Kochi Chapter

Indian Geotechnical Conference
IGC 2022
15th – 17th December, 2022, Kochi

Prediction of California Bearing Ratio (CBR) of Soils using AI-based Techniques

Likhith K M[1], Vaishnavi Bherde[2], Ramu Baadiga[3]and Umashankar Balunaini[4*]

[1]Postgraduate Student, Department of Civil Engineering, Indian Institute of Technology, Hyderabad, 502284, India
[2]Doctoral Student, Department of Civil Engineering, Indian Institute of Technology, Hyderabad, 502284, India
[3]Postdoctoral Researcher, Department of Civil Engineering, Indian Institute of Technology, Hyderabad, 502284, India
[4*]Professor, Department of Civil Engineering, Indian Institute of Technology, Hyderabad, 502284, India; e-mail: buma@ce.iith.ac.in (Corresponding author)

Abstract. The California bearing ratio (CBR) is an important input parameter in the design of flexible pavements. CBR is often determined in the laboratory involving a laborious and time-consuming testing procedure. In recent years, artificial intelligence (AI) and machine learning (ML) techniques have gained popularity in geotechnical engineering and can circumvent the laborious process of conducting laboratory testing to determine soil properties. This study presents the application of two AI models, viz., random forest regressor (RFR) and artificial neural network (ANN), to determine CBR based on soil basic and mechanical properties such as gradation, maximum dry density (MDD), optimum moisture content (OMC), liquid limit (LL), and plastic limit (PL). A large dataset of 652 data points was gathered from an extensive literature review consisting of all the basic and mechanical properties of soil along with the CBR value. The findings from the study reveal that the RFR model gave a high prediction performance with the coefficient of determination (R^2) and mean squared error (MSE) equal to 0.92 and 16.2 respectively, whereas the ANN model resulted in the coefficient of correlation (R) and MSE equal to 0.95 and 28 respectively. Furthermore, sensitivity analysis was carried out to evaluate the most influencing soil parameters affecting the CBR. The results show that MDD has the greatest influence, followed by the percentage of fines, whereas PL has the least importance.

Keywords: California bearing ratio, Random Forest regressor, Artificial neural network, Machine learning, and Soil index properties.

1 Introduction

In India, there is a significant growth in highway infrastructure aimed at improving the connectivity between rural and urban areas. The California Bearing Ratio (CBR) is one of the key input parameters in the design of flexible pavements. It indicates the stiffness modulus and shear strength of the pavement layer or subgrade. CBR is an indirect measure of comparing the subgrade material strength to the strength of typical crushed rock expressed as percentage values [1]. The CBR can be obtained through laboratory testing (following IS: 2720 (1987)) or in the field (IS 10042 (1981)). The CBR value of a

remolded specimen prepared by the Proctor compaction test at OMC and MDD is obtained only after four days of soaking and subsequent loading of the soaked specimen [2][3]. As a result, determining CBR in the laboratory is not only laborious but also time-consuming process. The test results may also be erroneous due to sampling interruption and insufficient laboratory testing conditions. This experiment cannot be extrapolated to the full road stretch due to variations in engineering properties as soil samples are collected from certain sites. The CBR testing along the entire project road length attracts high cost and time. Hence newly emerging methods, such as AI/ML techniques, have been reported to predict CBR of soil based on input soil properties. For example, Taskiran [4] used an artificial intelligence technique to established the relation between CBR of fine-grained soils from the southeast part of Anatolia. Similarly, Ikeagwuani [5] used machine learning techniques such as Multivariate adaptive regression spline (MARS), random forest (RF), and gradient boosting (GB) algorithms to predict the CBR of expansive soil subgrade. This study reported the power of AI/ML to develop the robust models which are capable of predicting the geotechnical properties stated with better accuracy.

On the other hand, Yildrin et al. [1] correlated the CBR with the index properties of soil to obtain the correlation equation for CBR based on multilinear regression analysis (MLRA). Over the last decade, several researchers have attempted to relate CBR values to other soil properties. For instance, National Cooperative Highway Research Program (NCHRP) stated the two statistical relation for predicting the CBR value using soil index properties for both fine- and coarse- grained soils [6]. However, the previous research works correlating the CBR with the other soil properties and relations were established by taking a simple regression analysis such as multilinear regression, genetic expression programming (GEP), artificial neuron network (ANN), MARS, RF, and GB, and so on. Most of the above models used limited data from specific region and reported high accuracy for predicted models. However, it is well known that when considerable variations in soil properties exist in the dataset, its accuracy is questionable.

Predicting the appropriate CBR value of soil using AI/ML techniques requires big data and thorough analysis. In this study, extensive data of different soils were collected from a thorough review of the published research paper. In general, the machine learning model is adversely affected by bias, variance, and noise present in the dataset leading to low accuracy and high error. In order to combat this drawback newly emerging ensemble methods have been successfully used so far. However, the present study focuses on one of the ensemble methods based on the bagging technique popularly known as bootstrap aggregation as it is effective in reducing overall variance resulting in better accuracy and lower error. Random forest works on the principle of bootstrap aggregation and shows convincing results in many fields [5]. Hence this study is an attempt to predict CBR based on non-linear approaches using a random forest regressor and compare the results with the artificial neural network. Based on the authors' recent work, these approaches were found to result in good accuracy in predicting the CBR value. The proposed RFR and ANN techniques could assist in predicting CBR in addition to conventional methods.

2 Overview of RFR and ANN

2.1 Random Forest Regressor (RFR)

Random forest is a type of supervised machine learning algorithm. This approach combines a large number of regression trees and is an ensemble-learning algorithm. A regression tree represents a group of conditions or limitations arranged hierarchically and gradually applied from the tree's root to its leaf [27–29]. Fig. 1 shows the structure of an RFR for regression analysis. RFR regression technique predicts the value of a variable by combining the output of several decision tree algorithms [31], [32]. This is when RFR generates N regression trees and averages the results after receiving an input vector (x) containing the values of the various evidentiary features that have been evaluated for a certain training area. Following the growth of N such trees $\{T(x)\}_1^N$, the RFR regression predictor is expressed in Eq. 1

$$\hat{f}_{rf}^{N}(x) = \frac{1}{N}\sum_{n=1}^{N} T(x)` \quad (1)$$

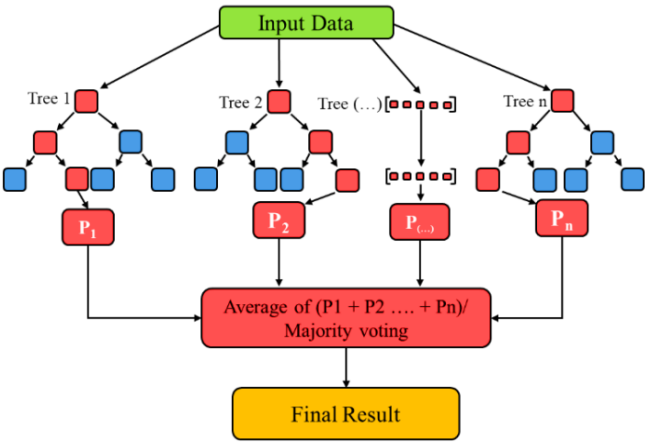

Fig. 1. Structure of random forest regressor

2.2 Artificial Neural Network (ANN)

ANN is a well-known and commonly used machine learning model for mapping any system's nonlinear response. Neural computing necessitates the connection of several simple processing units, known as neurons, into a neural network. These models are made up of a multilayer network of linked basic processing components known as neurons. At its core, The ANN model architecture is composed of three layers: the input layer (number of features), hidden layers with some activation function along with dropouts, and the output layer. Fig. 2 illustrates a simple ANNs architecture. Each layer is made up of a collection of processing components known as nodes (neurons), which communicate with one another via weighted connections.
However, the relationships are unknown and physical meaning is difficult to explain. The backpropagation algorithm [7] is a component of the neural network model that is mostly utilized for data prediction, modelling and used in layered feed-forward ANN. The general ANN computation process can be described as shown in Fig. 2 [14].

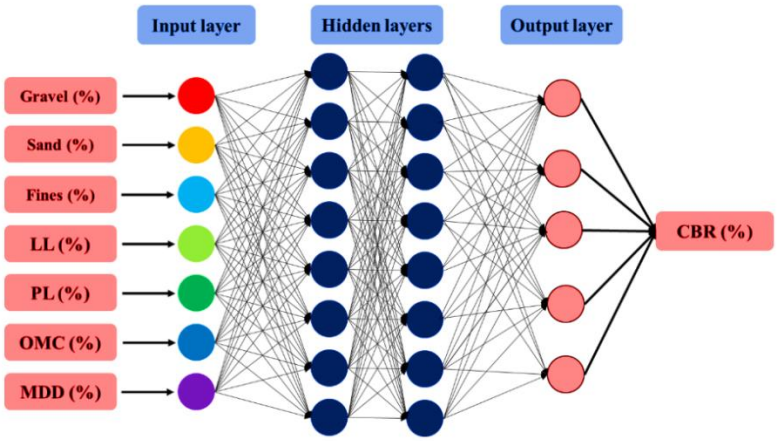

Fig. 2. Architecture of ANN

3 Data Acquisition and Methodology

A systematic literature review was conducted to collect a significant dataset and a total of 652 data points were collected from published research for use in the analysis. Additionally, data on soil gradation characteristics (such as the percentages of gravel, sand, and fines), consistency limits (such as the liquid and plastic limits), compaction characteristics (such as OMC and MDD), and soaked CBR values were gathered [1], [4]-[5], [7]–[31] . The obtained data are summarised in Table 1, along with a statistcal depiction of the data.

Table 1. Statistical representation of the dataset.

Statistical	**Gravel (%)**	**Sand (%)**	**Fines (%)**	**LL (%)**	**PL (%)**	**OMC (%)**	**MDD (g/cc)**	**CBR (%)**
Count	652	652	652	652	652	652	652	652
Mean	14.69	22.55	63.07	43.26	24.42	16.22	1.77	10.81
SD	22.64	15.4	26.78	15.74	8.11	6.31	0.23	16.25
Minimum	0	0	0	15	11	4.5	1.21	0.48
25^{th} percentile	0	11	46.75	33	19	12.15	1.62	2.1
50^{th} percentile	2	21	69	39	22.35	15	1.76	4.61
75^{th} percentile	21.55	31	85	52	28	20	1.94	9.57
Maximum	94	100	100	92.6	62.1	40.2	2.33	81

Note: SD = Standard deviation

Fig. 3 shows the box plot; the reduced section of every box represents the median value of that sample, whereas black dots above the whisker represents the outliers. It represents the data which lies above the range of expected value for that sample, but in this case, all the points were found to be within the general range of that particular feature. Hence, none of the data points were removed in the analysis. While for some features

like MDD, the maximum value was 2.3; for some features like percent fines, the maximum value was 100. Thus, this data type must be normalized before further training in the ML model.

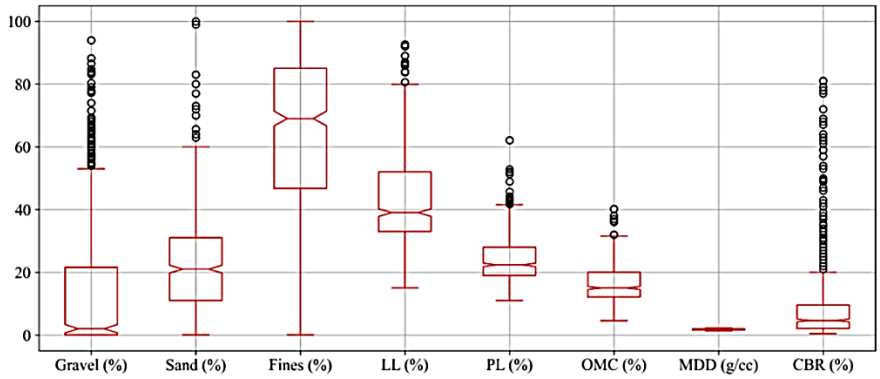

Fig. 3. Box and whisker plot of different soil properties

Furthermore, to analyze the relationship between soil properties, the dataset was represented as a correlation matrix, as shown in Fig. 4. Each feature in the dataset was linearly correlated with the other features without considering the influence of surrounding features. CBR and percentage gravel from the given dataset showed the maximum positive correlation coefficient of 0.77, followed by MDD with a value of 0.66. The maximum negative coefficient was observed for percentage fines.

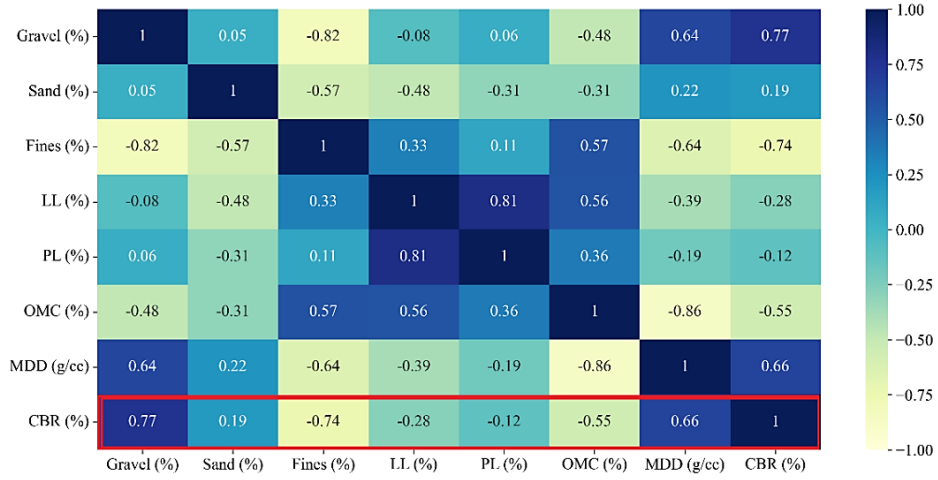

Fig. 4. Correlation matrix obtained between soil properties

4 Result and Discussion

In the current study, out of 652 soil samples, 70% was used for training, and 30% was used for testing. Further, model was trained on the training data and later tested on 30% of the data, and actual and predicted test values were compared. Fig. 5 shows a detailed overview of the process.

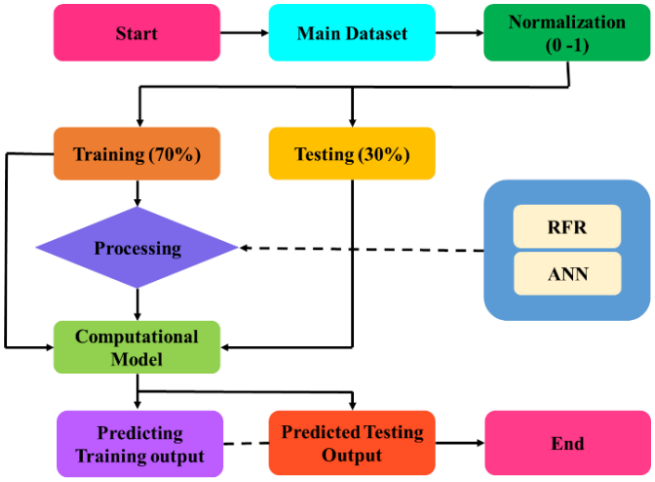

Fig. 5. Step involved in developing model

For judging the performance of the ML model, different accuracy and error matrices were taken into account, viz., coefficient of determination (R^2), coefficient of correlation (R), root mean squared error (RMSE), mean squared error (MSE), and mean absolute error (MAE). Equations (2) to (5) give the expression for R^2, MSE, RMSE, and MAE.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(CBR_i^a - CBR_i^p)^2}{\sum_{i=1}^{N}(CBR_i^a - \overline{CBR}^a)^2} \quad (2)$$

$$MSE = \left[\frac{\sum_{i=1}^{N}(CBR_i^a - CBR_i^p)^2}{N}\right] \quad (3)$$

$$RMSE = \left[\frac{\sum_{i=1}^{N}(CBR_i^a - CBR_i^p)^2}{N}\right]^{1/2} \quad (4)$$

$$MAE = \left[\frac{\sum_{i=1}^{N}|CBR_i^a - CBR_i^p|}{N}\right] \quad (5)$$

where, N represents the testing or training dataset number, CBR_i^p represents the predicted CBR values, CBR_i^a and $\overline{CBR}^a$ denote the actual and mean of the actual CBR values, respectively. R^2 represents the goodness of fit; it can take any value between zero and 1; if this value is equal to 1, we can say that data is linearly correlated. The coefficient of determination with $R^2 = 1$ does not represent zero RMSE or MAE. However, zero RMSE or MAE represents $R^2 = 1$. Literature reveals that, if the generated model gives coefficient of correlation (R) higher than 0.8, there exists a satisfactory correlation between actual and predicted values for that dataset Smith [33]. Sometimes R-value becomes a more significant criterion because some models with a high R^2 value may exhibit high MSE and MAE.

As discussed above, the RFR algorithm was trained on 70% of the data, which accounts for about 457 data points. The optimization method was used to select the best-generated features based on optimization for the number of trees for low MSE total 125 trees were generated, as shown in Fig. 6. The minimum sample split was maintained as two. Fig. 7(a) shows the regression plot for RFR with an R^2 of 0.98 and observed MAE and MSE of 0.82 and 3.09, respectively for the training process and Fig. 7(b) shows the regression plot for RFR with an R^2 of 0.92 and observed MAE and MSE of 2.15 and 16.2, respectively for the testing process.

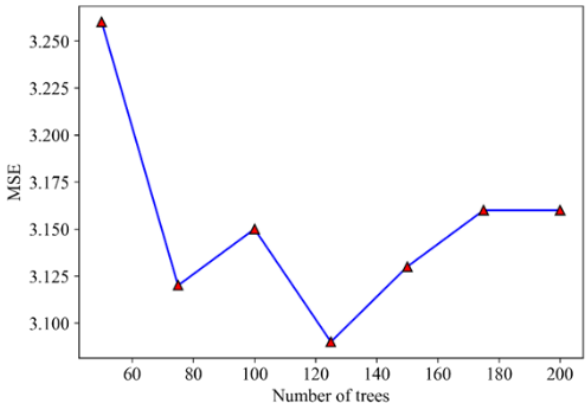

Fig. 6. Optimization for number of trees in RFR

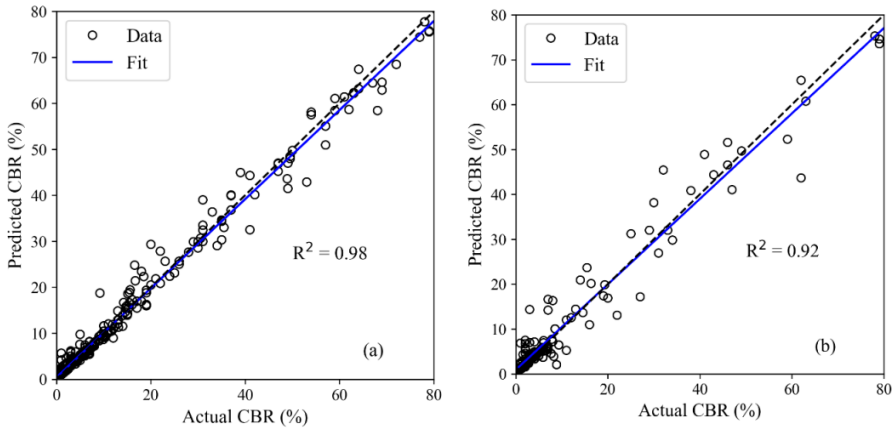

Fig. 7. Comparison of actual and predicted CBR values from RFR (a) for training dataset (b) for testing dataset

ANN toolbox of MATLAB computer-aided software was utilized to assess the performance of neural network on the same previously used training and testing samples. The total number of neurons in the hidden layer and a variety of multilayer networks along with transfer functions such as sigmoid was used to predict optimum CBR of soil. As a

result, they were changed until convergence in the MSE was attained. This investigation used a two-layer feed-forward network along with a sigmoid activation function for hidden neurons and a linear activation function for output neurons. The performance of this network was assessed by MSE and R matrix, and the training algorithm used was Levenberg-Marquardt (LM). The Levenberg-Marquardt (LM) approach is an iterative method for finding the minimum of a multivariate function represented as the sum of squares of non-linear real-valued functions [34, 35]. It has become a typical technique for non-linear least-squares problems [36], and it is widely utilized in many domains. The regression analysis was carried out for the training and testing phases by ANN from the MATLAB toolbox and the results are plotted by using matplotlib in Python as shown in fig. 8. The obtained results showed excellent R values from the ANN Model for both phases.

Similarly, Fig.9 compares the actual and predicted CBR values generated from ML models. Both models performed well on the non-linear dataset (refer to Fig. 9 and Table 2), whereas RFR performs exceptionally well as compared to the ANN model because, this type of model is not biased as predictions are generated from several decision trees. Additionally, the principle of double selection in RFR (results from tree decision in the first stage and random selection of trees in the second stage) gives more accurate results. Moreover, RFR is less computationally expensive and does not require a high GPU to complete its training. In the present study, total 652 data points were considered from different parts of the world, but still, it cannot fully represent the complete complex behavior of soil. As lot of uncertainties are involved in the formation of soil and laboratory experimental testing process, some samples may result into noisy data points. RFR algorithm can perform better for this type of dataset. Moreover, neural network requires a large amount of data and a lot of variety for better performance; hence, with increase in the size of the dataset, ANN model can also be used as an alternative.

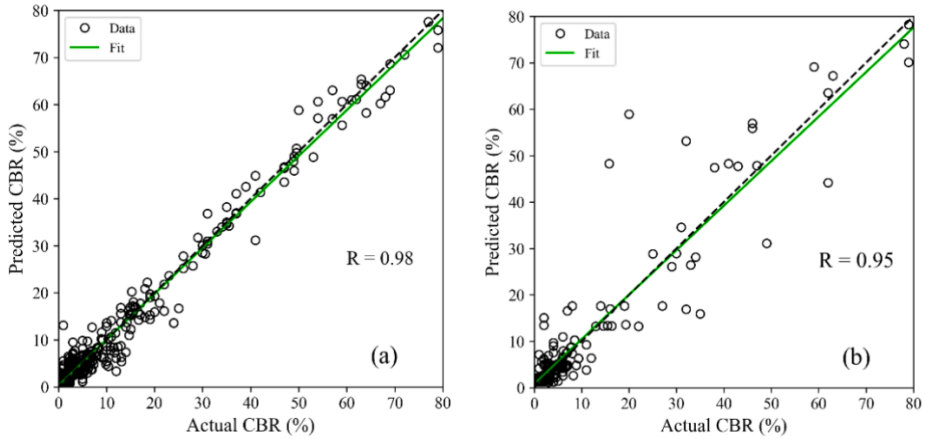

Fig. 8. Comparison of actual and predicted CBR values from ANN using MATLAB toolbox for (a) for training dataset (b) for testing dataset

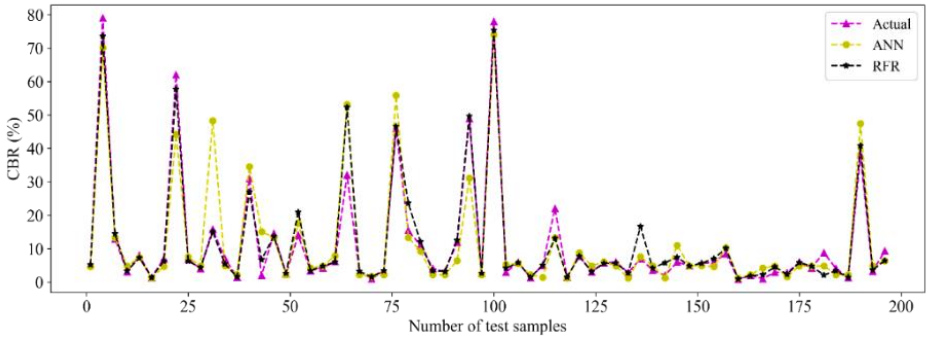

Fig. 9. Comparison of the actual and predicted CBR values from ML models considered in the study

Table 2. Comparison between RFR and ANN model

Models	R^2 (testing)	R (training)	R (testing)	MAE	MSE
Random forest Regressor	0.92	-	-	0.82	3.09
Artificial Neural Network	-	0.98	0.95	-	28.52

Furthermore, sensitivity analysis was carried out to check which feature has more influence on prediction. In the present study, the RFR model was found to perform better than ANN; hence, for further sensitivity analysis, the RFR model was only taken into account. There are different ways to measure feature importance, but this study focuses more on permutation feature importance as it is fast to calculate, most widely used, and easy to understand. Table 3 shows permutation feature importance for CBR prediction using the RFR model.

Table 3. Permutation feature importance for CBR prediction using the RFR model.

Feature	Weight
MDD (g/cc)	0.4609 ± 0.0349
Fines (%)	0.0977 ± 0.0218
LL (%)	0.0758 ± 0.0143
Gravel (%)	0.0723 ± 0.0215
OMC (%)	0.0371 ± 0.0055
Sand (%)	0.0036 ± 0.0021
PL (%)	0.0022 ± 0.0021

Table 3 shows that the top computational interpretation, such as MDD, was the most influential parameter, and the least important parameter was found to be PL. These results from the present study were consistent with the findings of Doshi et al. [37]. The

dry unit weight of deposit is very significant among many engineering properties of soils, like settlement characteristics and bearing capacity. Given that CBR is an indirect measure of soil stiffness and shear modulus, such an effect of dry unit weight on CBR is to be expected.

The first value in each row of Table 3 indicates how much model performance was reduced due to random shuffling ("accuracy" as the performance metric). The precise performance improvement by rearranging columns and repeating the procedure with several shuffles may determine the level of unpredictability. The number after ± shows how the performance is changing from one reshuffle to the next.

5 Summary and Conclusion

The CBR determination in the laboratory is often a laborious and time-consuming procedure. As a result, this study is an attempt for replacing real laboratory tests with prominent AI/ML techniques for calculating CBR of soil on the basis of extensive experimental information available from the literature. In the present study, two different AI models, viz., random forest regressor (RFR) and artificial neural network (ANN), were taken into account for predicting CBR of soil. About 652 soil samples were collected and analysed. Furthermore, several performance metrics were used to evaluate the prediction accuracies of the deployed models. Experimental results show that the RFR model achieved the most accurate prediction with R^2 = 0.98, MSE = 3.09, MAE = 0.82 in training phase; and R^2 = 0.92, MSE = 16.2, MAE = 2.15 in testing phase, whereas ANN model produced R = 0.98 and MSE = 6.63 in training phase and R = 0.95 and MSE = 28.25 in testing phase. R^2 =1 may sometimes not represent ideal model whereas MSE and MAE equal to zero represents perfectly fitted model. Hence, it was concluded that RFR is a better predictor of this type of dataset with less computational cost. However, the neural network performs better on a large dataset with a wide variety. Additionally, sensitivity analysis was carried out to check which features influenced the prediction the most. For this analysis, the RFR model was considered using the permutation importance technique. The results showed that MDD has the greatest influence on CBR prediction, followed by the percentage of fines, whereas PL has the least importance. Based on the efficacy of the models used in the present study, the RFR may be recommended as a potential approach for predicting the CBR of soil and can be efficiently used in many civil engineering applications.

References

1. Yildirim, B., and Gunaydin, O. “Estimation of California bearing ratio by using soft computing systems,” Expert Syst. Appl., vol. 38, no. 5, pp. 6381–6391 (2011).
2. Congress, I. R. IRC: 37-guidelines for the design of flexible pavements (Fourth Revision). In Indian Roads Congress, New Delhi (2018).
3. El-Badawy, S. M., Bayomy, F. M., Santi, M., and Clawson, C. W. “Comparison of Idaho pavement design procedure with AASHTO 1993 and MEPDG methods,” T DI Congr. 2011 Integr. Transp. Dev. a Better Tomorrow - Proc. 1st Congr. Transp. Dev. Inst. ASCE, vol. 41167, no. October, pp. 586–595 (2011).

4. Taskiran, T. “Prediction of California bearing ratio (CBR) of fine-grained soils by AI methods,” Adv. Eng. Softw., vol. 41, no. 6, pp. 886–892 (2010).
5. Ikeagwuani, C. C. “Estimation of modified expansive soil CBR with multivariate adaptive regression splines, random forest and gradient boosting machine,” Innov. Infrastruct. Solut., vol. 6, no. 4 (2021).
6. NEW, O., & PAVEMENT, R. “Guide for Mechanistic-Empirical Design APPENDIX LL: PUNCHOUTS IN CONTINUOUSLY REINFORCED,” no. July, pp. 1–412 (2003).
7. Alam, S. K., Mondal, A., and Shiuly, A. “Prediction of CBR Value of Fine-Grained Soils of Bengal Basin by Genetic Expression Programming, Artificial Neural Network and Krigging Method,” J. Geol. Soc. India, vol. 95, no. 2, pp. 190–196 (2020).
8. Sreelekshmypillai. G., and Vinod, P. “Prediction of CBR value of fine-grained soils at any rational compactive effort,” Int. J. Geotech. Eng., vol. 13, no. 6, pp. 560–565 (2019).
9. Selçuk, L., and Seker, V. “Predicting California bearing ratio of foundation soil using ultrasonic pulse velocity,” Proc. Inst. Civ. Eng. Geotech. Eng., vol. 172, no. 4, pp. 320–330 (2019).
10. Zumrawi, M. “Prediction of CBR from index properties of cohesive soils,” Geotech. Geol. Eng., vol. 2, no. 6, pp. 561–565 (2012).
11. Rakaraddi, P. G., & Gomarsi, V. “Establishing Relationship Between CBR With Different Soil Properties,” Int. J. Res. Eng. Technol., vol. 04, no. 02, pp. 182–188 (2015).
12. Suman, S., Mahamaya, M., and Das, S. K. “Prediction of Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilised Soil Using Artificial Intelligence Techniques,” Int. J. Geosynth. Gr. Eng., vol. 2, no. 2, pp. 1–11 (2016).
13. On, S., Procurement, E., Aik., Terhad, T., Tainia, J. A., and Ehsan, S.D. “Borang Pengesahan Status Tesis,” Program, vol. 16 (2007).
14. Raja, M. N. A., Shukla, S. K., and Khan, M. U. A. “An intelligent approach for predicting the strength of geosynthetic-reinforced subgrade soil,” Int. J. Pavement Eng., vol. 0, no. 0, pp. 1–17.
15. Zumrawi, M. (2012). “Prediction of CBR from index properties of cohesive soils,” Adv. Civ. Eng. Build. Mater., no. November, pp. 561–565 (2021).
16. Johari, A., and Hooshmand Nejad, A. “Prediction of soil-water characteristic curve using gene expression programming,” Iran. J. Sci. Technol. - Trans. Civ. Eng., vol. 39, no. C1, pp. 143–165 (2015).
17. Katte, V. Y., Mfoyet, S, M., Manefouet, B., Wouatong, A, S, L., and Bezeng, L. A. “Correlation of California Bearing Ratio (CBR) Value with Soil Properties of Road Subgrade Soil,” Geotech. Geol. Eng., vol. 37, no. 1, pp. 217–234 (2019).
18. González Farias, I., Araujo, W., and Ruiz, G. “Prediction of California Bearing Ratio from Index Properties of Soils Using Parametric and Non-parametric Models,” Geotech. Geol. Eng., vol. 36, no. 6, pp. 3485–3498 (2018).
19. Zumrawi, M. “Prediction of CBR from index properties of cohesive soils,” Adv. Civ. Eng. Build. Mater. - Sel. Peer Rev. Pap. from 2012 2nd Int. Conf. Civ. Eng. Build. Mater. CEBM 2012, vol. 22, no. Ii, pp. 561–565 (2012).
20. Gratchev, I., Pitawala, S., Gurung, N., and Monteiro, E. “A chart to estimate CBR of plastic soils,” Aust. Geomech. J., vol. 53, no. 1, pp. 1–5 (2018).

21. Ambrose, P., and Rimoy, S. "Prediction of four-days Soaked California Bearing Ratio (CBR) Values from Soil Index Properties," Tanzania J. Eng. Technol., vol. 40, no. 1, pp. 28–38 (2021).
22. Bello, A. A. Regression analysis between properties of subgrade lateritic soil. Leonardo Journal of Sciences, 11(21), 99-108 (2012).
23. Afolabi, O, A., and Afolayan, O, D. "Strength modelling of soil geotechnical properties from index properties," Jordan J. Civ. Eng., vol. 12, no. 4, pp. 619–628 (2018).
24. Satyanarayana Reddy, C. N. V., and Raghava Rao, E. V., "A study on impact of variables on precision of correlation equations of CBR of sandy soils," Int. J. Earth Sci. Eng., vol. 9, no. 3, pp. 1029–1032 (2016).
25. Al-Busultan, S., Aswed, G. K., Almuhanna, R. R. A., and Rasheed, S. E. "Application of Artificial Neural Networks in Predicting Subbase CBR Values Using Soil Indices Data," IOP Conf. Ser. Mater. Sci. Eng., vol. 671, no. 1 (2020).
26. Shirur, N. B., and Hiremath, S. G. "Establishing Relationship between Cbr Value and Physical Properties of Soil," IOSR J. Mech. Civ. Eng., vol. 11, no. 5, pp. 26–30 (2014).
27. Bardhan, A., Gokceoglu, C., Burman, A., Samui, P., and Asteris, P. G. "Efficient computational techniques for predicting the California bearing ratio of soil in soaked conditions," Eng. Geol., vol. 291, no. May, p. 106239 (2021).
28. Abdella, D., Abebe, T., and Quezon, T. "Regression Analysis of Index Properties of Soil as Strength Determinant for California Bearing Ratio (CBR)," Gsj, vol. 5, no. 6, pp. 1–12 (2017).
29. Korde, M., and Yadav, P. R. K. "A Study of Correlation between CBR Value and Physical Properties of Some Soils," Int. J. Emerg. Technol. Adv. Eng., vol. 5, no. 7, pp. 237–239 (2015).
30. Talukdar, D. K. "A Study of Correlation Between California Bearing Ratio (CBR) Value With Other Properties of Soil," Int. J. Emerg. Technol. Adv. Eng., vol. 4, no. 1, pp. 559–562 (2014).
31. Guo, L., Chehata, N., Mallet, C., and Boukir, S. "Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests," ISPRS J. Photogramm. Remote Sens., vol. 66, no. 1, pp. 56–66 (2011).
32. Jin, Z., Shang, J., Zhu, Q., Ling, C., Xie, W., and Qiang, B. "RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 12343 LNCS, pp. 503–515 (2020).
33. Smith, G. N. "Probability and statistics in civil engineering: an introduction," Collins Professional and Technical Books, London (1986).
34. Levenberg, K. "A Method for the Solution of Certain Non-linear Problems in Least Squares," Quarterly of Applied Mathematics, 2(2):164–168 (1944).
35. Marquardt, D . W. "An Algorithm for the Least-Squares Estimation of Nonlinear Parameters," SIAM Journal of Applied Mathematics, 11(2):431–441 (1963).
36. Mittelmann, H. D. The Least Squares Problem. [web page] http://plato.asu.edu/topics/problems/nlolsq.html, Jul. 2004.
37. Doshi, S. N., Mesdary, M. S., & Guirguis, H. R. A statistical study of laboratory CBR for Kuwaiti soils. In Road Engineering Association of Asia and Australasia, Conference, 4th, 1983, Jakarta, Indonesia (Vol. 2).11 (1983).