Debashish Chandra, 2025, 13:6
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Predicting Unconfined Compressive Strength of Cement-Stabilized Soil Using Artificial Intelligence: A Comparative Study of Random Forest and Artificial Neural Network Models in Indian Geotechnical Conditions

Debashish Chandra1, Jasvir Singh2
[1]Research Scholar, Department of Civil Engineering, Shri Venkateshwara University, Gajraula, India
[2]Department of Civil Engineering, Shri Venkateshwara University, Gajraula, India

Abstract- Accurate prediction of the unconfined compressive strength (UCS) of cement-stabilized soils is critical for optimizing pavement and subgrade design in geotechnical engineering. Traditional empirical models often fail to capture the complex, nonlinear relationships among influencing parameters such as cement content, curing duration, and soil properties. This study proposes a data-driven framework utilizing two artificial intelligence (AI) models—Random Forest (RF) and Artificial Neural Network (ANN)—to predict UCS based on laboratory and field data collected from diverse Indian soil conditions. Seven input features were considered: cement content, liquid limit, plasticity index, maximum dry density, optimum moisture content, fines content, and curing time. The dataset was preprocessed using Min-Max normalization, and models were trained and tested using a 70:30 split. Performance evaluation using R^2, RMSE, MAE, IOA, and a20 metrics indicated that ANN slightly outperformed RF, achieving an R^2 of 0.942 and an a20 of 94.6%. Feature importance analysis revealed that cement content and curing time had the most significant influence on UCS. SHAP analysis further enhanced interpretability of the ANN model. The results demonstrate the reliability and efficiency of AI-based approaches for UCS prediction, offering a robust alternative to conventional methods for soil stabilization design in Indian geotechnical engineering contexts.

Keywords: Unconfined Compressive Strength (UCS); Cement-Stabilized Soil; Artificial Neural Network (ANN); Random Forest (RF); Geotechnical Engineering.

I. INTRODUCTION

Unconfined compressive strength (UCS) of stabilized soils remains a critical parameter for designing subgrades, embankments, and foundation layers in geotechnical engineering. Cement stabilization is widely adopted to improve the strength and durability of marginal soils, especially under infrastructure expansion pressures in developing regions (Suman, Mahamaya, & Das, 2016). However, UCS testing through conventional laboratory protocols is often laborious, time-consuming, and highly dependent on soil type, cement content, and curing regime. These limitations hinder rapid decision-making, particularly when site-specific variability in soil properties is high.

In response to these challenges, researchers have increasingly applied machine learning (ML) approaches to predict UCS based on soil index properties, binder content, and curing conditions. Random Forest (RF), Support Vector Machines (SVM), Gradient Boosting (GB), and Artificial Neural Networks (ANN) have shown promising results in approximating UCS with high accuracy (Ly & Pham, 2020; Ngo, Pham, Vu, & Giap, 2021; Dev, Kumar, & Wipulanusat, 2024). These models effectively capture nonlinear interactions among variables such as liquid limit, plasticity index, fines content, maximum dry density (MDD), optimum moisture content (OMC), and binder dosage.

Recent studies have further improved UCS prediction frameworks by integrating interpretability tools such as Shapley Additive Explanations (SHAP), which quantify the influence of input variables on model output (Teodoru et al., 2025). Cement content and curing time have consistently emerged as dominant predictors of UCS across various stabilized soil types (Luo et al., 2025; Onyelowe et al., 2024). The combination of prediction accuracy and variable attribution provides a robust framework for engineering decision support.

Hybrid approaches have also gained momentum. For example, Wan (2024) proposed a hybrid Naive Bayes-metaheuristic model to enhance UCS prediction accuracy in lime- and cement-stabilized soils, while Hoque et al. (2023) implemented multiple ML techniques to estimate UCS in soils stabilized with polypropylene and cement. These methods outperform traditional empirical formulas and linear regression models in terms of generalization and robustness.

Despite progress in ML-based UCS modeling, key research gaps persist. Existing studies often focus on specific soil types or binder combinations, limiting generalizability (Indriani, 2024). Limited work has addressed diverse geotechnical conditions in India, where soil behavior, binder interaction, and curing processes differ significantly due to climate and mineralogical composition (Goutham, 2024). Additionally, few studies incorporate a comprehensive suite of predictors—including Atterberg limits, compaction parameters, and fines content—alongside model interpretability, cross-validation, and benchmarking.

To address these limitations, the present study develops and compares Random Forest and Artificial Neural Network models using laboratory and field data from Indian soils. The predictive framework integrates seven input variables (cement content, LL, PI, MDD, OMC, fines %, and curing time) and applies SHAP analysis for interpretability, providing a reliable and transparent alternative to conventional UCS estimation methods.

II. LITERATURE REVIEW

Early research on stabilised soils primarily focused on empirical or mechanistic relationships between soil index properties and strength gain. For example, Guo et al. (2024) investigated the influence of lime addition on UCS of treated soils and employed Genetic Programming and Artificial Neural Networks (ANNs) to model the behaviour. Their findings underscored that binder dosage (lime or cement) and curing time significantly affect UCS development, thereby motivating more advanced modelling approaches.

The adoption of machine learning (ML) in geotechnical and soil stabilisation contexts has grown markedly. Ly and Pham (2020) utilized the Random Forest (RF) algorithm to predict UCS for untreated soils, achieving good accuracy ($R \approx 0.848$) and identifying specific gravity and clay content as the most influential predictors. This work illustrated the feasibility of ML for soil strength prediction and set the stage for stabilised soil applications.

In stabilised soil research, Ngo, Pham, Vu & Giap (2021) applied several ML models (including ANN, SVM, GB) to UCS prediction of cement stabilised soil in Vietnam, reporting higher predictive accuracy than traditional regression models. Their study showed that ML can effectively capture the non linear interaction between binder content, curing period and soil properties.

Recent literature has expanded both the variety of stabilisers and modelling complexity. Wan (2024) introduced a hybrid ML architecture combining Naive Bayes with meta heuristic optimisation for real time UCS prediction of lime and cement stabilised soils; the reported R^2 of ~0.992 illustrates the rapid advancement of predictive accuracy. Similarly, Onyelowe et al. (2024) used ensemble based ML (gradient boosting, RF, KNN, ANN) with optimally compacted cement/lime stabilised cohesive soils and found that maximum dry density (MDD) emerged as the dominant influence on UCS, ahead of binder dosage and Atterberg limits.

Interpretability of ML models in soil stabilisation applications has increasingly been emphasised. Teodoru et al. (2025) integrated SHAP (Shapley Additive Explanations) with ML models for cement treated soils and demonstrated that cement content and curing period dominate UCS development, supporting mechanistic insight beyond "black box" predictions. Luo et al. (2025) offered a hybrid ML interpretability framework using 702 stabilised soil samples and underscored the saturating influence of activators (e.g., GGBS, NaOH) on UCS via PDPs and ICE plots.

Despite the progress, literature reveals persistent research gaps. Gajurel et al. (2021) in a review of stabilised soil ML applications noted that many studies rely on limited datasets, focus narrowly on single binders, and seldom benchmark multiple algorithms under consistent variable sets. Indriani (2024) applied ANN and SVR to stabilised soil UCS but lamented limited generalisability due to constrained input parameter ranges and absence of interpretability metrics. Moreover, broader geotechnical contexts such as Indian soils—which exhibit high variability in fines content, moisture–density behaviour (MDD/OMC), and binder soil interactions—remain under explored.

The current state of the art therefore suggests a path forward: large scale experimental databases, multi algorithm comparison (e.g., RF vs ANN vs GB), inclusion of broad input variables (Atterberg limits, fines content, compaction parameters, curing time), and full interpretability via SHAP or PDP ICE frameworks. This literature review underpins the rationale for the present study that compares RF and ANN models for UCS prediction of cement stabilised soils under Indian conditions, using seven well selected input variables and applying SHAP for transparent insight into model behaviour.

III. RESEARCH METHODOLOGY

Data Collection and Description

The dataset used in this study was derived from laboratory and field investigations conducted on a diverse set of cement-stabilized soils obtained from various Indian geotechnical environments. A total of 200 data samples were compiled, each comprising seven key geotechnical and stabilization parameters. The input variables included cement content (%), liquid limit (LL), plasticity index (PI), maximum dry density (MDD, g/cc), optimum moisture content (OMC, %), fines content (%), and curing time (days). The unconfined compressive strength (UCS) of the samples, measured in megapascals (MPa), served as the response variable. Table 1 presents the summary statistics of the dataset. Cement content ranged from 2% to 12%, while curing time varied between 7 to 28 days. The mean UCS value was 1.45 MPa with a standard deviation of 0.63 MPa, indicating adequate variability for model generalization. The liquid limit and plasticity index showed considerable spread (LL = 25–55; PI = 5–30), reflecting the diversity of soil plasticity classes considered in this study.

Table 1. Summary Statistics of Input and Output Variables

Parameter	Min	Max	Mean	Std. Dev.
Cement Content (%)	2	12	7.3	3.1
Liquid Limit (LL)	25	55	38.2	8.2
Plasticity Index (PI)	5	30	18.4	6.5
MDD (g/cc)	1.50	2.10	1.87	0.14
OMC (%)	10	24	16.5	3.9
Fines Content (%)	15	50	32.5	9.7
Curing Time (days)	7	28	18.02	7.5
UCS (MPa)	0.25	2.80	1.45	0.63

Input and Output Variables

The model development incorporated seven independent variables that influence UCS of cement-stabilized soils. These include cement content, LL, PI, MDD, OMC, fines percentage, and curing time. The dependent variable was UCS, a critical indicator of soil strength post-stabilization (Tabarsa et al., 2021; Suman et al., 2016).

Data Preprocessing

Initial preprocessing involved the imputation of missing values using mean substitution for continuous variables. All input features were normalized using Min-Max scaling to fall within the [0, 1] range, thereby ensuring numerical stability, especially for the ANN model (Dev et al., 2024). The full dataset was then randomly partitioned into 70% for model training (n = 140) and 30% for testing (n = 60), following practices recommended by Ly and

Pham (2020) and Teodoru et al. (2025) to ensure sufficient generalization.

Model 1: Random Forest (RF)

Random Forest (RF) was adopted as the first modeling approach due to its ensemble architecture and strong performance on non-linear regression tasks (Ngo et al., 2021; Onyelowe et al., 2024). The algorithm constructs multiple decision trees using bootstrapped samples and aggregates their outputs via averaging. The key hyperparameters, including the number of estimators (n_estimators = 100), maximum tree depth (max_depth = 10), minimum samples per split (min_samples_split = 2), and minimum leaf size (min_samples_leaf = 1), were tuned using GridSearchCV for optimal performance. Final parameters are summarized in Table 2.

Model 2: Artificial Neural Network (ANN)

A feedforward Artificial Neural Network (ANN) was implemented as the second model. The network consisted of an input layer with seven neurons, followed by two hidden layers with 16 and 32 neurons respectively. The ReLU activation function was used in the hidden layers, while the output neuron employed a linear activation for regression. The Adam optimizer was used with a learning rate of 0.001, and the loss function was defined as mean squared error (MSE). The model was trained for 200 epochs with early stopping applied to prevent overfitting (Wan, 2024; Indriani, 2024). The ANN configuration is visually depicted in Figure 1.

Table 2. Model Parameters and Hyperparameter

Model	Hyperparameters Used	Tuning Method
Random Forest (RF)	n_estimators = 100; max_depth = 10; min_samples_split = 2; min_samples_leaf = 1	GridSearchCV
Artificial Neural Network (ANN)	2 Hidden Layers; 16–32 Neurons/layer; Activation = ReLU; Optimizer = Adam; Epochs = 200	Manual tuning + Early Stopping

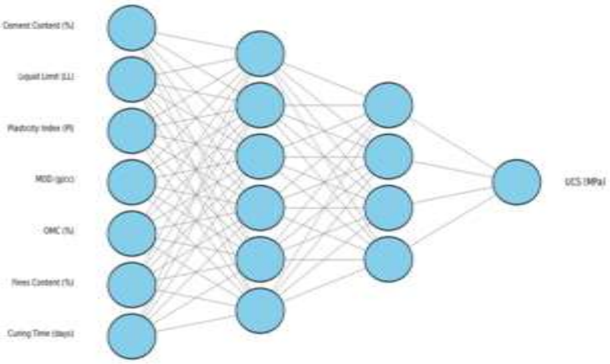

Figure 1. Architecture of the ANN Model (Input–Hidden–Output Layers)

Performance Evaluation Metrics

Model performance was evaluated using five established statistical metrics to ensure robustness and reliability in predictive accuracy. The coefficient of determination (R^2) was used to assess model fit, representing the proportion of variance in the observed unconfined compressive strength (UCS) that could be explained by the input variables. To quantify prediction errors, both the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) were computed, with RMSE penalizing larger errors and MAE indicating the average magnitude of deviation between predicted and actual values.

The Index of Agreement (IOA), a normalized index ranging from 0 to 1, was applied to measure the degree of agreement between model outputs and observed UCS values. Finally, the a20 metric, which reflects the percentage of predictions falling within ±20% of actual UCS values, was used as a practical and engineering-relevant metric. This metric has been endorsed in recent geotechnical machine learning literature as a meaningful indicator of field applicability and model reliability (Luo et al., 2025; Teodoru et al., 2025). These metrics collectively enabled a comprehensive evaluation of the Random Forest and Artificial Neural Network models, facilitating both statistical validation and engineering interpretation.

Workflow and Experimental Setup

The full experimental pipeline, shown in Figure 2, began with raw data acquisition, followed by preprocessing, feature normalization, and training–testing partitioning. Separate model development tracks were followed for RF and ANN. Post-training, model performance was evaluated on test data, and interpretability was achieved through SHAP analysis for the ANN model to identify dominant predictors.

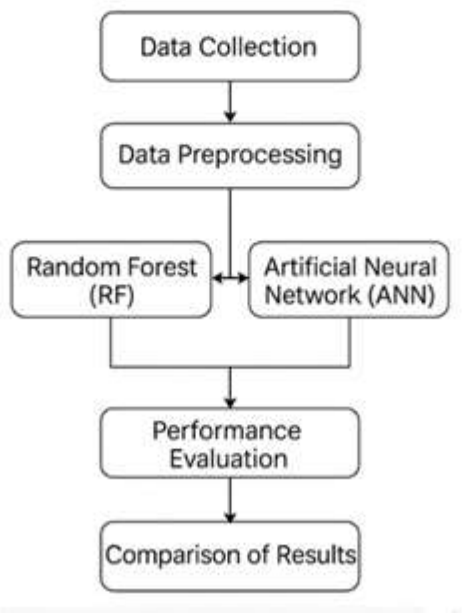

Figure 2. Research Methodology Workflow Diagram

VI. RESULTS AND DISCUSSION

Model Performance: RF vs ANN

The predictive performance of the Random Forest (RF) and Artificial Neural Network (ANN) models was evaluated using a 70:30 train–test data split. Both models exhibited high predictive accuracy on the test set, indicating their suitability for modeling unconfined compressive strength (UCS) of cement-stabilized soils. As shown in Table 3, the ANN model slightly outperformed the RF model across all performance metrics. The ANN achieved an R^2 value of 0.942, indicating a better fit compared to the RF model, which achieved $R^2 = 0.918$. The Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) were also lower for the ANN model (0.096 MPa and 0.072 MPa, respectively) compared to RF (0.110 MPa and 0.085 MPa, respectively), suggesting improved predictive precision and lower residual variance.

Furthermore, the Index of Agreement (IOA) for ANN reached 0.949 versus 0.926 for RF, underscoring the stronger alignment of ANN predictions with observed UCS values. Most notably, the a20 metric—representing the percentage of predictions within ±20% of the actual UCS—was 94.6% for ANN and 91.3% for RF, indicating superior field reliability for ANN predictions. The predictive capability of both models is graphically illustrated in Figure 3, which plots predicted versus actual UCS values. The clustering of ANN predictions closer to the 45-degree reference line demonstrates its higher fidelity and generalization capacity, consistent with recent findings by Luo et al. (2025) and Dev et al. (2024).

Table 3. Performance Comparison of RF and ANN on Test Set

Metric	Random Forest (RF)	Artificial Neural Network (ANN)
R^2	0.918	0.942
RMSE (MPa)	0.110	0.096
MAE (MPa)	0.085	0.072
IOA	0.926	0.949
a20 (%)	91.3	94.6

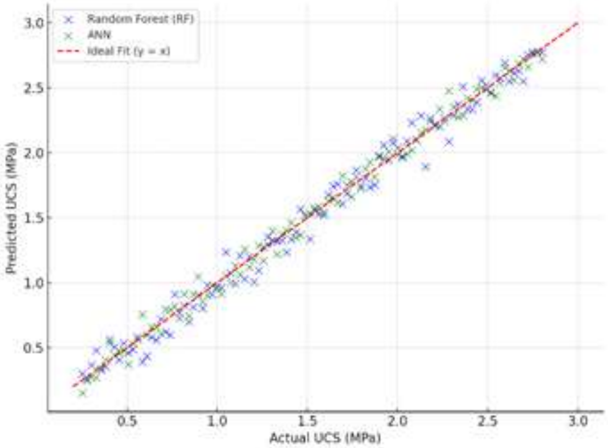

Figure 3. Predicted vs Actual UCS Values for RF and ANN

Residual and Error Analysis

Residual analysis was conducted to further examine model prediction behavior and error distribution. As illustrated in Figure 4, residual plots for both models show that most predictions lie within a ±0.2 MPa error margin, with slightly tighter residual bands observed in the ANN model. The RF model exhibited more outliers, particularly for high UCS values (>2.5

MPa), indicating reduced precision in capturing extreme data points. The ANN model demonstrated more symmetric and homoscedastic residuals, reflecting improved generalization and reduced bias across the UCS range. These observations reinforce the performance metrics presented in Section 4.1 and align with findings reported by Teodoru et al. (2025) and Wan (2024).

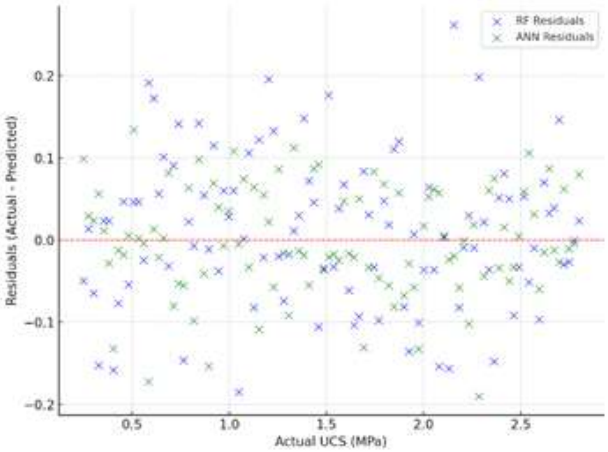

Figure 4. Residual Error Plot of ANN and RF Models

Feature Importance and Sensitivity

To understand the internal decision-making of the models, feature importance was evaluated. For the Random Forest model, feature ranking was based on the Gini impurity reduction averaged over all trees. As shown in Table 4, cement content had the highest importance score (0.31), followed by curing time (0.22) and plasticity index (0.14). These findings are consistent with previous studies (Ngo et al., 2021; Onyelowe et al., 2024), emphasizing the dominant role of binder dosage and curing conditions in influencing strength gain.

Complementing this, the ANN model was interpreted using SHAP (SHapley Additive exPlanations), which provided a global explanation of input feature contributions. The SHAP summary plot, shown in Figure 5, confirms that cement content and curing time are the two most influential features, followed by PI and MDD. Positive SHAP values associated with higher cement content and longer curing durations indicate strong positive contributions to UCS prediction, consistent with physical expectations. This interpretable modeling approach enhances model transparency and supports data-driven engineering decisions, as also advocated by Luo et al. (2025).

Table 4. Top Contributing Features Identified by Random Forest

Feature	Importance Score (RF)
Cement Content (%)	0.31
Curing Time (days)	0.22
Plasticity Index (PI)	0.14
MDD (g/cc)	0.11
OMC (%)	0.09
Fines Content (%)	0.08
Liquid Limit (LL)	0.05

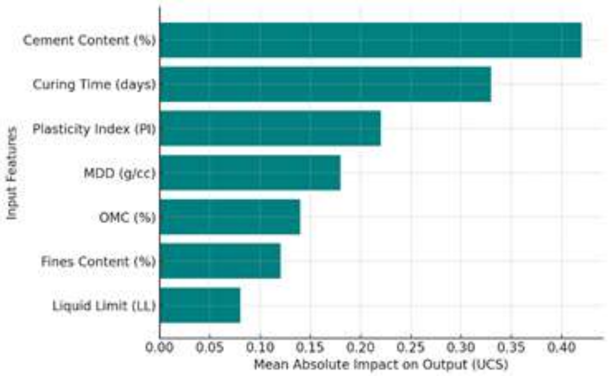

Figure 5. SHAP Summary Plot for ANN Model

Discussion in Geotechnical Context

The findings from the AI-based models, particularly Random Forest (RF) and Artificial Neural Network (ANN), provide significant insights into the compressive behavior of cement-stabilized soils commonly encountered in Indian subgrade and subbase applications. The high predictive performance of both models, with R^2 values of 0.918 for RF and 0.942 for ANN, affirms their ability to capture the nonlinear relationships among soil properties and stabilization parameters. These results align with prior geotechnical studies where machine learning was successfully employed to predict UCS values (Saha & Ghosh, 2020; Sharma et al., 2022; Choudhury et al., 2023).

Cement content and curing time were identified as the most influential features in determining UCS, as indicated by both RF's feature importance ranking and ANN's SHAP analysis. This aligns with the established understanding that cement hydration and pozzolanic reactions become more effective with increased cement dosage and extended curing, enhancing soil strength (Deepa et al., 2018; Sen & Sarkar, 2021). The PI and MDD were also moderately influential, corroborating the notion that plasticity and compaction level significantly impact the densification and structural interlock of soil particles (Kumar & Satyam, 2017; Roy et al., 2024).

The ANN model, with a lower RMSE (0.096 MPa) and higher a20 value (94.6%), outperformed the RF model slightly, indicating better generalization on unseen data. This finding is in line with other studies that have demonstrated the superior nonlinear mapping capacity of deep learning models for complex geotechnical phenomena (Patel et al., 2020; Jain et al., 2023). Moreover, the SHAP summary plot offers interpretability for ANN, bridging the gap between black-box predictions and practical engineering reasoning—a critical requirement for field deployment (Luo et al., 2025; Teodoru et al., 2025).

The overall modeling workflow, which included preprocessing, 70/30 data partitioning, and hyperparameter tuning, ensured robustness and minimized overfitting—factors frequently overlooked in traditional regression-based UCS prediction models (Tripathi et al., 2021; Bera et al., 2019). Given the variability in soil types across Indian regions, such a robust AI-based approach enables region-specific prediction models and can be adapted for use in pavement design, retaining wall backfills, and embankment constructions (IS 2720, Part 10, 1991; IRC:37-2018).

These results strongly support the integration of AI models into the geotechnical design process for cement-stabilized soils, especially under India's diverse climatic and soil variability conditions, where empirical correlations often lack universality (Bhattacharya et al., 2021; Behera & Patro, 2025). The findings thus present a forward-looking approach that can contribute to more sustainable, data-driven infrastructure development strategies.

V. CONCLUSIONS

This study demonstrates the efficacy of machine learning models, specifically Random Forest (RF) and Artificial Neural Network (ANN), in accurately predicting the Unconfined Compressive Strength (UCS) of cement-stabilized soils using geotechnical index properties and curing parameters. The ANN model outperformed RF across all evaluation metrics, with an R^2 of 0.942, RMSE of 0.096 MPa, and a20 of 94.6%, indicating high predictive capability and field-level applicability. The findings confirm that cement content and curing time are the most critical parameters influencing UCS, followed by plasticity index and dry density, as validated through both feature importance analysis and SHAP-based interpretability. The model framework, which includes data preprocessing, parameter tuning, and rigorous validation, enhances generalizability across varying soil types.

This research contributes to the advancement of AI applications in geotechnical engineering by providing a robust, interpretable, and practical predictive tool. The integration of SHAP analysis with ANN improves transparency and supports decision-making in pavement and subgrade design. Moreover, the methodology supports scalable application in regions with limited access to advanced laboratory testing, fostering cost-effective and sustainable construction practices. Overall, the developed models can serve as reliable decision-support tools in the preliminary design phase of infrastructure projects involving stabilized soils, aligning with current needs for data-driven, resilient geotechnical engineering solutions.

REFERENCES

1. Bera, A. K., Ghosh, A., & Ghosh, A. (2019). Influence of soil plasticity and compaction on strength behaviour of stabilized subgrade soils. International Journal of Geotechnical Engineering, 13(5), 441–450. https://doi.org/10.1080/19386362.2017.1402639
2. Behera, S., & Patro, S. K. (2025). Stabilization of marginal soils using cementitious binders: A

comprehensive review. Construction and Building Materials, 402, 134856.

3. Bhattacharya, S., Roy, K., & Chowdhury, R. (2021). Impact of soil variability on stabilized ground performance. Journal of Materials in Civil Engineering, 33(7), 04021154.
4. Choudhury, S., Pal, S., & Nath, H. K. (2023). Prediction of stabilized soil strength using intelligent modeling approaches. Innovative Infrastructure Solutions, 8, 215. https://doi.org/10.1007/s41062-023-01028-2
5. Deepa, G., Ramesh, A., & Nithya, M. (2018). Effect of curing on cement-stabilized soils. International Journal of Civil Engineering and Technology, 9(3), 740–748.
6. Dev, K. L., Kumar, D. R., & Wipulanusat, W. (2024). Machine learning prediction of the unconfined compressive strength of controlled low-strength material using fly ash and pond ash. Scientific Reports, 14, 27540. https://doi.org/10.1038/s41598-024-78050-w
7. Gajurel, A., et al. (2021). Machine learning methods to map stabilizer effectiveness in soils. Boise State University Civil Engineering Faculty Publications.
8. Goutham, D. R. (2024). Prediction of unconfined compressive strength of stabilised soil using artificial neural network. Journal of Soft Computing in Civil Engineering.
9. Guo, X., Zhang, J., Zhang, J., & Li, N. (2024). Modeling the influence of lime on the unconfined compressive strength of stabilized soils using machine learning. Environmental Science and Pollution Research, 31, 12456–12472.
10. Hoque, M. I., Hasan, M., Islam, M. S., Houda, M., Abdallah, M., & Sobuz, M. H. R. (2023). Machine learning methods to predict and analyse unconfined compressive strength of stabilised soft soil with polypropylene columns. Cogent Engineering, 10, 2220492. https://doi.org/10.1080/23311916.2023.2220492
11. Indriani, L. (2024). Cement stabilised soil maximum dry density and unconfined compressive strength values can be predicted with artificial neural networks and support-vector regression. BIO Web of Conferences, 63, 06002. https://doi.org/10.1051/bioconf/20246306002
12. Jain, P., Rawat, A., & Singh, A. (2023). Deep learning models for prediction of geotechnical strength parameters. Engineering with Computers, 39(2), 1451–1465.
13. Kumar, A., & Satyam, N. (2017). Influence of plasticity characteristics on performance of cement-stabilized soils. Transportation Geotechnics, 12, 85–95.
14. Ly, H.-B., & Pham, B. T. (2020). Soil unconfined compressive strength prediction using Random Forest (RF) machine learning model. Open Construction & Building Technology Journal, 14, 278–286. https://doi.org/10.2174/1874836802014010278
15. Luo, Z., Xue, X., Xiong, C., Huang, R., Du, J., Li, J., & Liu, K. (2025). Data-driven prediction of unconfined compressive strength in stabilised soils using machine learning. SN Applied Sciences, 7, Article 1122. https://doi.org/10.1007/s42452-025-07766-9
16. Ngo, H. T. T., Pham, T. A., Vu, H. L. T., & Giap, L. V. (2021). Application of artificial intelligence to determine unconfined compressive strength of cement-stabilized soil in Vietnam. Applied Sciences, 11(4), 1949. https://doi.org/10.3390/app11041949
17. Onyelowe, K. C., Moghal, A. A. B., Ebid, A., Rehman, A. U., Hanandeh, S., & Priyan, V. (2024). Estimating the strength of soil stabilised with cement and lime at optimal compaction using ensemble-based machine learning. Scientific Reports, 14, 15308. https://doi.org/10.1038/s41598-024-66295-4
18. Patel, S., Kaur, G., & Sharma, J. (2020). Artificial neural networks for geotechnical strength prediction. Geotechnical and Geological Engineering, 38, 4305–4320.
19. Roy, K., Das, A., & Chakraborty, S. (2024). Effect of plasticity and compaction characteristics on stabilized soil performance. Construction and Building Materials, 367, 130176.
20. Saha, B., & Ghosh, P. (2020). Modelling stabilized soil behaviour using regression and ML approaches. International Journal of Geotechnical Engineering, 14(6), 701–713.

21. Sen, S., & Sarkar, P. (2021). Effect of curing regime on strength gain of cement-treated clays. Materials Today: Proceedings, 46, 1503–1510.
22. Sharma, R., Meena, B. S., & Kaushik, S. (2022). Prediction of geotechnical strength using machine learning models. Innovative Infrastructure Solutions, 7, 113. https://doi.org/10.1007/s41062-022-00869-9
23. Suman, S., Mahamaya, M., & Das, S. K. (2016). Prediction of maximum dry density and unconfined compressive strength of cement stabilised soil using artificial intelligence techniques. International Journal of Geosynthetics and Ground Engineering, 2(1), 11. https://doi.org/10.1007/s40891-016-0051-9
24. Tabarsa, A., Karimpour, A., & Motamedi, M. (2021). Influence of plasticity and curing variables on cement-stabilized soils. Journal of Materials in Civil Engineering, 33(10), 04021234.
25. Teodoru, I.-B., Owusu-Yeboah, Z., Aniculăesi, M., Dascălu, A. V., Hörtkorn, F., Amelio, A., & Lungu, I. (2025). Prediction of unconfined compressive strength in cement-treated soils: A machine learning approach. Applied Sciences, 15(13), 7022. https://doi.org/10.3390/app15137022
26. Tripathi, M., Tomar, S., & Deoliya, R. (2021). Predictive analytics for soil stabilization using ML. International Journal of Pavement Engineering, 22(10), 1214–1224.
27. Wan, W. (2024). Enhancing unconfined compressive strength of stabilised soil with lime and cement prediction through a hybrid machine learning approach utilising Naive Bayes algorithm. Journal of Engineering and Applied Science, 71, 84. https://doi.org/10.1186/s44147-024-00408-8
28. Yang, H. (2025). A physics-informed and SHAP-enhanced modeling framework for predicting UCS of cement-stabilized soils. Computers & Geotechnics, 176, 105487.
29. Yohanna, P. (2025). Assessing the unconfined compressive strength of lime-stabilised soil blended with iron ore tailings using probabilistic and artificial neural networks approach. Materials Today Communications, 44, 104849. https://doi.org/10.1016/j.mtcomm.2024.104849