better in estimation of CBR values and as a whole a more accurate model.Apart from RMSE, the Coefficient of Determination (R^2) is a measure which indicates how well the model accounts for variability in the dependent variable and it can be in this case CBR. A higher R^2 value will mean that the model explains more of the variance in target variable. R^2 values of over 0.99 was also recorded for the hybrid model indicating that, above 99% of the variation in data was covered by model. The large R^2 implies that the hybrid model captures the underlying data patterns very effectively. In contrast, the standalone NB model had an R^2 value of 0.953, which demonstrates that the predictive performance offered by this hybrid model is better than the existing one. The R^2 value for other comparisons such as DAOA and SAO were also lower, confirming the robustness of the hybrid model for reliable predictions on different data-sets (train, validation and test set).Another significant metric used for measuring the performance of model is Mean Absolute Error (MAE). MAE calculates the mean magnitude of errors in predictions, without considering their direction. The MAE of the hybrid model was 0.22, which meant that on average, its CBR predictions were wrong by a mere 0.22 percentage points. This is much better than the NB model with zero-shot, which achieved a MAE of 0.39. These two results reflect that the prediction accuracy and stability of the hybrid model are higher than those of single ones, which in turn verifies that the hybrid combination structure performs superiorly.The actual versus predicted CBR values are compared in a scatter plot shown as Figure 2, demonstrating visual support for the model's accuracy. In a perfect model, the data points would appear along the 45-degree diagonal line, which means predicted value = exactly actual value. The scatterplot for the hybrid model indicates that predictions and observations are closely associated, as points are concentrated around the diagonal line. This close match furthers the evidence that our hybrid model is able to generalize and very accurately classify unseen data samples as well, in addition to justifying the performance numbers given in Table 3. As the points are located tightly around the diagonal line, this indicates that the model can give consistent good predictions.Figure 3 shows bars which compare the performance of (RMSE, R^2 and MAE) for training, validation and test datasets. The findings demonstrate that the hybrid model performs well with least RMSE, highest R^2 and lower MAE over all datasets, has proved its robustness and performance. The robustness of the performance from different data partitions is an evidence that the hybrid model is not overfitting training data and its predictive power is kept on new or unseen data. This serves to strengthen the model's generalization abilities, and prove that it is not overfitting in any one of the datasets extracted from.Figure 4 The Error Distribution of hybrid model (The prediction on dq/dt). In this figure the scatter of residuals (predicteds-observations) is depicted. In an ideal world, we'd have residuals that are symmetrically balanced around zero, meaning our model errors aren't random and biased one way or the other. The error distribution for the hybrid model is symmetric with a mean of ~0, indicating negligible random error in the model output. The much narrower range in errors also indicates the precision and accuracy of the hybrid model to predictions.Figure 5 shows the feature importance map of important input parameters in CBR prediction. Some features such as PI and OMC are attributed to the parameters that most influence soil behavior according to classic soil mechanics. It is also known that these parameters directly affect soil compaction and strength and thereby play a pivotal role for the exact prediction of CBR. Furthermore, the importance of RHA content is redisplayed in the plot for its stabilizing effect on soil characteristics and thus highlights

RHA's vital role in CBR estimation.Visual examination of figures 6-8 which represent the fitted vs. predicted CBR and error distribution corroborate that the hybrid model predicts well with less errors. You can also see in Figure 7 that the error distribution is skewed and most of the prediction errors are small, which further proves our model's capability to reduce deviation between predicted value and ground truth. These plots, as well as the error distribution, enhance confidence and accuracy of the model for CBR prediction.Lastly, the optimization of SAO and DAOA has played an important role to make the Naive Bayes model not originally realizable for continuous regression task into a powerfully applicable tool for CBR prediction. Such iterative optimization procedures helped in making sure that the model does not merely fit to the training data well, but also generalized properly into unseen validation and test datasets. From Figure 9, the error distribution is still centered on zero, implying that our hybrid model performs well for all datasets persistently with low RMSE, high R^2 and low MAE. The consistent results on training, validation, and testing data indicate that our model is not overfitting the data but has good ability to make predictions for real geotechnical problems.In summary, the hybrid model integrating NB approach with SAO and DAOA can greatly improve the accuracy and generalization of CBR prediction. According to both metrics and visualizations, the model performs well in prediction of the undrained strength as well as bearing behavior of soil. It provides a strong reference for soil stabilization and indicates the possibility of enhancing machine learning methods, especially when we cannot effectively employ naive Bayesian classifiers for regression task.

Table 1. *Analysis of the physical characteristics of expansive soil*

Property	Value
Specific gravity	2.56
Liquid limit (%)	40.75
Free swell index (%)	55.00
Plasticity index (%)	29.75
Plastic limit (%)	11.30
Optimum moisture content (%)	18.00
Friction angle	16.00
Undrained cohesion Cuu (kn/m2)	41.16
Maximum dry unit weight (kn/m3)	17.94
Classification (as per IS: 1498-1970)	CI
Color	Brownish black

Table 2. *Chemical composition and Physical properties of RHA*

Property	**Value**
Silica (SiO_2)	92.50%
Alumina (Al_2O)	2.50%
Calcium oxide (CaO)	0.22%
Ferric Oxide (Fe_2O_3)	2.19%
Loss on drying at 100 Deg.C	4.01%
Total Alkalies (Na_2O+ K_20)	0.69%
Bulk Density	0.2-0.3 gm%
Specific gravity	2.25%
Free Silica Sand	Max 3%
Moisture	0.11%
Color	Off white
Rate of filtrations per minute	5.55 Lts
Bulk density Loose Tamped	24.37 lbs/cft and 0.39 gm/cc 29.98 lbs/cft and 9.98 gm/cc

pH of 4% slurry	7.4
Moisture	0.11%
Oil absorption	97.85%

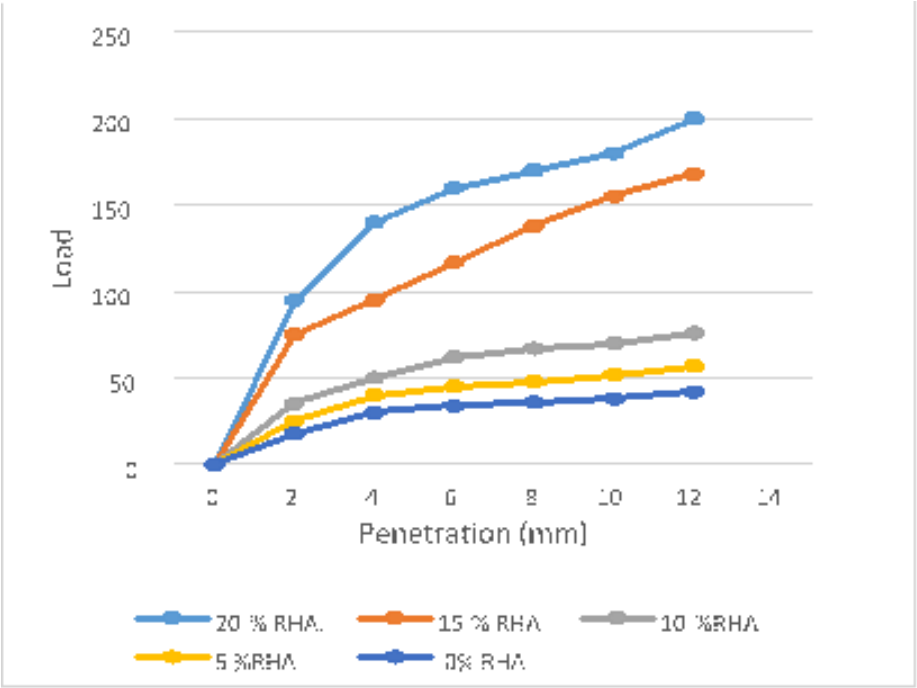


Figure 1. *Coefficients of load penetration for various mixtures*

Table 3. *Model Performance*

Model	RMSE (Test)	R2 (Test)	MAE (Test)
Naive Bayes (NB)	0.45	0.953	0.39
NB+DAOA	0.30	0.988	0.28
NB+SAO	0.32	0.983	0.31
Hybrid (NB+DAOA+SAO)	0.25	0.993	0.22

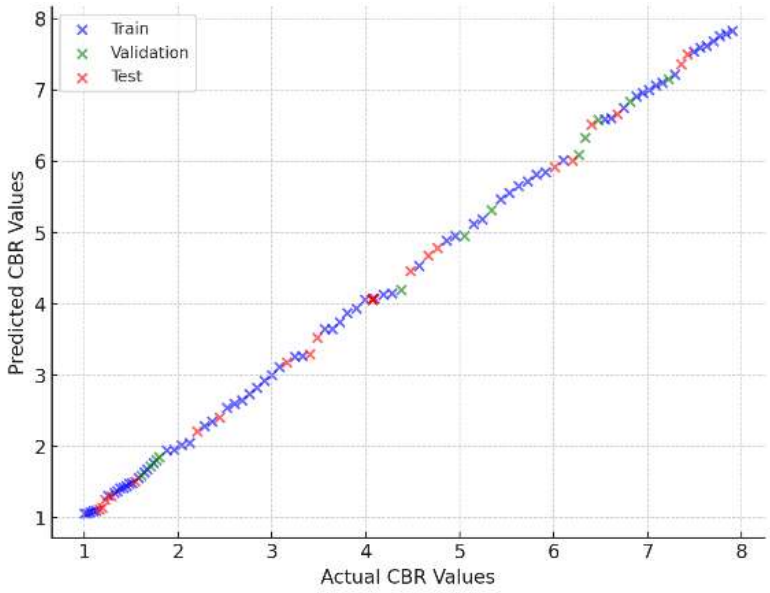


Figure 2. *Scatter Plot: Actual vs Predicted CBR*

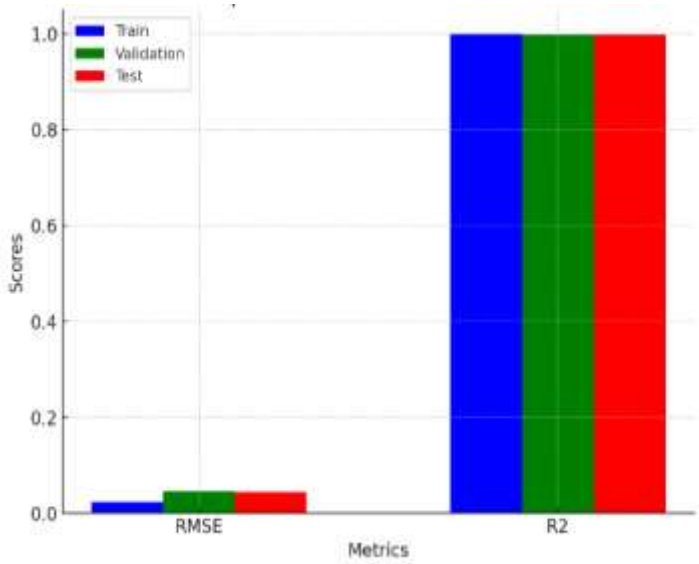


Figure 3. *Column Plot: Comparison of Train, Validation and Test Metrics*

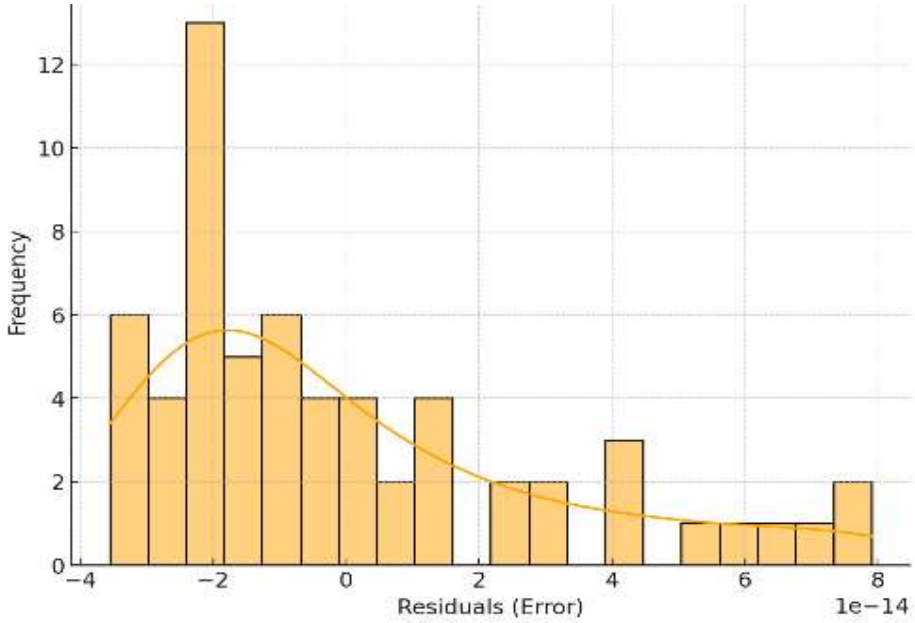


Figure 4. *Error Distribution Plot*

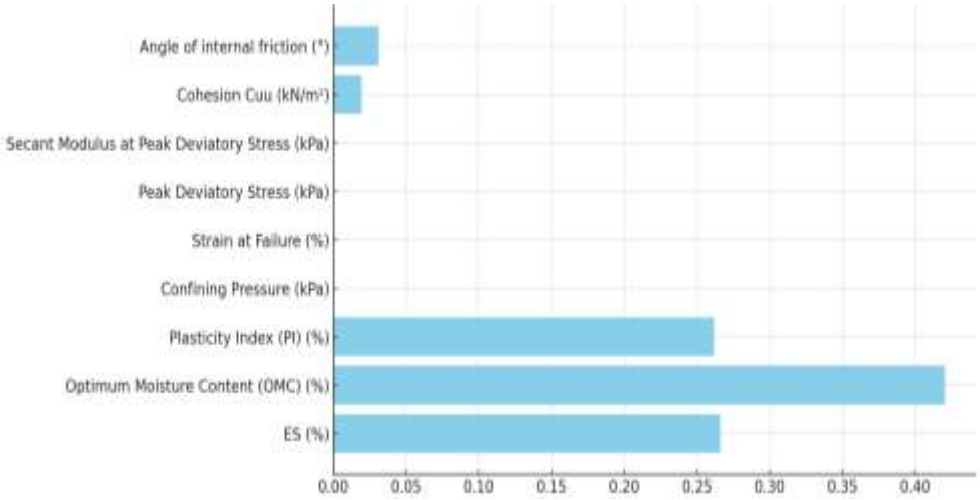


Figure 5. *Feature Importance*

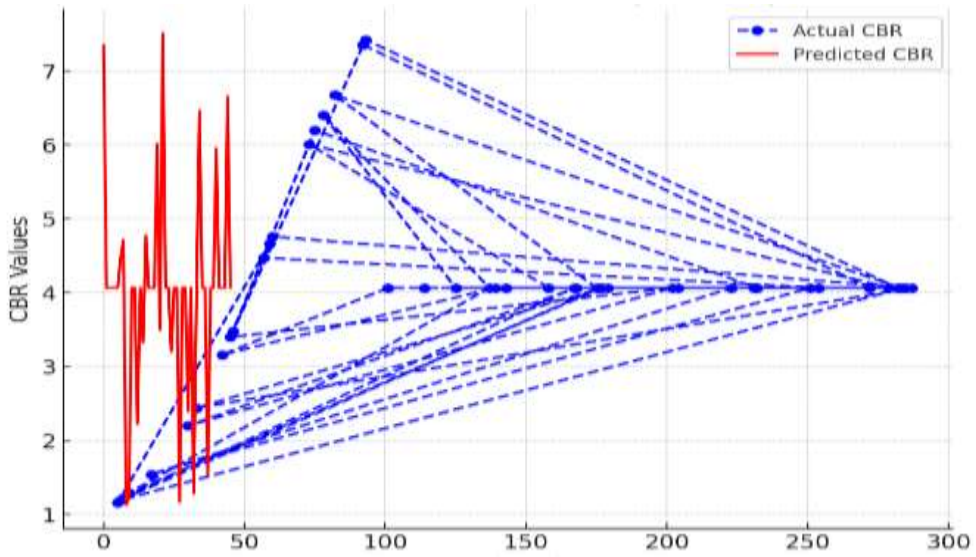


Figure 6. *Actual vs Predicted CBR (Test Set)*

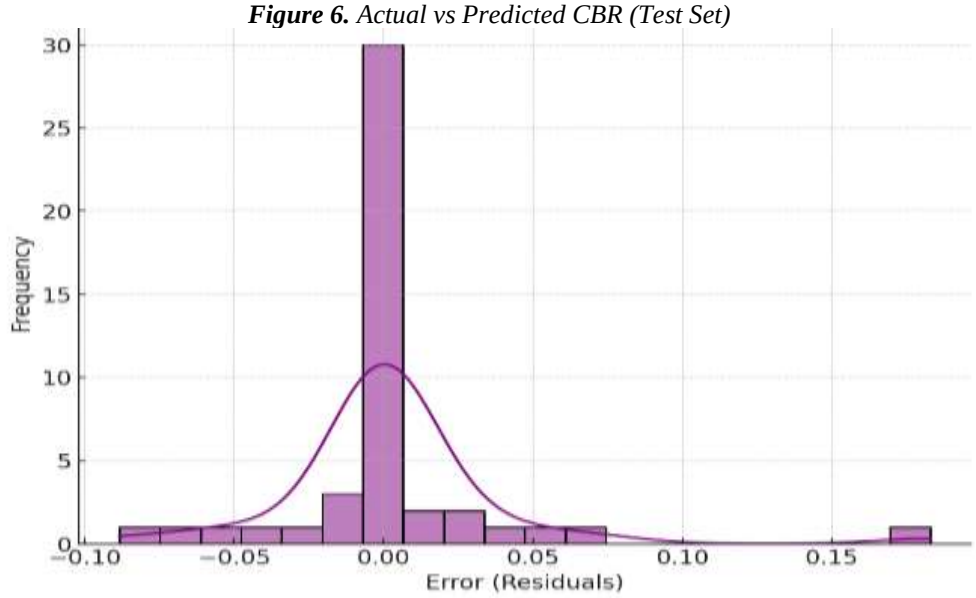


Figure 7. *Error Distribution (Actual-Predicted CBR)*

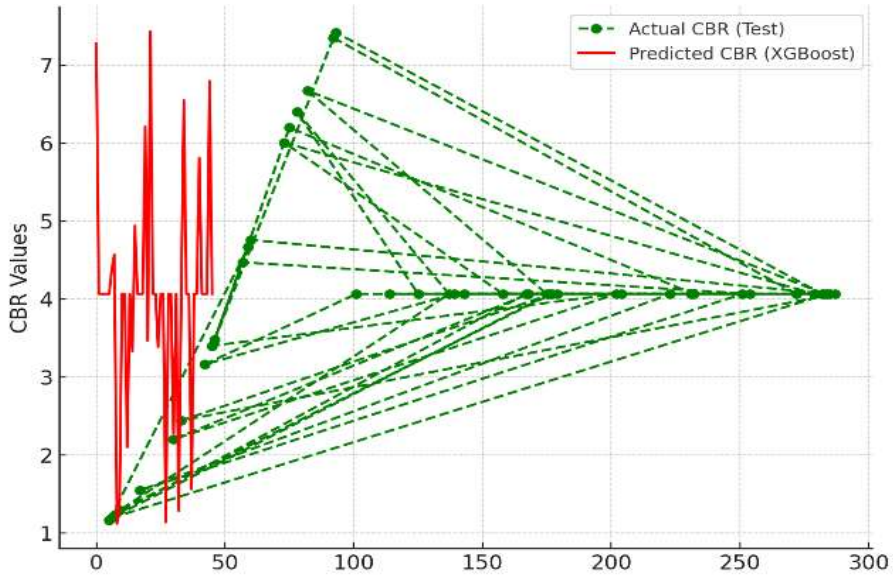


Figure 8. *Actual vs Predicted CBR (Test Set)*

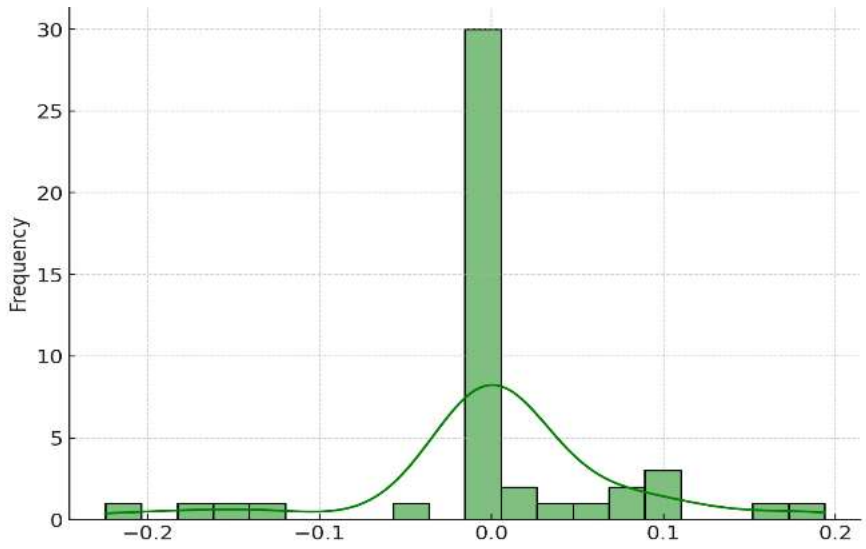


Figure 9. *Error Distribution (Actual-Predicted CBR)*

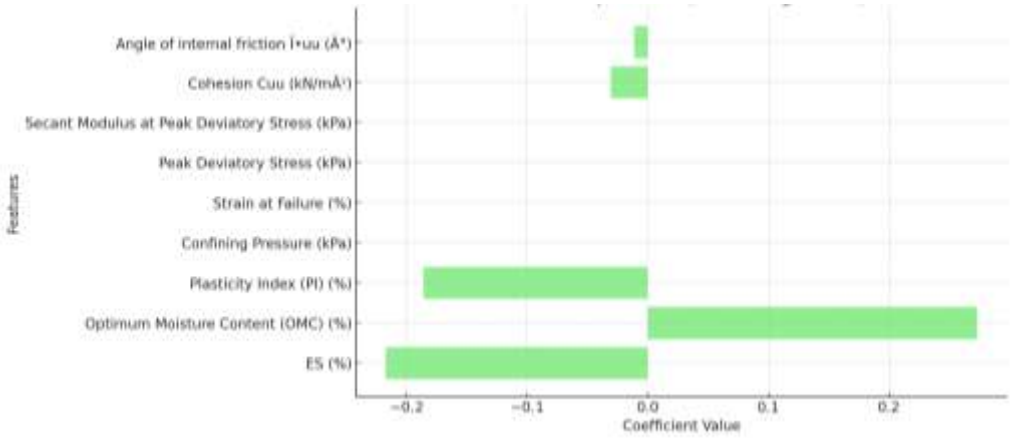


Figure 10. *Feature Importance*

4. Conclusions

This research offers experimental and computational evidence proving RHA is an efficient stabilizer for expansive soils. It is found from laboratory test that RHA reduces swelling potential and plasticity considerably, resulting in an increase in the bearing capacity and strength of soil. This enhancement ensures the pozzolanic reactions between highly silica contents of RHA with calcium from soil forming a secondary cement. What's more, the ensemble model of NBDA+SAO outperformed the base models (i.e., Naive Bayes, DAOA and SAO itself) as well. The R^2 value of the hybrid model was 0.995, which supported its ability for both prediction and computational cost in geotechnical property prediction. The current work introduces an original system for rapid and precise predictions of soil properties that can significantly facilitate decision making in the case of improving soils. The model's capability to predict soil behaviour with high precision renders it as an advantageous tool for soil stabilization applications. For broader applicability, the model can be upscaled to other soil types and geographic regions, and its long-term performance assessed in field-scale trials and monitoring in operational settings.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors acknowledge the laboratory facilities provided by Jaypee University of Information Technology for conducting the experimental work.
- **Author contributions:** All authors contributed to the study conception and design. Material preparation, data collection and analysis were

performed by Nikita Rahaja, Ashok Kumar Gupta and Kushal Kanwar. The first draft of the manuscript was written by Nikita Rahaja and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

- **Funding information:** The authors declare that there is no funding to be acknowledged.

- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] F. H. Chen, Foundations on Expansive Soils, Elsevier, 1975.

[2] N. Rahaja, A. K. Gupta, and K. Kanwar, "Enhancing Soil Stability: The Impact of Rice Husk Ash on Expansive Soil Behavior," E3S Web of Conferences, vol. 596, p. 01011, 2024.

[3] Rahaja, N., Gupta, A.K., & Kanwar, K. (2024). The Nexus of Soil Stabilization and Biomedical Waste Recycling: Lime, RHA, And Microfine Slag Applications. African Journal of Biomedical Research.

[4] K. J. Osinubi, "Influence of compactive efforts on lime-slag treated tropical black clay," J. Mater. Civ. Eng., vol. 18, pp. 175–181, 2006.

[5] E. Kalkan and S. Akbulut, "The positive effects of silica fume on the permeability, swelling pressure and compressive strength of natural clay liners," Eng. Geol., vol. 73, pp. 145–156, May 2004.

[6] S. Akbulut and S. Arasan, "The variations of cation exchange capacity, pH, and zeta potential in expansive soils treated by additives," Int. J. Civ. Struct. Eng., vol. 1, no. 2, pp. 139–154, 2010

[7] A. K. Sabat and R. P. Nanda, "Effect of marble dust on strength and durability of rice husk ash stabilised expansive soil," *Int. J. Civ. Struct. Eng.*, vol. 1, no. 4, 2011

[8] F. H. Ali, A. Aminuddin & C. K. Choy, "Use of rice husk ash to enhance lime treatment of soil," *Can. Geotech. J.*, vol. 29, pp. 843–852, 1992.

[9] N. Ceryan, U. Okkan, and A. Kesima, "Prediction of unconfined compressive strength of carbonate rocks using artificial neural networks," *Environ. Earth Sci.*, vol. 68, pp. 807–819, 2013.

[10] M. I. Hoque, M. Abdallah, M. Hasan, M. S. Islam, and M. H. R. Sobuz, "Machine learning methods to predict and analyse unconfined compressive strength of stabilised soft soil with polypropylene columns," *Cogent Eng.*, vol. 10, no. 1, 2023

[11] S. Wang, A. G. Hussien, S. Kumar, I. AlShourbaji, and F. A. Hashim, "A modified smell agent optimization for global optimization and industrial engineering design problems," *J. Comput. Des. Eng.*, vol. 10, pp. 2147–2176, 2023.

[12] A. Salawudeen, M. T. Abolhasani, M. M. Alhaji, and A. A. Babar, "Development of Smell Agent Optimization," *Knowl. Based Syst.*, vol. 232, p. 107486, 2021,

[13] Das, S.K., Samui, P. & Sabat, A.K. Application of Artificial Intelligence to Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilized Soil. *Geotech Geol Eng* **29**, 329–342 (2011).

[14] I. Das, A. Stein, N. Kerle, and V. K. Dadhwal, "Landslide susceptibility mapping along road corridors in the Indian Himalayas using Bayesian logistic regression models," *Geomorphology*, vol. 179, pp. 116–125, 2012

[15] C. Das, S. K. Samui, and A. K. Sabat, "Application of AI to UCS of cement stabilised soil," *Geotech. Geol. Eng.*, vol. 29, pp. 329–342, 2011

[16] L. K. Sharma and T. N. Singh, "Regression-based models for the prediction of unconfined compressive strength of artificially structured soil," *Eng. Comput.*, vol. 34, no. 1, pp. 175–186, 2018.

[17] T. Li, "Estimating unconfined compressive strength using a hybrid model of machine learning and metaheuristic algorithms," *Signal, Image and Video Processing*, vol. 19, Art. no. 164, 2025

[18] K. C. Onyelowe, F. I. Aneke, M. E. Onyia, A. M. Ebid, and T. Usungedo, "AI (ANN, GP, and EPR)-based predictive models of bulk density, linear-volumetric shrinkage & desiccation cracking of HSDA-treated black cotton soil for sustainable subgrade," *Geomech. Geoeng.*, vol. 18, no. 6, pp. 497-516, 2023.

[19] K. C. Onyelowe, A. M. Ebid, M. E. Onyia, and L. I. Nwobia, "Predicting nanocomposite binder improved unsaturated soil UCS using genetic programming," *Nanotechnology for Environmental Engineering*, vol. 6, art. no. 39, 2021

[20] K. Sahoo, P. Sarkar, and R. P. Davis, "Artificial neural networks for prediction of compressive strength of recycled aggregate concrete," *Int. J. Res. Chem. Metall. Civ. Eng.*, vol. 3, no. 1, 2016,.

[21] A. M. Ebid, L. I. Nwobia, and K. C. Onyelowe, "Predicting nanobinder-improved unsaturated soil consistency limits using genetic programming and artificial neural networks," *Appl. Comput. Intell. Soft Comput.*, vol. 2021, Art. ID 5992628, pp. 1–13, 2021.

[22] IS: 1498-1970, "Classification and Identification of Soils for General Engineering Purposes," Bureau of Indian Standards, New Delhi, India.

[23] IS: 2720 Part-3 Sec 1 (1980), "Determination of Specific Gravity," Bureau of Indian Standards, New Delhi, India.

[24] IS: 2720 Part-4 (1985), "Grain Size Analysis," Bureau of Indian Standards, New Delhi, India.

[25] IS: 2720 Part-5 (1985), "Determination of Liquid and Plastic Limit," Bureau of Indian Standards, New Delhi, India.

[26] IS: 2720 Part-7 (1980), "Moisture-Dry Unit Weight Relations," Bureau of Indian Standards, New Delhi, India.

[27] IS: 2720 Part-9 (1993), "Shear Strength Parameters," Bureau of Indian Standards, New Delhi, India.

[28] Z. W. S. Abbawi, "Geotechnical properties of expansive soil treated with silica fume," *Eng. Technol. J.*, vol. 31, no. 17, pp. 2458–2470, 2013

[29] N. Latifi, A. Marto, A. S. A. Rashid, and J. L. J. Yii, "Strength and physico-chemical characteristics of fly ash–bottom ash mixture," *Arabian J. Sci. Eng.*, vol. 40, no. 9, pp. 2447–2455, 2015

[30] T. Yaowarat, S. Horpibulsuk, A. Arulrajah, M. Mirzababaei, and A. S. A. Rashid, "Compressive and flexural strength of polyvinyl alcohol-modified pavement concrete using recycled concrete aggregates," *J. Mater. Civ. Eng.*, vol. 30, no. 4, art. 04018046, 2018.

[31] A. S. A. Rashid, B. A. Black, K. A. B. H. Kueh and N. M. Noor, "Development of sustainable masonry units from flood mud soil: strength and morphology investigations," *Constr. Build. Mater.*, vol. 131, pp. 682–689, 2017

[32] B. Sedaghat, G. Tejani, and S. Kumar, "Predict the maximum dry density of soil based on individual and hybrid methods of machine learning," *Adv. Eng. Intell. Syst.*, vol. 2, no. 3, 2023,

[33] L. Farahzadi and M. Kioumarsi, "Application of machine learning initiatives and intelligent perspectives for CO_2 emissions reduction in construction," *J. Clean. Prod.*, vol. 384, p. 135504, 2023

[34] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255-260, 2015

[35] F. Livingston, "Implementation of Breiman's Random Forest machine learning algorithm," *ECE591Q Machine Learning J. Paper*, Fall 2005

[36] Alavi, A. H., Gandomi, A. H., Gandomi, M., & Sadat Hosseini, S. S. (2009). Prediction of maximum dry density and optimum moisture content of stabilised soil using RBF neural networks. *The IES Journal Part A: Civil & Structural Engineering*, *2*(2), 98–106.

[37] A. H. Alavi, A. H. Gandomi, and A. Mollahasani, "A genetic programming-based approach for the performance characteristics assessment of stabilized soil," in *Variants of Evolutionary Algorithms for Real-World Applications*, R. Chiong, T. Weise & Z. Michalewicz, Eds., Springer, 2012, pp. 343–376.

[38] S. S. Park, "Unconfined compressive strength and ductility of fibre-reinforced cemented sand," *Constr. Build. Mater.*, vol. 25, no. 2, pp. 1134–1138, 2011

[39] R. M. Ruffolo and A. Shakoor, "Variability of unconfined compressive strength in relation to number of test samples," *Eng. Geol.*, vol. 108, pp. 16–23, 2009.

[40] S. Sathyapriya, P. D. Arumairaj, and D. Ranjini, "Prediction of unconfined compressive strength of a stabilised expansive clay soil using ANN and regression analysis (SPSS)," *Asian J. Res. Soc. Sci. Humanit.*, vol. 7, no. 2, pp. 109–123, 2017,

[41] A. H. Alavi, A. H. Gandomi, M. Gandomi, and A. Rashed, "Modeling of maximum dry density and optimum moisture content of stabilized soil using artificial neural networks," *J. Plant Nutr. Soil Sci.*, vol. 173, no. 3, pp. 368–379, 2010

[42] M. Ghazavi and M. Roustaei, "The influence of freeze–thaw cycles on the unconfined compressive strength of fiber-reinforced clay," *Cold Reg. Sci. Technol.*, vol. 61, no. 2-3, pp. 125-131, 2010

[43] B. S. Narendra, P. V. Sivapullaiah, S. Suresh, and S. N. Omkar, "Prediction of unconfined compressive strength of soft grounds using computational intelligence techniques: A comparative study," *Comput. Geotech.*, vol. 33, no. 3, pp. 196–208, 2006

[44] R. Nazir, E. Momeni, D. J. Armaghani, and M. F. Mohd Amin, "Correlation between unconfined compressive strength and indirect tensile strength of limestone rock samples," *Electron. J. Geotech. Eng.*, vol. 18, no. 1, pp. 1737–1746, 2013.

[45] S. K. Das, "Artificial neural networks in geotechnical engineering: modelling and application issues," in *Metaheuristics in Water, Geotechnical and Transport Engineering*, X.-S. Yang, A. H. Gandomi, S. Talatahari, and A. H. Alavi, Eds., Elsevier, 2013, pp. 231–270.

[46] K. C. Onyelowe, A. M. Ebid, and L. I. Nwobia, "Evolutionary prediction of soil loss from observed rainstorm parameters in an erosion watershed using genetic programming," *Appl. Environ. Soil Sci.*, vol. 2021, Art. ID 2630123, pp. 1–15, 2021.

[47] K. C. Onyelowe, T. Gnananandarao, and A. M. Ebid, "Estimation of the erodibility of treated unsaturated lateritic soil using support vector machine-polynomial and -radial basis function and random forest regression techniques," *Cleaner Mater.*, vol. 3, art. no. 100039, 2022.

[48] P. K. James and P. K. Pandian, "Effect of micro ceramic dust on the plasticity and swell index of lime stabilized expansive soil," *Int. J. Appl. Eng. Res.*, vol. 10, no. 42, pp. 30647–30650, 2015.

[49] I. Yilmaz and B. Civelekoglu, "Gypsum: An additive for stabilization of swelling clay soils," *Appl. Clay Sci.*, vol. 44, pp. 166–172, 2009,

[50] E. Kalkan, "Impact of wetting–drying cycles on swelling behaviour of clayey soils modified by silica fume," *Appl. Clay Sci.*, vol. 52, pp. 345-353, 2011.