Research on Engineering Structures & Materials
www.jresm.org
MIM Research Group

Review Article

Forecasting California bearing ratio (CBR) of soil using machine learning algorithms: A review

Nabam Tado[a], Salam Medhajit[b], Dibyendu Pal[*,c]

Department of Civil Engineering, North Eastern Regional Institute of Science and Technology, Nirjuli-791109, Arunachal Pradesh, India

Article Info

Article History:

Received 15 Jan 2025

Accepted 29 Jan 2025

Keywords:

California bearing ratio; Machine learning; Support vector machine; Artificial neural networks; Atterberg limits; Compaction parameters

Abstract

Traditionally California bearing ratio (CBR) is obtained by conducting laboratory testing, which is often time-consuming, laborious, and costly. This delays the design and construction processes of important structures. Recently, several researchers have predicted CBR using ML algorithms. This study focused on understanding the uses of various ML algorithms in the prediction of CBR of treated and natural soils, and other applications. Factors like OMC (30%), MDD (29%), LL (25%), PL (20%), and PI (19%) were mostly used as contributing factors for estimating CBR. ANN, RF, and CNN were the best models for predicting settlement of shallow foundation, bearing capacity of piles and slope stability, and landslide identification, respectively. DNN, GEP, and ELM-CSO were the best models in estimating CBR for granular soil, fine-grained soil, and lateritic soil, respectively, and RFR, AB-DT, LR, and ANN for other types of soils. ANN and BBO-MLP were the best models for expansive clay soil treated with HARHA, and pond ash treated with lime and lime sludge, whereas ANN was for lateritic soil treated with cement and RHA, sand with quartz, feldspar, calcite, corund, amorphous, and clay with pozzolan and lime powder, respectively. The quality and quantity of available training data were fundamental to observing the capacity of models, highlighting the importance of richer, better-labeled datasets.

1. Introduction

California bearing ratio (CBR) is widely used as a key factor to assess the strength of subgrade and determine pavement thickness. CBR determines the relative force required to penetrate a soil sample compared to a standard material. CBR is determined in the laboratory by placing a standard diameter plunger into a sample of compacted soil [1]. This evaluates the relative quality of subgrade soil. These laboratory tests encountered numerous obstacles such as inappropriate compaction, difficulties in preserving proper moisture content, improper alignment of the pistol during penetration, swelling in soil samples after soaking, and errors in recording test results. The traditional method of CBR test requires a substantial amount of time (minimum 4 days per sample per composition), money, and experienced and trained laboratory staff. The field test of CBR comprises driving a piston into the soil mass and soil subgrade using a loading jack. This requires more labor for carrying bulky instruments, making boreholes in the field, and skilled operators to get proper results. Field compaction may not accurately match with laboratory compaction which may lead to erroneous results [2]. This delays the design and construction processes of important structures. Recently, there has been an increasing demand for a faster, more precise, and much less expensive alternative, especially given the increasing complexity of infrastructure projects and new sustainable construction practices. This has attracted

*Corresponding author: dpl@nerist.ac.in / pal_here00@yahoo.co.in
[a]orcid.org/0009-0001-2756-6835; [b]orcid.org/0009-0009-1934-005X; [c]orcid.org/0000-0002-6419-4639
DOI: http://dx.doi.org/10.17515/resm2025-623ml0115rv

considerable interest in exploring machine learning (ML) algorithms as predictive tools for estimating CBR [3]. Over the last decade, numerous ML methods have been investigated to forecast the CBR value. This includes Artificial Neural Network (ANN) [4], Decision Tree (DT) [5], Support Vector Machine (SVM) [6], and ensemble models such as Random Forest (RF) [7], Extreme Gradient Boosting (XGB) [1], and Light Gradient Boosting (LightGBM) [8]. The training of such models mainly takes on data relating to soil characteristics; including particle size distribution (fine content (FC), gravel content (GC), sand content (SC)), Atterberg limits (liquid limit (LL), plasticity index (PI), plastic limit (PL), shrinkage limit(SL)), compaction parameters (maximum dry density (MDD), optimum moisture content (OMC)), natural moisture content (NMC), Specific gravity(G) [9]. The MDD is a key variable, with studies indicating its substantial impact on CBR predictions [10]. While MDD is important, other factors like particle size distribution, including gravel and sand content, were essential for accurate CBR prediction [9, 5]. ML models depend on data quality, quantity, and choices of appropriate algorithms. Issues such as overfitting, model interpretability, and generalizability for diverse soil types were still a challenge [3]. This study is an effort to overview the ML model that functions superior in predicting the CBR value of soil. Efforts have also been made to find the most common soil properties being used to build such a model. This paper elaborates on the studies conducted using ML algorithms to predict various soil properties and narrates the review related to the prediction of CBR from natural soil as well as treated/stabilized soil. It also briefly summarizes the reviews and highlights the conclusions and future direction of this study.

2. Uses of Machine Learning (ML) in Geotechnical Engineering

The application of ML offers numerous advantages with respect to conventional approaches in the field of geotechnical engineering, through enhanced forecasting capabilities and efficiency in data processing. In comparison to regression, the ML model can handle several outputs and responses, whereas regression models can handle a single response at a time [11]. ML techniques, such as ANN, DT, and SVM, excel in modeling complex, non-linear relationships of geotechnical materials like soil, which traditional methods often struggle to capture due to their reliance on simplified assumptions [12-13].

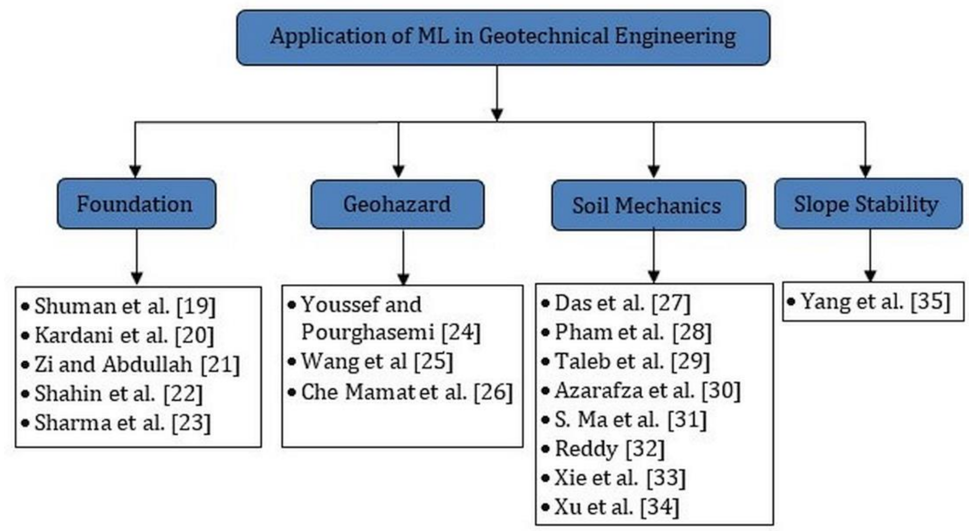

Fig. 1. Structure of the study on application of ML in Geotechnical Engineering

ML can significantly enhance site characterization and scour assessment, providing more reliable estimation and real-time monitoring capabilities [14]. However, challenges persist, including the need for large, high-quality datasets and the chances of overfitting, which can undermine model reliability [15]. Additionally, traditional methods, while laborious, often provide a more straightforward interpretable framework that can be beneficial in certain contexts, highlighting the need for a balanced approach that integrates both methodologies for optimal outcomes in geotechnical applications [13]. The ML techniques, contrary to the traditional statistical and empirical methods that depend on previously available data regarding relationships in data, have very much been adapted to represent the complex behavior of geotechnical engineering materials [16]. Similarly, a novel parameter estimation model was presented by [17] in a study on Backpropagation (BP) neural networks along with geotechnical properties using MATLAB. They chose the expert system method from the BP neural network theory to forecast geotechnical parameters because it had a convenient programming function and strong nonlinear fitting ability and did not need the distribution of the drilling hole. The data source for their study was mainly from geotechnical investigation and engineering geological survey including hydrogeological survey. It was concluded that the forecasting capability of the model satisfied the requirement. A structure of the study showing the flow of the application of ML in geotechnical engineering is presented in Fig. 1.

2.1 Foundations

Finite Element Methods (FEM) are precise but resource-demanding and Boundary Element Methods (BEM) reduce computational load but are incapable of handling heterogeneous and nonlinear conditions [18]. Researchers [19] predicted the settlement of bigger-diameter helical piles in cohesive soil using ML algorithms. They have created a complete database by linking fields and calibrating numerical models. A total of 40 load cases were considered to develop a database of 3600 numerical models for training and validating four different ML algorithms such as DT, RF, ANN, and Adaboost. They evaluated the models using cross-validation techniques and assessed them on a separate dataset. The results confirmed that the DT (R^2 values of 0.92) and RF (R^2 values of 0.96) models had high accuracy. Moreover, ML algorithms also demonstrated great potential in projecting the bearing capacity of piles. Researchers [20] carried out an ML analysis to estimate the bearing capacity of piles. They used cohesionless soil for the study. A dataset of 59 cases was employed to train and validate six different ML models. They were optimized by using the particle swarm optimization (PSO) algorithm. From the results, it was observed that among all, the optimized XGB model was superior with an R^2 value of 0.9615, and the PSO algorithm is efficient in hyperparameter tuning. Researchers [21] investigated the use of ML in predicting the geotechnical axial capacity of reinforced concrete-driven piles. A dataset of 439 piles from six projects in Penang, Malaysia was used. 80% of the data set was used to train the model. The remaining 20% was used for testing the model. The RF model stood to be the superior for forecasting pile geotechnical axial capacity with an R^2 value of 0.962. Researchers [22] examined the ability of ANNs to predict foundation settlement more effectively. The model predictions were also evaluated with three traditional methods. A total of 189 case histories were compiled from existing literature. Input parameters included the footing width, average SPT blow count, footing geometry, applied pressure, and embedment ratio; and the target variable was the settlement. It was found that ANN outperformed the traditional approaches. The ANN was particularly excellent in predicting a wide interval of settlements between 0.6 and 121 mm, while the other traditional methods showed limitations for large settlements. Sensitivity analysis showed SPT blow count, footing width, and applied pressure as the most relevant factors with 33.3%, 23.2%, and 17.7%, respectively. Researchers [23] discussed the usage of ML techniques to forecast the dynamic response of geogrid-reinforced foundation beds for industrial machines using ANN and GP. The consultant ANN and GP models based on field test data were written for parameters such as geogrid depth, shear strain, excitation angle, damping ratio, natural frequency, shear modulus, and operating frequency. They found that both the GP and ANN models can be successfully used to estimate the dynamic response, although GP turned out to be a slightly better model. The most critical parameter was operating frequency.

2.2 Geohazard Assessment

Researchers [24] used ML algorithms for mapping landslide susceptibility prediction in the Abha Basin, Saudi Arabia. They used an inventory map of landslides and twelve landslide-conditioning factors to train and validate seven different ML models. The models were assessed through a comparison of the predicted susceptibility to the actual landslide occurrences. It was found that linear discriminant analysis (LDA) and RF model performed with good accuracy. Researchers [25] gave a new approach that combined both deep learning and ML methods to recognize landslides of natural terrain through integrated geotechnical data. In this research, they used landslide-related data, such as topographic, geological, and rainfall, to build three general-purpose geodatabases. They implemented and compared five different algorithms RF, LR, SVM, CNN, and boosting methods, over the datasets. They have also shown an application of the current method with a case study that was executed at Lantau in Hong Kong. They found that CNN was the best model with an accuracy of 92.5% in the RecLD dataset and that it outperformed other algorithms. Boosting methods were the second most accurate and then RF, LR, and SVM. Similarly, researchers [26] used SVM to predict maximum ground surface settlement (MGS) beneath a road embankment, with embankment height, applied surcharge, and side slope as inputs. Four kernel functions were used to design the SVM network. The RBF kernel function outperforms the SVM (mean absolute relative error (MARE) = 0.048 and root mean square error (RMSE) = 0.007). The SVM model was compared with an ANN for performance evaluation, and it was found that SVM RBF improves MGS prediction accuracy over neural networks.

2.3 Soil Mechanics

Researchers [27] estimated the residual angle of friction of clay soil using characteristics such as index properties of the soil to train and validate different models based on ANN and SVM. From the result, it was seen that the SVM model was superior compared to ANN models in forecasting the residual strength of clay soils. A comparison of four ML algorithms for forecasting the shear strength of soft soil was carried out [28]. The datasets were obtained from 188 samples of plastic clay soil gathered from bridge construction projects in Vietnam. For predicting shear strength, the four ML techniques were Genetic Algorithm Adaptive Network based Fuzzy Inference System (GANFIS), ANN, Particle Swarm Optimization Adaptive Network based Fuzzy Inference System (PANFIS), and SVR. Evaluation of performance was done using different criteria such as RMSE and Pearson coefficient of correlation (R). The PANFIS model showed the highest accuracy in prediction (R=0.601, RMSE=0.038) of the strength of the soil. Researchers [29] used ANNs to estimate the plasticity index, MDD, and OMC of Clayey soil (CH, CL, and MH classes) stabilized with lime. The ANN model was validated on a fresh data set. The sensitivity of each parameter was also carried out. They concluded that the ANN model predicted PI, MDD, and OMC of a lime-stabilized clayey soil with high accuracy. Researchers [30] provided a series of procedures for creating a deep learning predictive model (DNN) predicting the geo-mechanical attributes of samples of marlstone sampled from the South Pars region of southwest Iran. They have predicted unconfined compressive strength (UCS), specific gravity (G), elasticity modulus(E), and an angle of internal friction. The models were used on a dataset obtained by conducting multiple geotechnical tests while evaluating the geotechnical parameters of 120 samples. They suggested that the proposed DNN-based model showed peak accuracy (=0.95), precision (=0.97), and the minimum error rate (RMSE = 0.17, mean absolute error (MAE) = 0.13, and mean square error (MSE) = 0.11). This model was an accurate predictor of geotechnical indices in terms of R^2 (0.925 for E, 0.941 for G, 0.933 for UCS, and 0.954 for the angle of internal friction). Researchers [31] developed an ML algorithm to predict permeability by optimizing the image sample data enhancement process to get a sufficient training dataset. They trained an extreme learning machine neural network to predict the permeability of sandstone and compared its relative error with other established ML methods. They applied the same data collection and forecasting method to granite and bentonite to verify the correctness of the technique. The recommended method accurately predicted the gas permeability of different geomaterials with low error rates (4.1782% for sandstone and granite, 3.2479% for bentonite). RF as an ML algorithm in geotechnical engineering was used by [32] for forecasting the UCS of soil. The key features were soil content, water-holding capacity, relative density, OMC, and plasticity index. The model was assessed using correlation coefficient (R), MAE, and RMSE. They observed that the RF model could efficiently predict UCS values for a wide range of soils with higher accuracy than those of conventional empirical models. Researchers [33] have developed a broad-based model for machine-learned geotechnical subsurface modeling that essentially includes spatial autocorrelations. They applied geotechnical distance fields (GDFs) to six local mapping ML methods: GB, ETs, RF, XGBoost, KNN, and general regression neural network (GRNN). These GDFs allowed the ML models to learn the spatial association among the sampled and unknown locations, and therefore, enhance the accuracy and spatial continuity compared with the conventional XY coordinate fields. GDF-ML is generic and applicable to multi-variable and high-dimensional datasets, as well as incomplete datasets. They concluded that the GDF-ET method results in an accurate and speedy interpretation of soil property profiles with quantified statistical uncertainty.

Researchers [34] used ML to predict key tunnel boring machines (TBM) operational parameters. The ML models employed to develop prediction models were Bayesian ridge regression (BR), nearest neighbors' regression, RF, GTB (Gradient Tree Boosting), and SVM. Two different DNNs, CNN (convolutional neural network), and long short-term memory network (LSTM) were also evaluated. The GTB and LSTM methods provided the best prediction accuracy.

2.4 Slope Stability Analysis

ML Algorithms are also being used to predict and analyze slope stability by finding the relationship between slope stability and influential factors based on available slope data. However, ML model accuracy depends greatly on the appropriate setting of hyperparameters. Different optimization algorithms, like the firefly algorithm (FA) and PSO were used to optimize hyperparameters while improving the accuracy of prediction [35].

Table 1. Summary of the literature on the application of ML in Geotechnical Engineering

Reference	Input parameter used	Model used	R^2/R	Number of data used	Target variable
Researchers [27]	LL, PI, CF, ΔPI	SVM	0.954	137	Residual friction angle
		ANN	0.888		
Researchers [28]	LL, PL, CC, NMC	PANFIS	0.601	188	Shear Strength
		GANFIS	0.569		
		SVR	0.549		
		ANN	0.49		
Researchers [20]	L, A, σ', ϕs, ϕt	PSO-DT	0.9428	59	Bearing Capacity of piles on cohesionless soil.
		PSO-KNN	0.7706		
		PSO-MLP	0.8408		
		PSO-RF	0.9235		
		PSO-SVR	0.8222		
		PSO-XGB	0.9807		
		ANN	N/A		
Researchers [29]	LL, PL, LC	ANN	0.94	280	PI
		ANN	0.94	122	OMC
		ANN	0.94	122	MDD
Researchers [22]	B, q, N, L/B, Df/B	ANN	0.865	N/A	Settlement of shallow foundations on cohesionless soils
Researchers [21]	Dp, As, Ab, Ab, Sp, Ns, Nb, S1, S2	SVM	0.956	439	Geotechnical axial capacity of reinforced concrete-driven pile
		RF	0.962		
		DT	0.959		
		KNN	0.919		

Notation Used: MGS = Maximum ground surface settlement, CC = Clay Content, MLP = Multilayer perceptron, L = Length of the pile, A = Cross-sectional area of the pile, ϕt = Soil shear resistance angle at the tip of the pile, ϕs = Soil shear resistance angle at the shaft of the pile, σ' = Effective stress at the tip of the pile, ΔPI = Deviation from the A-line in casagrande's classification chart = PI-0.73(LL-20), β = Side slope, h = Embankment height, q = Surcharge, RBF = Radial basis function, LC = Lime content, B = Footing width, N = Average SPT blow count, L/B = Footing geometry, Df/B = Footing embedment ratio, CF = Clay fraction, Dp= Pile depth, As = Pile shaft area, Ab = Pile base area, Sp = Pile shape, Ns = Shaft SPT-N, Nb = Base SPT-N, S1 = Soil along pile shaft, S2 = Soil at pile base, Duw=Dry unit weight, CoU=Coefficient of uniformity, CoC=Coefficient of curvature, ClayA=Clay activity, LS=Liner shrinkage, DFS=Differential free swell.

From the above studies, it was observed that the best ML models depended on the specific location and data set and the predicted variable. ML algorithms were used to predict UCS, ground settlement, internal angle of friction, specific gravity, modulus of elasticity, slope stability, geotechnical axial capacity, etc. It was also observed that ML algorithms outperformed traditional statistical techniques. ANN was the best ML algorithm in many cases. RMSE, MAE, R, and MSE were used as a validation parameter for verifying the precision of the ML model. Table 1 presents the summary of the literature on the application of ML in Geotechnical Engineering.

3. Prediction of CBR of Natural Soil

Estimating the CBR of the soil is primarily required in geotechnical and pavement engineering for its crucial role in the design of flexible pavement. Recent advancement in artificial intelligence (AI) and ML offers alternative solutions with much-enhanced accuracy and efficiency of prediction. Initial attempts at predicting CBR were based largely on statistical correlations between soil properties and CBR values. Researchers [36] proved ANNs to be effective in predicting CBR values while identifying important input parameters by sensitivity analysis.

Recent studies have also focused on hybrid and ensemble ML models. Researchers [37] studied the impact of the physical properties of soil on the Un-soaked CBR of soil. A total of 99 soil samples were collected from Nigeria to develop a simplified CBR model using ANN and Least Square Regression (LSR). Both the ANN and LSR models forecasted CBR quite similarly to its laboratory value. The model without the percentage passing through the 200-micron sieve ranked as top. Researchers [38] hybridized ANN with optimization algorithms, such as gradient-based optimization (GBO) and firefly algorithms, and nearly perfect R^2 values were obtained from the training and testing phases.

Researchers [6] compared SVM, RF, and ANN models and concluded that the RMSE and R2 values of RF were maximum. Researchers [10] forecasted the CBR of soils for pavement designs using ML algorithms. A total of 679 data were taken from published literature for this study. ML models such as MLP, KNN regressor, Support Vector Regression (SVR), Random Forest regressor (RFR), and Multilinear regression (MLR) were used. Analysis showed that RFR was the best model followed by KNN, MLP, SVR, and MLR. MDD of soils followed by the percentage of gravel was the supreme dependent parameter of CBR's outcome. Researchers [39] used five ANN models with a database of 521 records based on CBR and eight other index variables from standard laboratory tests of soil to predict CBR. Deep Neural Networks (DNNs) were applied in forecasting the CBR of subgrade soil [40] Various soil characteristics such as grain-size distribution, Atterberg limits, and compaction characteristics were considered as input variables. The results indicate that DNNs give better performance on CBR prediction compared with shallow ANNs and conventional MLR models.

Researchers [41] fused ELM with PSO to develop an accurate prediction model. They also optimized DT with meta-heuristic techniques for superior prediction accuracy ($R^2 = 0.996$). Such recent findings have led to novel hybrid methods compositing various ML algorithms with optimization techniques. Researchers [5] found that the hybrid model (AB-DT) was more reliable than the single model (DT). From the result, it was revealed that the suggested AB-DT model can forecast successfully and precisely the CBR values, and MDD was the most important parameter influencing CBR value. Researchers [42] combined meta-heuristic algorithms with Least Square Support Vector Regression for an overfitting solution. Researchers [43] highlighted the demerit of conventional statistical methods that failed to capture the complex interdependencies among the different soil properties like plasticity, gradation, and compaction-related characteristics. Researchers [44] introduced the GMDH model, which optimized network architecture based on input variables and showed better performance than MLR. Research like [45] strongly shows the ability of ML and AI to substitute empirical methods to make fast and economical CBR estimations. Table 2 represents a brief overview of various literature that employed ML algorithms to forecast the CBR value of natural soil. The ANN was the top performing and the most frequently used ML model for CBR prediction.

Table 2. Summary of the literature in predicting the CBR value of natural soil

References	Input Parameter	Model used	R^2	Number of soil samples
Researchers [6]	NMC, FC, SC, GC, G, LL, PL	SVM	0.7	480
		RF	0.94	
		MLR	0.3	
		ANN	0.43	
		M5 TREE	0.16	
Researchers [36]	MDD, OMC, LL, PI, FC, SC, GC	ANN	0.78	358
Researchers [10]	FC, SC, GC, LL, PL, OMC, MDD	KNNR	0.86	697
		SVR	0.86	
		RFR	0.93	
		MLP	0.9	
		MLR	0.7	
Researchers [46]	FC, SC, GC, LL, PL, OMC, MDD	ELM-CSO	0.996	149
		ELM	0.974	
Researchers [39]	FC, SC, GC, LL, PL, OMC, MDD, PI	ANN	0.96	70
Researchers [40]	FC, SC, GC, LL, PL, OMC, MDD, PI	ANN	0.945	77
Researchers [38]	FC, SC, GC, LL, PL, OMC, MDD, PI, SL	ANN	0.997	100
Researchers [43]	FC, SC, GC, LL, OMC, γd, PI	ANN	0.91	354
		GEP	0.918	
Researchers [5]	FC, SC, GC, LL, PL, PI, OMC, MDD	DT	0.815	214
		AB-DT	0.967	
Researchers [4]	FC, LL, OMC, MDD, PL, PI	DNN M-1	0.836	105
	FC, LL, OMC, MDD, PL, PI	DNN M-2	0.36	175
	FC, OMC, MDD	DNN M-3	0.965	282
Researchers [47]	D10, D60, D30, D50, Cu, Cc	DNN	0.999	90
		MLR	0.957	
Researchers [48]	MDD, G	LR	0.92	34

Hybrid models are the combination of one or more different types of ML models to enhance overall performance like combining the interpreting ability of DT and predicting ability of Neural Networks to achieve better results than a single model could achieve. For more information about

Hybrid Models, readers can refer to the cited article for RF [2], XGB [9], Light GBM [8], etc. From the above study, it was observed that ML and AI are the most accurate and fast techniques to predict the CBR of natural soil with the available dataset. However, from this study, it was not clear whether soil characteristics of different soil types can be combined or not. No study has shown the influence of different soil types on predicting CBR. Therefore, a study may be conducted to verify the impact of soil types on CBR estimation using ML and AI algorithms.

4. Prediction of CBR of Treated Soil

A soil sample mixed with admixtures such as lime, cement, rice husk, etc. to enhance the strength of the soil is known as treated soil. In this section prediction of CBR of treated soil sample data is reviewed. Prediction of CBR of treated soil is challenging as there are different additives and different treatment methods which makes it difficult to generalize prediction for different treatment methods. Developing models for this prediction requires a complex interaction between soil characteristics and additives. Moreover, Datasets of treated soil are hard to find. However, ML (for example ANN model) has the potential to address this complexity by learning the relationship and dependencies between soil properties and additive types from large datasets [49]. ML can also adapt to different soil conditions and additive types, identify patterns, and provide accurate predictions. ML can also optimize additive combinations and quantities for specific soil conditions. The geographical location of soil samples used for the training of the ML model can influence the model's predictive ability [50]. ML models offer enhanced precision in estimating CBR values compared to traditional methods [49-50].

Different researchers have used ML models for the forecasting of CBR of problematic soil such as expansive soil and black cotton soil treated with different admixtures. Researchers [51] investigated the stabilization of expansive soil subgrades for pavement construction by using bagasse ash (BA) as well as geotextile reinforcement. The CBR value of soil was estimated using the ANN model. The treated soil was 6.84% stronger than the natural soil. The MLR analysis tool was 91% reliable in predicting the CBR value. Similarly, researchers [52] used Gaussian process regression (GPR) in forecasting the CBR of expansive soil treated with hydrated lime-activated rice husk ash (HARHA). GPR performed better than the previously constructed models of ANN and GEP for the forecasting of CBR. Sensitivity analysis revealed that HARHA was the most important parameter. Researchers [53] explored the application of three distinct algorithms of an artificial neural network- namely, Levenberg-Marquardt Backpropagation, Bayesian Programming, and Conjugate Gradient algorithms-for assessing strength performance for expansive soils treated with HARHA. The study found that all the ANN algorithms could forecast the values of CBR and UCS of the HARHA-treated soil with good accuracy, and the LMBP algorithm performed better than the others. Researchers [8] validated the predictability of CBR of agricultural and industrial waste mixed expansive soils. This study used Light Gradient Boosting (LGB) as an ML algorithm. The result obtained was, Pearson correlation coefficient (R) = 0.9452, RMSE = 0.3225, and MAE = 0.2522. The important feature in predicting the CBR value was found in order as: ash content, MDD, ash type, OMC, LL, and PL. Researchers [54] presented and compared the predictive capacity of the three ML models namely, Multivariate Adaptive Regression Splines (MARS), RF, and Gradient Boosting Machines (GBM) for the estimation of the CBR value of expansive soil stabilized with sawdust ash, ordinary Portland cement (OPC) and quarry. The RF model indicated better predictive potentiality as compared to MARS and GBM. The influence of stabilizers, incorporated into the model creation process, had considerably improved predictive accuracy. Researchers [55] applied multiple regression analysis to forecast the CBR of cement and waste glass admixture-treated black cotton soil. The model obtained had precisely predicted the CBR with a coefficient of multiple determination, R^2 equal to 0.98 for the standard proctor compactive effort and 0.94 for the treated proctor compactive effort.

Similar studies were also conducted for lateritic soil. Researchers [46] discussed employing an integrated extreme learning machine-cooperation search optimizer approach toward forecasting the CBR of lateritic soils. Two models were developed namely ELM-CSO and ELM. It assesses the ability of models pertaining to minimizing the MAE, MSE, RMSE, or maximizing R and R^2. The values of CBR were better predicted with the model ELM-CSO than with the ELM model. Researchers [56] focused on efficient model development for forecasting the CBR value of cement and RHA (Rice husk ash) mixed lateritic soil. For this study, 1288 samples were used and the data gathered was analyzed using three ANN algorithms. The model's ability was assessed by a set of statistical performance indicators, namely R^2, and RMSE. Based on the results they reported that the ANN model was the best technique for forecasting the lateritic soil's CBR values.

Similar studies were also carried out using clay soil. Researchers [57] did the strength prediction of difficult clayey soils treated with nano-scale combinations of natural source pozzolan (NNP) and lime powder (NL) using MLR, ANN, and FL. In the developed model, the performance criteria suggested that ANN as well as FL techniques predict the value of CBR and PI accurately, but in better terms, ANN is preferred compared to FL. Similarly, researchers [58] used a soft computing approach to estimate Nigerian black clay CBR values, using ANNs and MLPs. The models were trained with the feed forward back propagation algorithm to predict the unsoaked and soaked CBR value of black clay stabilized with cement kiln dust.

Researchers [59] used MR and ANN models to foresee the CBR of pond ash from a thermal power plant. The plant was stabilized with lime and lime sludge. The period of curing was the most crucial parameter affecting the estimation of treated pond ash CBR value. The ANN model was superior among the two. Similarly, in another study, researchers [7] studied the application of two ML techniques, namely RF and M5P model tree, for forecasting the CBR value of pond ash from a thermal power plant. The pond ash was mixed with lime and industrial waste lime sludge. The output of statistical parameters shows that RF performed better than M5P. The curing period was the dominating factor in estimating the CBR value.

Some researchers studied the application of ML algorithms in determining the CBR of fine-grained soil treated using admixtures. Researchers [60] estimated the soaked CBR of fine-grained plastic soils using ML algorithms. A total of 1011 data sets were taken from a highway project for this study. Three ML algorithms, GPR, Kernel Ridge Regression (KRR), and KNN were used in this study. The GPR model developed by the FCM data division method (GPRF) proved to be the superior model for predicting the CBR of fine-grained plastic soils. The K-fold data division method was also proved to be effective in preventing overfitting of the models. It was also found that the geological location of the soil samples used for developing the models can significantly affect the models' predictive ability. Researchers [61] applied soft computing systems in estimating the value of CBR for the fine-grained soil combined with QD and lime, and RHA and lime. Then, the CBR values of the soils were estimated using Simple Linear Regression (SLR), MLR, ANN, and SVM. They found that the ANN model surpassed all the models because it gave the highest R^2 value for both QD-lime and RHA-lime stabilized soils. The optimal content of QD-lime stabilized soil was found to be 40% QD and 4% lime and that for RHA-lime stabilized soil was 12% RHA and 4% lime.

Similar experiments were also conducted with sand. Researchers [62] explored the application of evolutionary polynomial regressions (EPR) and ANN in forecasting the UCS and CBR of micro silica-lime stabilized sulfate silty sand. They used 90 CBR and UCS tests on sulfate silty sand treated with various percentages of micro silica and lime. Similarly, researchers [63] used ANN and MR models to estimate the CBR of Aegean sands. The work had its basis on nine different sands of the Aegean with contrasting soil properties. Laboratory tests on the sands were conducted for an extensive dataset through various parameters like the distribution of particle size, Atterberg limits, and CBR. Among both, the ANN model was superior in predicting CBR values. Researchers [64] used eight ML models, which include RF, Least Median of Squares Regression (LMSR), ANN, Elastic Net Regularization Regression (ENRR), GPR, Lazy K-star (LKS), M-5 Model Trees Alternating Model Trees (AMT) to foresee the CBR of geosynthetic reinforced subgrade. The data used for calibrating and validating the models were collected from earlier studies, covering various soil types. The input parameters for the model's development include soil properties, geosynthetic reinforcement properties, and the position of reinforcement layers.

The results showed that ANN models showed better prediction accuracy in predicting the CBR value. Numerous studies [9, 65] have shown that the accuracy of the ANN model was the highest in predicting CBR value. Researchers [9] found the effects of gradation and compaction characteristics on the CBR of granular materials in subbase and landfill liner construction. Experimental data and six different computational intelligence models ANN, GP, Evolutionary Polynomial Regression (EPR), RF, XGBoost, and Response Surface Methodology (RSM) were used to predict CBR value. Accuracy for ANN was 88%, followed by GP, EPR, and RF with a similar accuracy of 85%, while XGBoost was the least 81%. It was also found that the optimum performance of CBR depends on D_{50} (Particle size at which 50% of the sample is finer) and D_{60} (Particle size at which 60% of the sample is finer). Researchers [66] applied ANNs to foresee the CBR of remolded soils. Three types of soil normally found in the central region of India such as yellow soil, copra soil, and murum soil were used. Two ANN models, specifically GRNN and MLPN with the Levenberg-Marquardt back-propagation algorithm were developed using MATLAB. GRNN model was superior in forecasting the CBR in terms of R^2. Researchers [65] explored the usability of chemically stabilized Coal Gangue (CG), a coal mining byproduct, as a sustainable filling material for earthworks by estimating its engineering properties using ANN and RF models. The Chemicals used were lime and gypsum. Parametric analysis revealed that for obtaining maximum unsoaked and soaked CBR, the optimum contents of gypsum and lime were 1.50% and 4%, respectively, and for UCS, it was 1.50% and 6%, respectively. Both ANN and RF models demonstrated high accuracy, with ANN being slightly superior.

Table 3. Summary of literature predicted the CBR value of treated soil

References	Techniques used	Additives used	Sample size	R^2/R	Soil Properties used	Soil type
Researchers [53]	ANN	hydrated-lime activated rice husk ash	121	More than 0.9	OMC, CA, MDD, LL, PL, PI	Expansive clay Soil
Researchers [66]	MLPN	Remolded the Soil	60	0.97	LL, SC, FC, PI, OMC, MDD	Yellow, Copra and Murum
	GRNN			0.99		
Researchers [69]	BBO-MLP	Lime and Lime sludge	51	0.997	MDD, NMC	pond ash
Researchers [56]	ANN	Cement and Rice husk ash	1288	0.99	LL, PI, MDD, OMC	Lateritic
Researchers [59]	ANN	Lime and Lime sludge	51	More than 0.96	MDD, OMC	pond ash
	MR					
Researchers [63]	ANN	Quartz, Feldspar, Calcite, Corund, Amorphous	N/A	0.978	G, Cu, Cc, MDD, OMC	Sand
	MR			0.812		
Researchers [57]	MLR	Pozzolan and lime powder	120	0.869	G, LL, PL, OMC, MDD	Clay
	ANN			0.989		
	FL			0.975		

Researchers [67] investigated the possibilities of AI methods in estimating the CBR of soil stabilized with alum sludge. They concluded that AI could estimate CBR with high accuracy for MAE values ranging between 0.30 to 0.51, and R^2 values between 0.94 to 0.99. The number of hammers for compaction was the most important parameter, and MDD and mixture were the least important. Researchers [68] utilized the MLR algorithm to forecast the CBR values of the Makkah area soil of Saudi Arabia, based on less complicated tests like the Los Angeles Abrasion test, sieve analysis, OMC, and MDD. The results provided a good relationship between the CBR and sieve analysis parameters, OMC, and MDD, with $R^2 = 0.95$. Table 3 summarizes the literature which predicted the CBR value of treated soil. From the above studies, it can be said that the CBR values obtained from treated or treated soil could easily be predicted using ML algorithms. Many researchers have considered only one or two types of additives to treat the soil to improve its capacity. However, limited research could be seen in predicting the CBR using different additives. The most effective additives of all were not compared to predicting the CBR using ML algorithms. Fig. 2 represents the soil properties used as an input parameter for CBR prediction. It is observed from Fig. 2 that OMC is the most frequently used as input for modeling CBR followed by MDD, LL, PL, and PI.

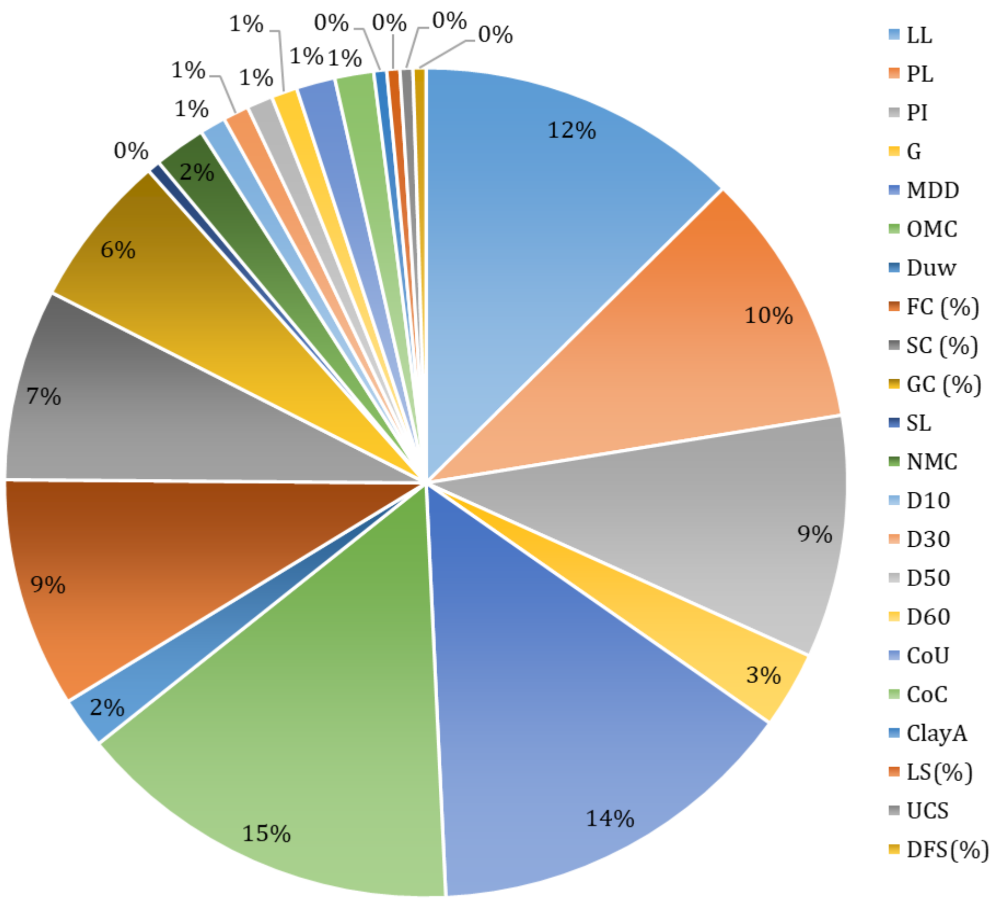

Fig. 2. Most commonly used soil properties as input parameters for CBR prediction

5. Summary and Conclusion

This study focuses on reviewing the literature related to the application of machine learning in forecasting the CBR of natural and treated soils. In addition to that, a few applications of machine learning in geotechnical engineering such as foundation engineering, slope stability, geohazard assessment, and soil mechanics were carried out. The important findings of the reviews are given below.

- ANN (0.865) was found to be the best model for forecasting the settlement of shallow foundations.
- The rank of ML models was RF (0.962), DT (0.959), SVM (0.956), and KNN (0.919) for forecasting load bearing capacity of the driven pile, whereas PSO-XGB (0.9807) was found to be the best model for forecasting bearing capacity of piles on cohesionless soils.

- Though SVM (0.964) was the best in terms of AUC, however, for practical consideration, RF was found to be the best model for forecasting slope stability.
- RF (AUC=0.951) was the best model in forecasting landslide susceptibility, whereas CNN was found to be the best in landslide identification.
- PANFIS, SVM, and ANN were the best models in forecasting shear strength, residual friction angle, and plasticity index, respectively for clay soil. The superior models for predicting PI, MDD, and OMC was ANN.
- In the case of CBR of natural soils, DNN (0.99), GEP (0.918), and ELM-CSO (0.996) were the best models for granular soil, fine-grained soil, and lateritic soil, respectively. Several studies have found the best models as RF (0.94), RFR (0.93), AB-DT (0.967), and LR (0.92). However, multiple studies have found that ANN (0.96/0.945/0.997/0.91/0.78) was the best model, and the most frequently used model was the ANN.
- In the case of CBR of treated soils, ANN (>0.90), BBO-MLP (0.997), ANN (0.99), ANN/MR (>0.96), ANN (0.978), ANN (0.989) was found to be the best models for expansive clay soil with HARHA, pond ash with lime and lime sludge, lateritic soil with cement and RHA, pond ash with lime and lime sludge, sand with Quartz, feldspar, calcite, corund, amorphous, and clay with pozzolan and lime powder, respectively. GRNN (0.99) was found to be the best model for yellow, copra, and murum soil remolding.
- OMC was the most frequently used input for modeling CBR followed by MDD, LL, PL, and PI.

Though ML models are advantageous over other techniques and tedious laboratory experiments, they have also certain limitations. While ML made CBR prediction easy and reliable, ML also faced challenges such as limited data availability, variation of soil properties for different soil, and difficulty in selecting input features. ML models may also suffer from overfitting, underfitting, or poor generalizability when applied to different datasets. ML also faced significant challenges when additives were used as it required complex interaction between soil characteristics and additives.

RF has superior predictive accuracy and reduces overfitting by ensemble learning but it can also be intensive with large datasets, and the complication of the model can hinder interpretability. ANN can learn complex nonlinear relations and is suitable for various types, but the training process is complex and difficult to select hyperparameters. SVM performs well with high dimensional data and has different kernel functions that can adapt to various types of data, but it takes a lot of time to train large datasets, and performance is largely dependent on parameter selection. DT is easy to visualize. It can perform numerical and categorical problems and has a certain level of robustness against missing data, but this model uses a greedy tree which may lead to getting stuck in local optimal solutions. DNN can detect complex features and has better generalizability than simple networks, but it has problems too and needs modification when problems arise since network architecture is initialized at the beginning and it depends on the family of ANN.

Future research on CBR prediction should focus on using deep learning techniques like CNN to detect complicated associations among soil properties and CBR values and developing hybrid models that integrate traditional methods with ML algorithms that could enhance interpretability and applicability. Similarly, ensemble models like RF and GB can be explored to combine the strengths of multiple algorithms and improve prediction accuracy. Moreover, Big data analytics should be utilized to handle and extract insights from large-scale geotechnical datasets, incorporating regional and global soil variability. Advanced engineering features such as automated feature selection, dimensionality reduction, and the use of domain-specific knowledge to create interaction terms, are essential to improve model inputs and understanding the interplay between influencing factors. These directions would pave the way for more robust, accurate, and generalizable predictive models in forecasting CBR.

References

[1] Vardhan KH, Rani KS, Babu ChVN, Satyanarayana G. Prediction of California Bearing Ratio Using Machine Learning for Stabilized Soil Subgrade. International Journal of Research Publication and Reviews, 2023;4(10):621–630. https://ijrpr.com/uploads/V4ISSUE10/IJRPR18113.pdf

[2] Kassa SM, Wubineh BZ. Use of Machine Learning to Predict California Bearing Ratio of Soils. Advances in Civil Engineering, 2023;2023:1-11. https://doi.org/10.1155/2023/8198648

[3] Díaz E, Spagnoli G. Natural Gradient Boosting for Probabilistic Prediction of Soaked CBR Values Using an Explainable Artificial Intelligence Approach. Buildings, 2024;14(2):352. https://doi.org/10.3390/buildings14020352

[4] Yabi CP, Agongbe SW, Koto Tamou BC, Farsangi EN, Alamou E, Gibigaye M. Prediction of CBR by Deep Artificial Neural Networks with Hyperparameter Optimization by Simulated Annealing. Indian Geotechnical Journal, 2024; 54(6):2318–2334. https://doi.org/10.1007/s40098-024-00870-4

[5] Ngo HTT, Bui Q-AT, Vi N Van, Thuy NTB. Application of hybrid modeling to predict California bearing ratio of soil. Vietnam Journal of Earth Sciences, 2024;46(3):399–410. https://doi.org/10.15625/2615-9783/20766

[6] Tunbosun AJ, Ehiorobo JO, Obinna OS, Nwankwo E. Application of Soft Computing Techniques in Modelling of Soaked and Unsoaked California Bearing Ratio. Asian Soil Research Journal, 2022;6(2):32–46. https://doi.org/10.9734/asrj/2022/v6i230128

[7] Suthar M, Aggarwal P. Modeling CBR value using RF and M5P Techniques. Mendel, 2019;25(1):73–78. https://doi.org/10.13164/mendel.2019.1.073

[8] Tran VQ, Do HQ. Prediction of California Bearing Ratio (CBR) of Stabilized Expansive Soils with Agricultural and Industrial Waste Using Light Gradient Boosting Machine. Journal of Science and Transport Technology, 2021;1(1):1–8. https://jstt.vn/index.php/en/article/view/3

[9] Alzara M, Onyelowe KC, Ebid AM, Hanandeh S, Yosri AM, Alshammari TO. Modeling of the Effect of Gradation and Compaction Characteristics on the California Bearing Ratio of Granular Materials for Subbase and Landfill Liner Construction. Sci Rep, 2024;14:23630. https://doi.org/10.1038/s41598-024-74106-z

[10] Bherde V, Mallikarjunappa LK, Baadiga R, Balunaini U. Application of Machine-Learning Algorithms for Predicting California Bearing Ratio of Soil. Journal of Transportation Engineering, Part B: Pavements, 2023;149(4): 1-12. https://doi.org/10.1061/jpeodx.pveng-1290

[11] Mahajan L, Bhagat S. Machine learning approaches for predicting compressive strength of concrete with fly ash admixture. Res. Eng. Struct. Mater., 2023;9(2):431–456. https://doi.org/10.17515/resm2022.534ma0927

[12] Liu H, Su H, Sun L, Dias-da-Costa D. State-of-the-art review on the use of AI-enhanced computational mechanics in geotechnical engineering. Artificial Intelligence Review, 2024;57:196. https://doi.org/10.1007/s10462-024-10836-w

[13] Shao W, Yue W, Zhang Y, Zhou T, Zhang Y, Dang Y, et al. The Application of Machine Learning Techniques in Geotechnical Engineering: A Review and Comparison. Mathematics, 2023;11(18):3976. https://doi.org/10.3390/math11183976

[14] Yousefpour N, Liu Z, Zhao C. Machine Learning Methods for Geotechnical Site Characterization and Scour Assessment. Transp Res Rec, 2024;0:0. https://doi.org/10.1177/03611981241257512

[15] Yaghoubi E, Yaghoubi E, Khamees A, Vakili AH. A systematic review and meta-analysis of artificial neural network, machine learning, deep learning, and ensemble learning approaches in field of geotechnical engineering. Neural Computing and Application, 2024;36:12655–12699. https://doi.org/10.1007/s00521-024-09893-7

[16] Puri N, Prasad HD, Jain A. Prediction of Geotechnical Parameters Using Machine Learning Techniques. Procedia Computer Science, 2018;125:509–517. https://doi.org/10.1016/j.procs.2017.12.066

[17] Cui K, Jing X. Research on prediction model of geotechnical parameters based on BP neural network. Neural Computing and Applications, 2019;31:8205–8215. https://doi.org/10.1007/s00521-018-3902-6

[18] Mourya VK, Pandey G, Patel D, Kumar R. Approaches considering non-linearity in soil-foundation-interaction: A State of the Art Review. Res. Eng. Struct. Mater., 2023;9(3):989–1013. https://doi.org/10.17515/resm2023.646me0117

[19] Shuman NM, Khan MS, Amini F. Efficient machine learning model for settlement prediction of large diameter helical pile in c—Φ soil. AI in Civil Engineering, 2024;3(10):1–26. https://doi.org/10.1007/s43503-024-00028-4

[20] Kardani N, Zhou A, Nazem M, Shen SL. Estimation of Bearing Capacity of Piles in Cohesionless Soil Using Optimised Machine Learning Approaches. Geotechnical and Geological Engineering, 2020;38:2271–2291. https://doi.org/10.1007/s10706-019-01085-8

[21] Xun OZ, Abdullah RA. Predicting Geotechnical Axial Capacity of Reinforced Concrete Driven Pile Using Machine Learning Technique. Malaysian Journal of Civil Engineering, 2023;35(3):11–23. https://doi.org/10.11113/mjce.v35.20544
[22] Shahin MA, Maier HR, Jaksa MB. Predicting Settlement of Shallow Foundations using Neural Networks. Journal of Geotechnical and Geoenvironmental Engineering, 2002;128(9):785–793. https://doi.org/10.1061/(ASCE)1090-0241(2002)128:9(785)
[23] Sharma S, Venkateswarlu H, Hegde A. Application of Machine Learning Techniques for Predicting the Dynamic Response of Geogrid Reinforced Foundation Beds. Geotechnical and Geological Engineering, 2019;37:4845–4864. https://doi.org/10.1007/s10706-019-00945-7
[24] Youssef AM, Pourghasemi HR. Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. Geoscience Frontiers, 2021;12(2):639–655. https://doi.org/10.1016/j.gsf.2020.05.010
[25] Wang H, Zhang L, Yin K, Luo H, Li J. Landslide identification using machine learning. Geoscience Frontiers, 2021;12(1):351–364. https://doi.org/10.1016/j.gsf.2020.02.012
[26] Che Mamat R, Ramli A, Che Omar MBH, Samad AM, Sulaiman SA. Application of machine learning for predicting ground surface settlement beneath road embankments. International Journal of Nonlinear Analysis and Applications, 2021;12(Special Issue):1025–1034. https://doi.org/10.22075/IJNAA.2021.5548
[27] Das SK, Samui P, Khan SZ, Sivakugan N. Machine learning techniques applied to prediction of residual strength of clay. Central European Journal of Geosciences, 2011;3(4):449–461. https://doi.org/10.2478/s13533-011-0043-1
[28] Pham BT, Son LH, Hoang TA, Nguyen DM, Tien Bui D. Prediction of shear strength of soft soil using machine learning methods. Catena, 2018;166:181–191. https://doi.org/10.1016/j.catena.2018.04.004
[29] Bahmed IT, Harichane K, Ghrici M, Boukhatem B, Rebouh R, Gadouri H. Prediction of geotechnical properties of clayey soils stabilised with lime using artificial neural networks (ANNs). International Journal of Geotechnical Engineering, 2017;13(2):191–203. https://doi.org/10.1080/19386362.2017.1329966
[30] Azarafza M, Bonab MH, Derakhshani R. A Deep Learning Method for the Prediction of the Index Mechanical Properties and Strength Parameters of Marlstone. Materials, 2022;15(19):6899. https://doi.org/10.3390/ma15196899
[31] Ma S, Liu J, Lin Y, Wang Z, Zhang T, Jian Z. Prediction of permeability of various geotechnical materials under different temperatures based on physical characteristics and machine learning. Fuel, 2025;379:13109. https://doi.org/10.1016/j.fuel.2024.133109
[32] Reddy YR. Prediction of Ground Properties Using Machine Learning Techniques. International Journal of Current Science (IJCSPUB), 2011;1(1):12-15. https://rjpn.org/ijcspub/papers/IJCSP11A1004.pdf
[33] Xie J, Huang J, Zeng C, Huang S, Burton GJ. A generic framework for geotechnical subsurface modeling with machine learning. Journal of Rock Mechanics and Geotechnical Engineering, 2022;14(5):1366–1379. https://doi.org/10.1016/j.jrmge.2022.08.001
[34] Xu C, Liu X, Wang E, Wang S. Prediction of tunnel boring machine operating parameters using various machine learning algorithms. Tunnelling and Underground Space Technology, 2021;109:103699. https://doi.org/10.1016/j.tust.2020.103699
[35] Yang Y, Zhou W, Jiskani IM, Lu X, Wang Z, Luan B. Slope Stability Prediction Method Based on Intelligent Optimization and Machine Learning Algorithms. Sustainability, 2023;15(2):1169. https://doi.org/10.3390/su15021169
[36] Al-Busultan S, Aswed GK, Almuhanna RRA, Rasheed SE. Application of Artificial Neural Networks in Predicting Subbase CBR Values Using Soil Indices Data. IOP Conference Series: Material Science and Engineering, 2020; 671:012106. https://doi.org/10.1088/1757-899X/671/1/012106
[37] Tunbosun AJ, Ehiorobo JO, Obinna OS, Nwankwo E. Application of Machine Learning Techniques to Estimate Unsoaked California Bearing Ratio in Ekiti Central Senatorial District. Current Journal of Applied Science and Technology, 2021;40(35):30–37. https://doi.org/10.9734/cjast/2021/v40i3531571
[38] Thapa I, Ghani S. Estimation of California bearing ratio for hill highways using advanced hybrid artificial neural network algorithms. Multiscale and Multidisciplinary Modeling, Experiments and Design, 2024;7:1119–1144. https://doi.org/10.1007/s41939-023-00269-3
[39] Ojeda JMP, Calderon BAC, Huatangari LQ, Pintado WR. Determination of the California Bearing Ratio of the Subgrade and Granular Base Using Artificial Neural Networks. International Journal of Engineering and Technology Innovation, 2023;13(3):175–188. https://doi.org/10.46604/ijeti.2023.11053
[40] Othman K, Abdelwahab H. The application of deep neural networks for the prediction of California Bearing Ratio of road subgrade soil. Ain Shams Engineering Journal, 2023;14(7):101988. https://doi.org/10.1016/j.asej.2022.101988

[41] Bardhan A, Samui P, Ghosh K, Gandomi AH, Bhattacharyya S. ELM-based adaptive neuro swarm intelligence techniques for predicting the California bearing ratio of soils in soaked conditions. Applied Soft Computing, 2021;110:107595. https://doi.org/10.1016/j.asoc.2021.107595
[42] Ma C, Shao Q, Yin M. California Bearing Ratio Prediction Based on LSSVR and Multiple Optimizers. Journal of Applied Science and Engineering, 2024;28(5):1095–1105. https://doi.org/10.6180/jase.202505_28(5).0015
[43] Taskiran T. Prediction of California bearing ratio (CBR) of fine grained soils by AI methods. Advances in Engineering Software, 2010;41(6):886–892. https://doi.org/10.1016/j.advengsoft.2010.01.003
[44] Kurnaz TF, Kaya Y. Prediction of the California bearing ratio (CBR) of compacted soils by using GMDH-type neural network. The European Physical Journal Plus, 2019;134:326. https://doi.org/10.1140/epjp/i2019-12692-0
[45] Shukla DK, Murthy YI. California bearing ratio of black cotton soil using soft computing techniques. Asian Journal of Civil Engineering, 2024;25:3961–3972. https://doi.org/10.1007/s42107-024-01023-x
[46] Nagaraju TV, Bahrami A, Prasad CD, Mantena S, Biswal M, Islam MR. Predicting California Bearing Ratio of Lateritic Soils Using Hybrid Machine Learning Technique. Buildings, 2023;13(1):255. https://doi.org/10.3390/buildings13010255
[47] Polo-Mendoza R, Duque J, Masin D. Prediction of California bearing ratio and modified proctor parameters using deep neural networks and multiple linear regression: A case study of granular soils. Case Studies in Construction Materials, 2024;20:e02800. https://doi.org/10.1016/j.cscm.2023.e02800
[48] Umar IH, Salisu IM, Lin H, Hassan JI. Soil characterization, CBR modeling, and spatial variability analysis for road subgrade: a case study of Danchuwa – Jajere Road, Yobe State, Nigeria. Engineering Research Express, 2024;6:035119. https://doi.org/10.1088/2631-8695/ad78a5
[49] Yousefpour N, Medina-Cetina Z, Hernandez-Martinez FG, Al-Tabbaa A. Stiffness and strength of stabilized organic soils—part ii/ii: Parametric analysis and modeling with machine learning. Geosciences, 2021;11(5):218. https://doi.org/10.3390/geosciences11050218
[50] Gowda S, Kunjar V, Gupta A, Kavitha G, Shukla BK, Sihag P. Prediction of the Subgrade Soil California Bearing Ratio Using Machine Learning and Neuro-Fuzzy Inference System Techniques: A Sustainable Approach in Urban Infrastructure Development. Urban Science, 2024;8(1):4. https://doi.org/10.3390/urbansci8010004
[51] Rajakumar C, Rao PKR, Babu GR, Sreenivasulu A. Experimental and numerical prediction of California bearing ratio of expansive soil stabilized by bagasse ash and geotextile reinforcement. IOP Conference Series: Earth and Environmental Science, 2021;796: 012057. https://doi.org/10.1088/1755-1315/796/1/012057
[52] Ahmad M, Al-Zubi MA, Kubińska-Jabcoń E, Majdi A, Al-Mansob RA, Sabri MMS, et al. Predicting California bearing ratio of HARHA-treated expansive soils using Gaussian process regression. Scientific Reports, 2023;13:13593. https://doi.org/10.1038/s41598-023-40903-1
[53] Onyelowe KC, Iqbal M, Jalal FE, Onyia ME, Onuoha IC. Application of 3-algorithm ANN programming to predict the strength performance of hydrated-lime activated rice husk ash treated soil. Multiscale and Multidisciplinary Modeling, Experiments and Design, 2021;4:259–274. https://doi.org/10.1007/s41939-021-00093-7
[54] Ikeagwuani CC. Estimation of modified expansive soil CBR with multivariate adaptive regression splines, random forest and gradient boosting machine. Innovative Infrastructure Solutions, 2021;6:199. https://doi.org/10.1007/s41062-021-00568-z
[55] Ikara IA, Kundiri AM, Mohammed A. Predicting CBR Values of Black Cotton Soil Stabilized with Cement and Waste Glass Admixture Using Regression Model. American Journal of Traffic and Transportation Engineering, 2019;4(1):31–36. https://doi.org/10.11648/j.ajtte.20190401.15
[56] Nnochiri ES, Okokpujie IP, Tartibu LK. Artificial Neural Network Models for Predicting California Bearing Ratio of Lateritic Soil Admixed with Reinforce and Rice Husk Ash. Revue d'Intelligence Artificielle, 2023;37(2):305–313. https://doi.org/10.18280/ria.370208
[57] Al-Swaidani AM, Meziab A, Khwies WT, Al-Bali M, Lala T. Building MLR, ANN and FL models to predict the strength of problematic clayey soil stabilized with a combination of nano lime and nano pozzolan of natural sources for pavement construction. International Journal of Geo-Engineering, 2024;15:2. https://doi.org/10.1186/s40703-023-00201-1
[58] Salahudeen AB, Sadeeq JA. California bearing ratio prediction of modified black clay using artificial neural networks. In: LS and EE, editors. 10th Anniversary Conference, Accra: West Africa Built Environment Research (WABER) Conference; 2019, 268-281. https://doi.org/10.33796/waberconference2019.19
[59] Suthar M, Aggarwal P. Predicting CBR Value of Stabilized Pond Ash with Lime and Lime Sludge Using ANN and MR Models. International Journal of Geosynthetics and Ground Engineering, 2018;4:6. https://doi.org/10.1007/s40891-017-0125-3

[60] Verma G, Kumar B, Kumar C, Ray A, Khandelwal M. Application of KRR, K-NN and GPR Algorithms for Predicting the Soaked CBR of Fine-Grained Plastic Soils. Arabian Journal of Science and Engineering, 2023;48:13901–13927. https://doi.org/10.1007/s13369-023-07962-y

[61] Islam MR, Roy AC. Prediction of California Bearing Ratio of Fine-Grained Soil Stabilized with Admixtures Using Soft Computing Systems. Journal of Civil Engineering, Science and Technology, 2020;11(1):28–44. https://doi.org/10.33736/jcest.2035.2020

[62] Ghorbani A, Hasanzadehshooiili H. Prediction of UCS and CBR of microsilica-lime stabilized sulfate silty sand using ANN and EPR models; application to the deep soil mixing. Soils and Foundations, 2018;58(1):34–49. https://doi.org/10.1016/j.sandf.2017.11.002

[63] Erzin Y, Turkoz D. Use of neural networks for the prediction of the CBR value of some Aegean sands. Neural Computing and Application, 2016;27:1415–1426. https://doi.org/10.1007/s00521-015-1943-7

[64] Raja MNA, Shukla SK, Khan MUA. An intelligent approach for predicting the strength of geosynthetic-reinforced subgrade soil. International Journal of Pavement Engineering, 2022;23(10):3505–3521. https://doi.org/10.1080/10298436.2021.1904237

[65] Amin MN, Iqbal M, Ashfaq M, Salami BA, Khan K, Faraz MI, et al. Prediction of Strength and CBR Characteristics of Chemically Stabilized Coal Gangue: ANN and Random Forest Tree Approach. Materials, 2022;15(12):4330. https://doi.org/10.3390/ma15124330

[66] Kumar SA, Kumar JP, Rajeev J. Application of Machine Learning Techniques to Predict Soaked CBR of Remolded Soils. International Journal of Engineering Research & Technology (IJERT), 2013; 2(6): 3019–3024. https://www.ijert.org/research/application-of-machine-learning-techniques-to-predict-soaked-cbr-of-remolded-soils-IJERTV2IS60881.pdf

[67] Baghbani A, Nguyen MD, Alnedawi A, Milne N, Baumgartl T, Abuel-Naga H. Improving Soil Stability with Alum Sludge: An AI-Enabled Approach for Accurate Prediction of California Bearing Ratio. Applied Sciences, 2023;13(8):4934. https://doi.org/10.3390/app13084934

[68] Alawi M, Rajab M. Prediction of California bearing ratio of subbase layer using multiple linear regression models. Road Materials and Pavement Design, 2013; 14(1):211–219. https://doi.org/10.1080/14680629.2012.757557

[69] Wu J, Li J, Hu W. Applying Biogeography-Based Multi-Layer Perceptron Neural Network to Predict California Bearing Capacity Value of Stabilized Pond Ash with Lime and Lime Sludge. Stavební Obzor - Civil Engineering Journal, 2022;31(2):349–359. https://doi.org/10.14311/cej.2022.02.0026