



Comparison of deep neural network (DNN) and recurrent neural network (RNN) for estimation of california bearing ratio of soils

Souvik Pal¹, May Huu Nguyen², Viet-Dung Do³, Dung Quang Vu^{4,*}

¹Department of Computer Science & Engineering, Sister Nivedita University, New Town, Kolkata, India; souvikpal22@gmail.com

²University of Transport Technology, 54 Trieu Khuc, Thanh Liet, Ha Noi, Viet Nam

³University of Transport Ho Chi Minh City, 2, Vo Oanh, Thanh My Tay, Ho Chi Minh 700000, Vietnam; dungdv@ut.edu.vn

⁴Geotechnical Engineering and Artificial Intelligence research group, University of Transport Technology, 54 Trieu Khuc, Thanh Liet, Hanoi, Vietnam; dungvq@utt.edu.vn

Article info

Type of article:

Original research paper

DOI:

<https://doi.org/10.58845/jstt.utt.2026.en.6.4.300-315>

*Corresponding author:

Email address:

dungvq@utt.edu.vn

Received: 05/06/2026

Received in Revised Form: 06/09/2026

Accepted: 17/09/2026

Abstract: This study applies deep learning models to predict the California Bearing Ratio (CBR) of road subgrade soils. The dataset consists of 214 soil samples, each characterized by ten geotechnical input parameters, including gravel content, coarse sand content, fine sand content, silt-clay content, organic content, liquid limit, plastic limit, plasticity index, optimum water content, and maximum dry density, with CBR as the output variable. Two deep learning models, the Deep Neural Network (DNN) and Recurrent Neural Network (RNN), were developed to model the nonlinear relationship between soil properties and CBR values. The evaluation results indicate that the DNN model outperforms the RNN model on both the training and testing datasets. Specifically, on the testing set, the DNN achieved $R^2 = 0.961$, RMSE = 1.791, and MAE = 1.319, whereas the RNN achieved $R^2 = 0.882$, RMSE = 3.121, and MAE = 2.059. In addition, supplementary analyses using Taylor diagrams, Partial Dependence Plots (PDP), and Individual Conditional Expectation (ICE) demonstrate that the DNN model not only provides higher prediction accuracy but also exhibits greater stability and better interpretability. These results highlight the strong potential of deep learning models, particularly DNN, for supporting rapid and reliable CBR prediction in road subgrade design and evaluation.

Keywords: Deep learning, California Bearing Ratio, Van Don – Mong Cai expressway.

1. Introduction

The strength of subgrade soils is a critical factor governing the performance and service life of pavement structures, in which the California Bearing Ratio (CBR) is widely used as an index to evaluate the load-bearing capacity of subgrade

materials in pavement design [1]. Essentially, the CBR test determines the resistance of soil to penetration under the action of a standard plunger and compares this resistance with that of a standard crushed stone material under identical testing conditions, thereby reflecting the relative

bearing capacity of the subgrade soil [2]. The resulting CBR value is directly used to determine pavement layer thicknesses and compaction requirements, contributing to the stable, long-term performance of the pavement–subgrade system under traffic loading [3]. On this basis, the CBR test has been standardized and widely adopted in accordance with U.S. standards, most notably AASHTO T 193 [4], which specifies detailed procedures for specimen preparation, testing conditions (soaked or unsoaked), penetration rate of the standard plunger, and determination of the characteristic CBR value, thereby ensuring consistency and reliability of test results in subgrade investigation and pavement design.

However, the traditional CBR test is time-consuming and costly, requiring extensive specimen preparation and relatively large-scale testing equipment [5]. For this reason, many previous studies have attempted to estimate CBR values using simpler and more readily available soil indices. Empirical and regression-based models have been developed by relating CBR to parameters such as particle size distribution, Atterberg limits, organic content, optimum water content, and maximum dry density. Although these models can provide preliminary estimates, they suffer from several limitations stemming from simplified assumptions and their inability to adequately capture the complex, nonlinear nature of soil mechanical behavior. In practice, linear or low-order regression relationships often exhibit limited accuracy, particularly for soils with high plasticity and complex structures, owing to heterogeneity and the interactions among different geotechnical properties.

In recent years, machine learning (ML) methods have emerged as effective tools to overcome the aforementioned limitations, owing to their ability to directly learn complex nonlinear relationships from data without relying on rigid mechanical assumptions [6]. ML and artificial intelligence algorithms have been successfully applied to a wide range of geotechnical problems,

such as bearing capacity prediction, slope stability analysis, and soil classification [7]. For CBR prediction, several recent studies have demonstrated the clear potential of ML approaches, showing that models such as Random Forest [8], Gradient Boosting [9], and artificial neural networks [10] can predict CBR with high accuracy based on basic soil properties. For example, studies conducted in India and Vietnam have confirmed that ML can effectively model the relationships between CBR and geotechnical parameters, including particle size distribution, plasticity index, moisture content, and maximum dry density. Beyond conventional ML models, deep learning techniques with multilayer network architectures, such as Deep Neural Networks (DNN) [11] and Recurrent Neural Networks (RNN) [12], offer even greater potential for modeling complex nonlinear relationships [13]. DNNs can automatically extract and learn high-order interactions among input variables, whereas RNNs and LSTM-based models can exploit sequential or dependency patterns in the data when such information is relevant [14].

In Vietnam, the rapid development of transportation infrastructure, with numerous road projects traversing areas characterized by complex geomorphological and geological conditions, from mountainous regions to lowland plains, has created an urgent need for early evaluation of subgrade bearing capacity. Moreover, the tropical monsoon climate, characterized by heavy rainfall and frequent flooding, further complicates the assessment of the CBR of road subgrade soils. Motivated by these practical demands, this study focuses on applying deep learning models to predict CBR, thereby providing effective support for the design and evaluation of road subgrade systems under Vietnam's diverse geological conditions.

Despite the growing use of ML models to predict CBR, limited attention has been paid to the suitability of different deep learning architectures for modeling tabular geotechnical data, particularly

when no explicit temporal or sequential information is available. In addition, the interpretability of deep learning-based CBR prediction models remains insufficiently explored, especially with respect to linking data-driven predictions to underlying geotechnical mechanisms [15, 16].

Therefore, this study aims to examine the use of DNN and RNN models for forecasting the CBR of subgrade soils, employing a dataset collected from the Van Don–Mong Cai expressway in Vietnam. The performance of the models is evaluated via multiple statistical indicators and a Taylor diagram. In addition, Partial Dependence Plots (PDPs) and Individual Conditional Expectation (ICE) analyses are employed to explain the influence of key soil properties on CBR

prediction.

2. Materials and Methods

2.1. Datasets

The dataset used in this study was collected from geotechnical investigations conducted along the Van Don–Mong Cai expressway in Quang Ninh Province, Vietnam. This expressway traverses a region characterized by complex and highly variable geological and geomorphological conditions, including mountainous terrain, hilly areas, and locally developed alluvial and colluvial deposits. The alignment crosses zones with strong lithological heterogeneity, weathered rock formations, and varying soil compositions, which result in significant spatial variability in subgrade properties.

Table 1. Descriptive statistics of variables

Parameter	Symbol	Unit	mean	std	min	max
Gravel content	X1	%	28.322	13.289	0.1	60.82
Coarse sand content	X2	%	18.171	8.818	1.9	39.6
Fine sand content	X3	%	9.035	6.468	2.5	41.5
Silt–clay content	X4	%	44.807	10.447	17.87	88.7
Organic content	X5	%	1.509	0.373	0.12	2.94
Liquid limit	X6	%	39.515	6.173	2.08	48.45
Plastic limit	X7	%	20.318	3.068	1.17	28.49
Plasticity index	X8	%	19.198	4.078	0.91	27.48
Optimum water content	X9	%	14.01	2.619	9.3	21.5
Maximum dry density	X10	(g/cm ³)	1.881	0.118	1.672	2.14
California Bearing Ratio	Y	%	11.804	8.175	3.09	41.26

Soil samples were obtained from borrow pits distributed along the expressway alignment to ensure that the dataset adequately represents the diverse ground conditions encountered along the route. A total of 214 soil samples were collected and tested in the laboratory, covering a wide range of soil types and engineering characteristics (Table 1). Laboratory tests were conducted in accordance with relevant international standards, including particle-size analysis, Atterberg limits, compaction characteristics, and California Bearing Ratio (CBR) testing, in accordance with AASHTO specifications. The diversity of geological and geomorphological settings, together with the

application of standardized laboratory testing procedures, enhances the dataset's representativeness and reliability. This provides a robust basis for developing and validating predictive models for geotechnical applications, particularly in regions characterized by complex terrain and heterogeneous subsurface conditions.

Each soil sample is characterized by ten input variables (X1–X10) representing key geotechnical properties, including gravel content (X1), coarse sand content (X2), fine sand content (X3), silt–clay content (X4), organic content (X5), liquid limit (X6), plastic limit (X7), plasticity index (X8), optimum water content (X9), and maximum

dry density (X10). These variables were selected due to their significant influence on soil strength, compaction behavior, and bearing capacity. The

corresponding California Bearing Ratio (CBR) values obtained from laboratory testing were used as the target variable for model development.

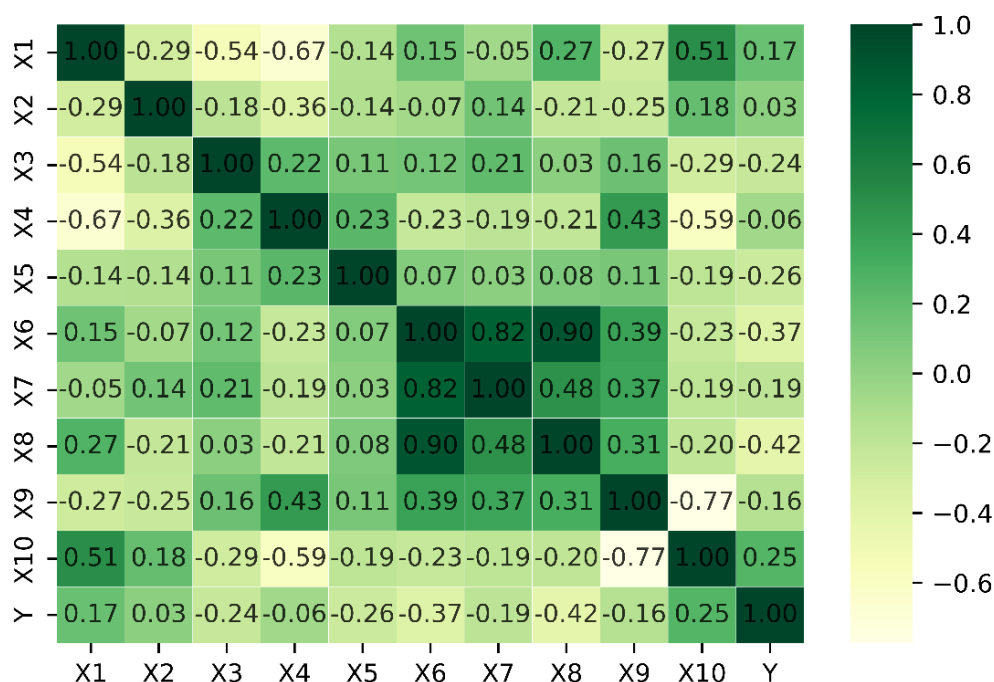


Fig. 1. Correlation matrix analysis of input variables in this study

After data collection, the dataset was examined and subjected to descriptive statistical analysis to evaluate the distribution characteristics and variability of both the input variables and the target variable [17]. Fig. 1 presents the Pearson correlation matrix between the input variables X1–X10 and the target variable Y, illustrating differences in the strength and direction of linear relationships among variable pairs. Overall, the correlation coefficients are mostly low to moderate, indicating weak linear relationships between the input variables and Y. Among the analyzed variables, X8 shows the strongest negative correlation with Y ($r = -0.42$), followed by X6 ($r = -0.37$) and X5 ($r = -0.26$), suggesting that increases in these variables tend to be associated with decreases in Y, although the relationships are not strictly linear. In contrast, X10 exhibits a relatively positive correlation with Y ($r = 0.25$), while the remaining variables display only very weak correlations with the target variable.

2.2. Methods

2.2.1. Deep Neural Network (DNN)

A Deep Neural Network (DNN) is an advanced machine learning architecture inspired by the biological neural system, consisting of multiple hidden layers positioned between the input and output layers [18]. By stacking successive nonlinear transformations, DNNs can learn highly complex, nonlinear mappings between input features and target variables, making them well-suited for regression and prediction tasks involving high-dimensional, intricate datasets.

In a typical DNN architecture, the input layer receives normalized feature vectors, which are then propagated through several fully connected hidden layers [19]. At each hidden layer, the inputs are linearly combined using trainable weights and bias terms, followed by a nonlinear activation function such as ReLU or tanh. This hierarchical learning process enables shallow layers to capture basic data patterns, while deeper layers progressively extract more abstract and higher-level representations of feature interactions. Model parameters are optimized via backpropagation in conjunction with gradient-based optimization

algorithms, such as stochastic gradient descent or Adam, to minimize a predefined loss function, typically the mean squared error in regression problems [20]. Compared with shallow neural networks, DNNs exhibit stronger representational capacity and often achieve superior predictive performance when sufficient training data are available.

2.2.2. Recurrent Neural Network (RNN)

A Recurrent Neural Network (RNN) is a class of neural networks specifically designed to process sequential data by incorporating feedback connections within its architecture [21]. Unlike feedforward neural networks, RNNs maintain an internal hidden state that functions as a memory, allowing information from previous inputs to influence the network's current output. This property enables RNNs to model dependencies across ordered data sequences.

In an RNN, the hidden state at each step is computed as a function of the current input and the hidden state from the previous step [22]. This recursive structure allows the network to capture temporal patterns and sequential dependencies within the data. For regression and forecasting tasks, the final hidden state or a sequence of hidden states is typically mapped to an output layer to generate predictions. In this study, a standard RNN was adopted as a baseline recurrent model to examine whether a simple sequential learning structure could capture dependencies among the input geotechnical variables for CBR prediction. Since the dataset consists of tabular soil properties rather than long temporal sequences, the standard RNN was considered appropriate for the comparative scope of this study [23].

2.2.3. Validation indicators

The predictive performance of the model was examined using three quantitative measures, namely R^2 , MAE, and RMSE, which jointly describe both the explanatory capability of the model and the characteristics of its prediction errors [24]. The combined use of these indicators enables a comprehensive assessment of model

effectiveness from multiple perspectives.

$$R^2 = 1 - \frac{\sum_{i=1}^n (h_i - t_i)^2}{\sum_{i=1}^n (h_i - \bar{h})^2} \quad (1)$$

$$MAE = \frac{\sum_{i=1}^n |h_i - t_i|}{n} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (h_i - t_i)^2}{n}} \quad (3)$$

where h_i and t_i denote the observed and predicted values for the i -th sample, respectively, $\bar{h}$ is the mean of the observed values, and n represents the total number of samples.

The R^2 index reflects the extent to which the model reproduces the overall trend of the observed data, with higher values indicating a closer correspondence between the predicted and observed patterns [25]. In contrast, RMSE places greater emphasis on large deviations and is therefore informative for identifying the severity of substantial prediction errors, whereas MAE provides a linear estimate of the typical prediction error and is less affected by extreme values [26]. Overall, a model can be considered reliable when it achieves a high R^2 together with low MAE and RMSE values. Employing MAE and RMSE simultaneously balances the evaluation of average and extreme errors, thereby enhancing the robustness of the performance assessment.

2.2.4. Taylor Diagram analysis

The Taylor diagram was employed to provide a comprehensive evaluation of model performance by simultaneously illustrating three key statistical indicators: the correlation coefficient (R), standard deviation (σ), and centered root mean square error (RMSE) [27]. In the diagram, the reference observation is shown along the horizontal axis, with a correlation coefficient of 1 and its corresponding standard deviation. Each model is depicted as a point whose radial distance indicates the standard deviation, angular position represents the correlation coefficient, and the distance from the reference point reflects the magnitude of the centered RMSE. Models located closer to the observation point with higher correlation

coefficients and similar standard deviations demonstrate superior predictive performance. Accordingly, the Taylor diagram offers an efficient and intuitive tool for comparing multiple predictive models in a unified framework [28].

2.2.5. Model interpretability using PDP and ICE

In this study, Partial Dependence Plots (PDP) and Individual Conditional Expectation (ICE) plots are employed to clarify the internal behavior of the model and to examine its response to variations in input variables [29]. These techniques are model-agnostic interpretability methods that enable a direct assessment of the magnitude and direction of influence of individual input features on the predicted output, while accounting for nonlinear relationships and latent interactions learned during model training. PDP is used to characterize the average effect of each input variable on the model prediction by marginalizing over the distribution of

the remaining features, thereby providing a global view of prediction trends, nonlinear patterns, and potential transition points in model behavior [30]. In contrast, ICE analyzes prediction responses at the level of individual observations, highlighting heterogeneity across samples and facilitating the identification of localized interactions that may be obscured in PDP representations by averaging [31].

3. Methodological flowchart

CBR prediction was conducted using an artificial intelligence approach formulated as a regression problem, in which deep learning models were trained to capture the nonlinear relationships between input variables and the CBR value [32]. The entire model development and evaluation workflow was implemented in PyCharm 2023.2.5 (see Fig. 2) and consisted of four main steps, as follows:

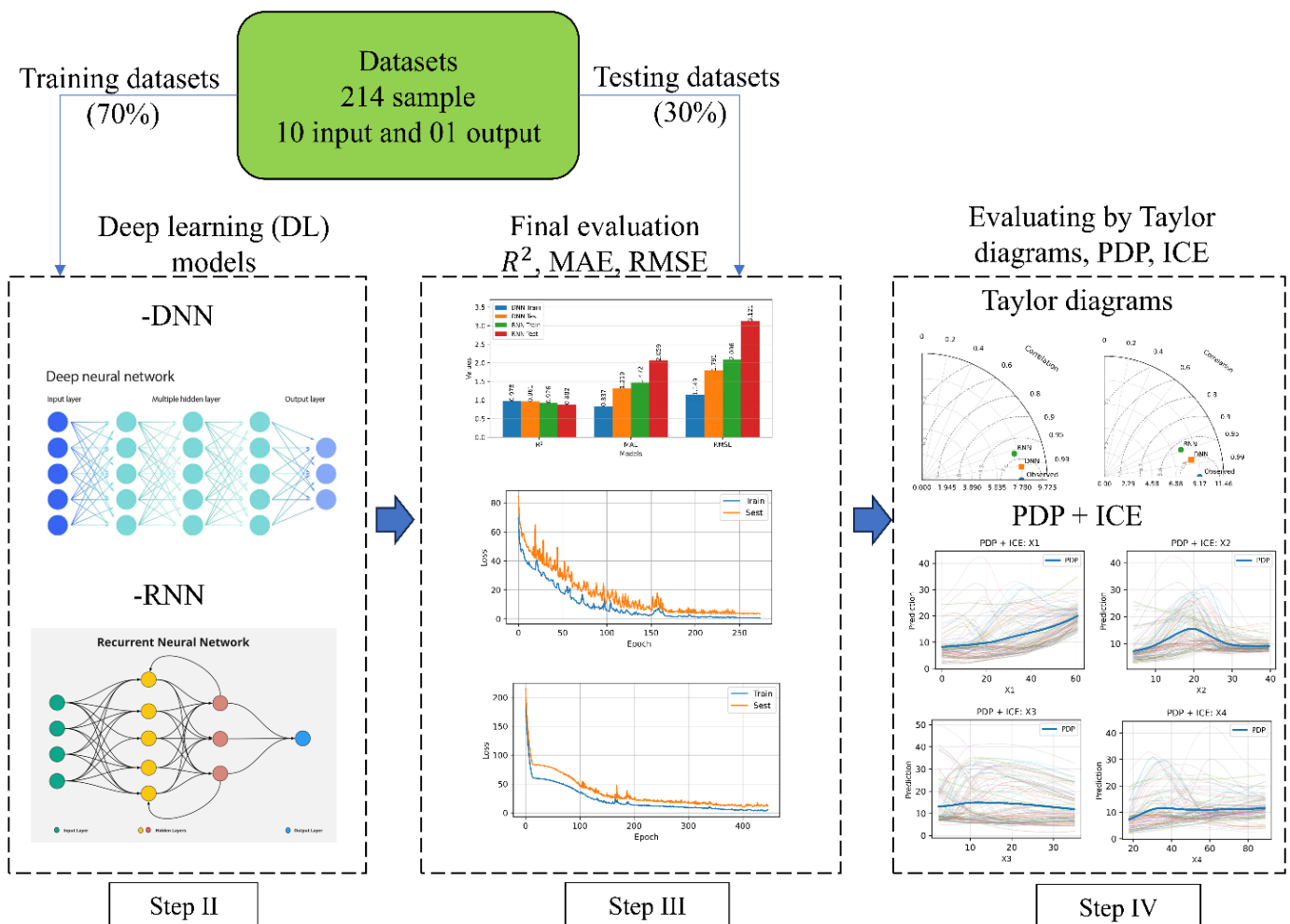


Fig. 2. Methodology flowchart of this study

(1) Data collection and preparation: The input dataset comprised ten variables (X1–X10) representing key geotechnical soil properties, including gravel content (X1), coarse sand content (X2), fine sand content (X3), silt–clay content (X4), organic content (X5), liquid limit (X6), plastic limit (X7), plasticity index (X8), optimum water content (X9), and maximum dry density (X10). The model output was the CBR value obtained from laboratory testing.

(2) Data preprocessing and partitioning: All input variables were normalized using Min–Max scaling to ensure consistent magnitudes [33]. The dataset was then divided into a training set (70%) and a testing set (30%) for model development and performance evaluation [34].

(3) Deep learning model construction and training: Two deep learning models, namely, DNN and RNN, were developed and trained on the training dataset to predict the CBR value. During the training process, regularization techniques were employed to mitigate overfitting, including dropout layers, early stopping based on the validation loss, and appropriate weight optimization strategies.

(4) Model evaluation and interpretation: Model performance was quantitatively assessed using the mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2). In addition, Taylor diagrams were employed to provide an integrated evaluation of model performance, while Partial Dependence Plots (PDP) and Individual Conditional Expectation (ICE) analyses were used to interpret the influence of input variables on the predicted CBR values.

4. Results and discussion

The model development process considered the main hyperparameters and training settings of both deep learning models. For the DNN model, three fully connected hidden layers were used, containing 512, 128, and 64 neurons, respectively, with ReLU activation functions. For the RNN model, a SimpleRNN layer with 64 hidden units and tanh activation was adopted, followed by a

dense layer with 32 neurons. Both models used a single-neuron output layer to predict the CBR value. The Adam optimizer was employed for model training, with mean squared error (MSE) used as the loss function and mean absolute error (MAE) used as the evaluation metric. Early stopping based on validation loss was applied to both models to reduce the risk of overfitting.

The performance evaluation results indicate that the DNN model consistently outperforms the RNN model across all evaluation metrics, including the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE), on both the training and testing datasets (Fig. 3). The DNN model achieves a high R^2 value of 0.978 on the training set and maintains strong generalization on the testing set with an R^2 of 0.961. In contrast, the RNN model yields lower R^2 values, particularly on the testing set ($R^2 = 0.882$), indicating inferior predictive performance. Furthermore, the DNN model exhibits substantially lower MAE and RMSE values on the test set (MAE = 1.319, RMSE = 1.791) than the RNN model (MAE = 2.059, RMSE = 3.121). These results demonstrate that the DNN not only provides higher prediction accuracy but also better error stability.

The scatter plots comparing predicted and actual values show that the DNN model performs better than the RNN model on both the training and testing datasets (Fig. 4). On the training dataset, the DNN predictions are closely aligned with the regression line, with a high coefficient of determination ($R^2 = 0.978$), indicating that the model can effectively learn the relationship between the input variables and the output. On the test dataset, the DNN model shows strong agreement with the observed values, achieving an R^2 of 0.961. In contrast, the RNN model shows lower performance. On the training dataset, the RNN model achieves an R^2 of 0.926, which is lower than that of the DNN model, and the predicted values are more scattered around the ideal line. On the testing dataset, the R^2 value further decreases to 0.882, with larger deviations, especially in the

medium and high value ranges. These results suggest that the RNN model has weaker generalization and is less suitable for the dataset under study.

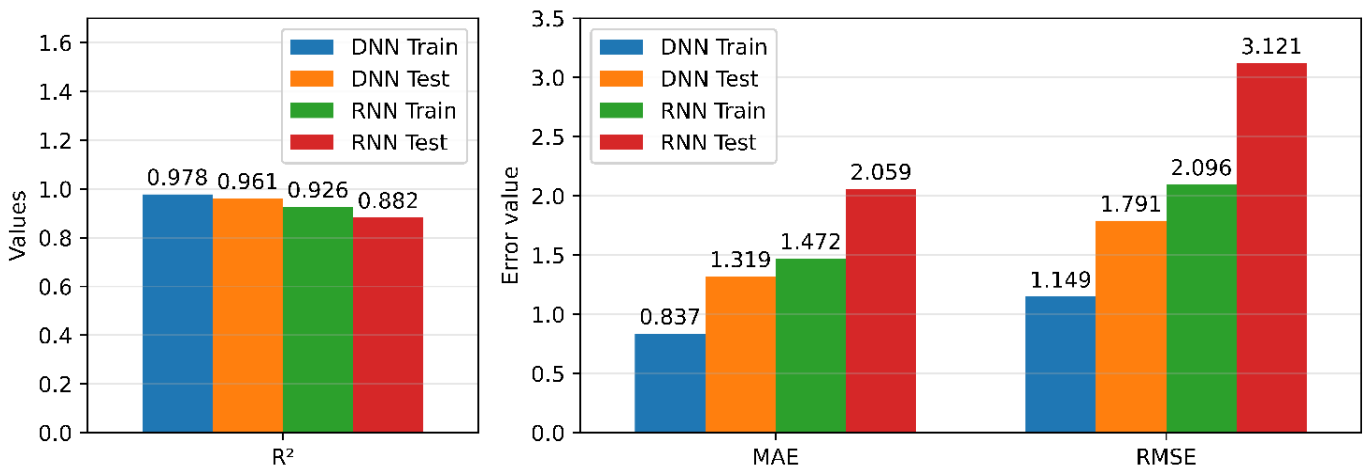


Fig. 3. Comparison of the models' performance results

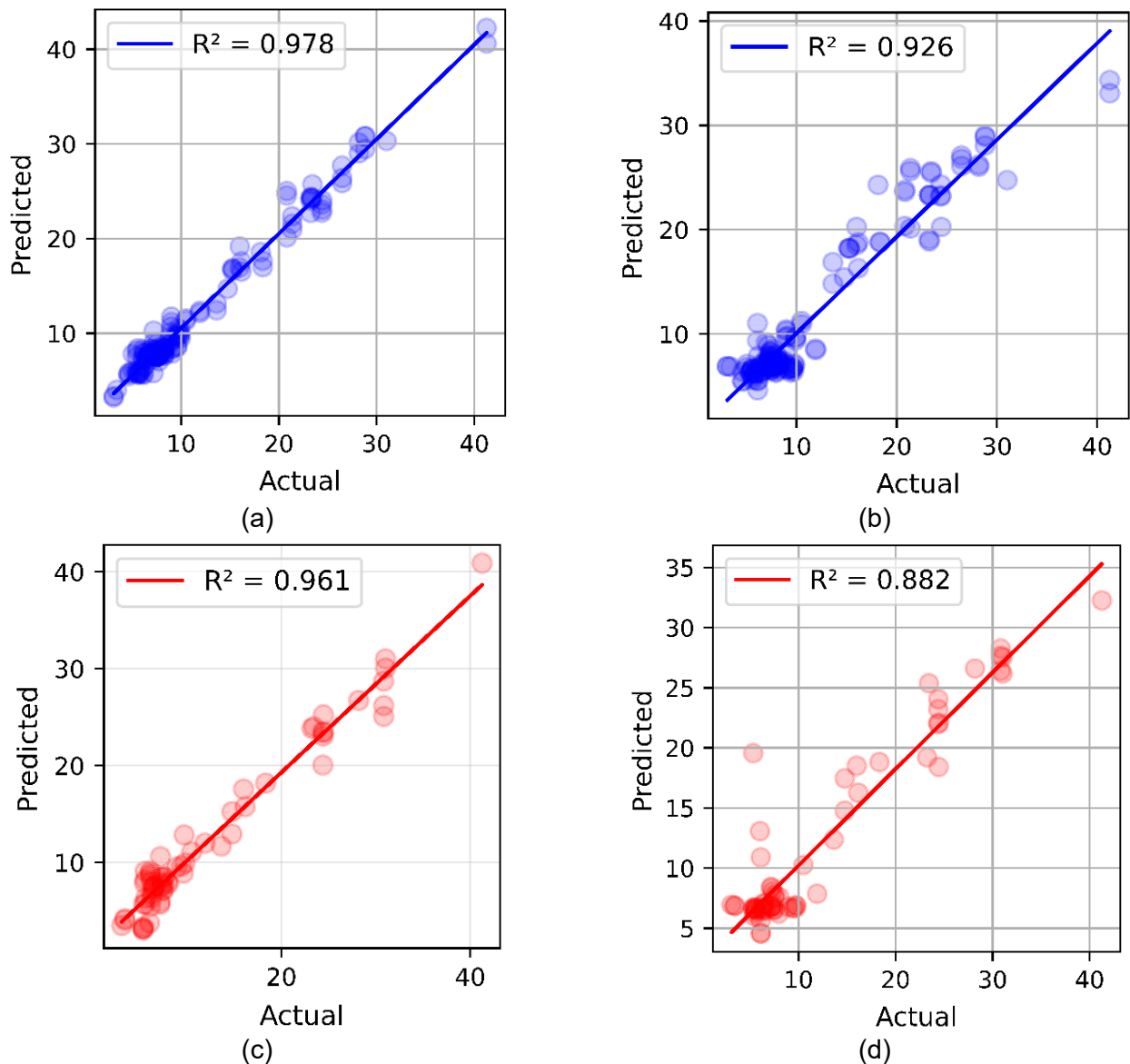


Fig. 4. Scatter plots of actual and predicted CBR values for the training: (a) DNN, (b) RNN, and testing: (c) DNN, (d) RNN

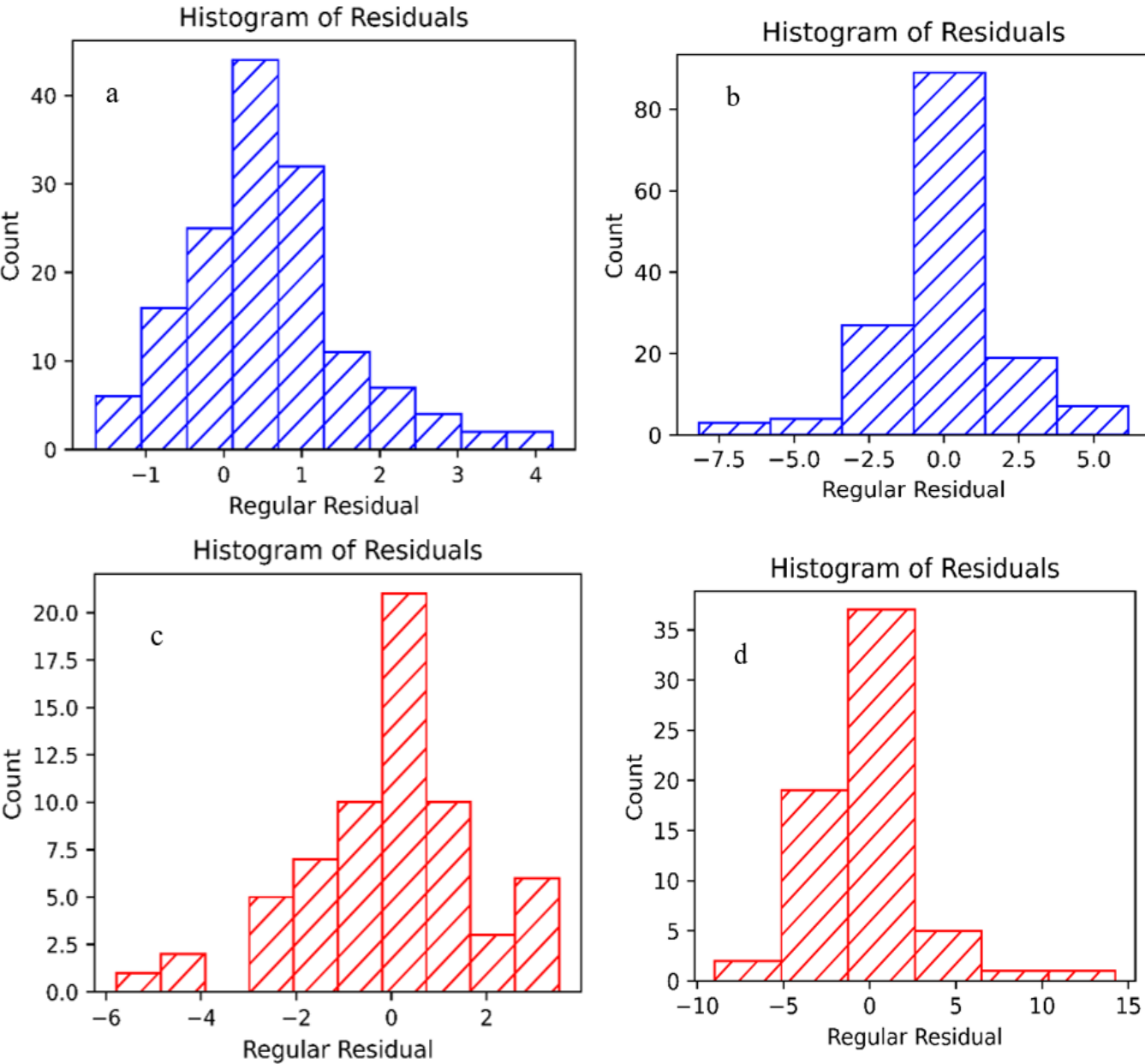


Fig. 5. Histograms of residual values for the training: (a) DNN, (b) RNN, and testing: (c) DNN, (d) RNN

The residual histograms reveal clear differences in error dispersion between the DNN and RNN models on both the training and testing datasets (Fig. 5). On the training dataset, the residuals of the DNN model are mainly concentrated around zero, with most errors falling within the approximate range of -1.5 to 4.0 , indicating small and stable prediction errors, whereas the RNN model exhibits a wider residual distribution, with errors spreading roughly from -8 to 6 , reflecting greater variability and lower stability. On the testing dataset, the DNN model maintains a relatively concentrated residual distribution, with

most residuals lying within the range of -6 to 4 , suggesting a reasonable generalization ability, while the RNN model shows a much broader residual distribution, with extreme residual values extending approximately from -10 to 15 , indicating larger prediction errors and higher sensitivity to medium and high-value samples, and thus a more limited generalization performance.

Fig. 6 illustrates the convergence behavior of the loss function on the training and testing datasets for the two deep learning models, DNN and RNN [35]. The DNN model shows a rapid decrease in loss from early epochs and maintains

a stable convergence throughout training. The small, relatively consistent gap between the training and test loss curves indicates that the model effectively captures the underlying patterns in the data without overfitting, reflecting strong generalization capability. In contrast, the RNN model shows a slower convergence rate, with the

testing loss remaining noticeably higher than the training loss and exhibiting pronounced fluctuations during the early and intermediate training stages. Overall, the comparison shows that the DNN model converges faster, achieves a lower final loss, and demonstrates greater training stability than the RNN model.

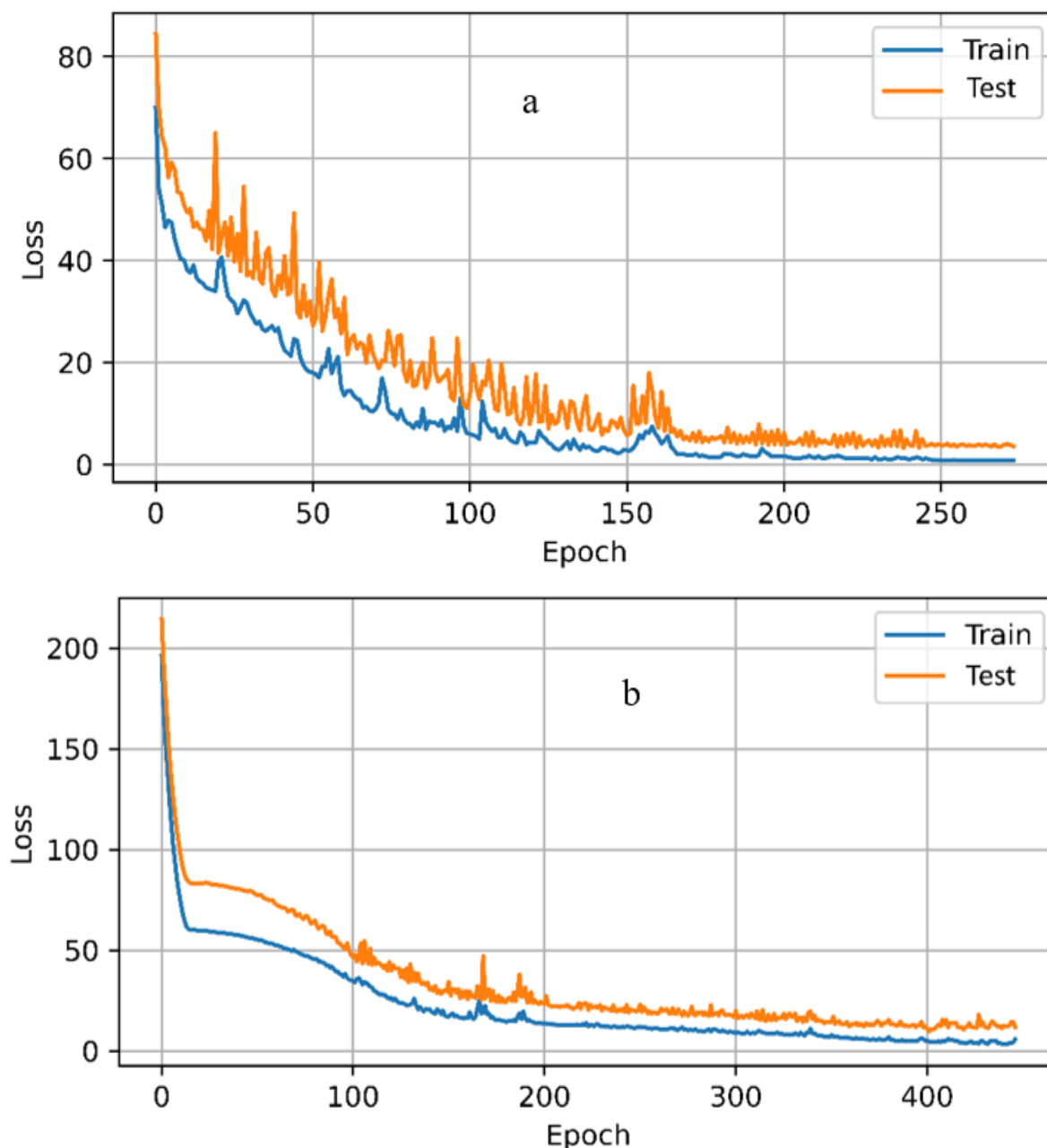


Fig. 6. Training and testing loss curves for the (a) DNN model and (b) RNN model

The Taylor diagrams reveal apparent differences in predictive performance between the DNN and RNN models on both the training and test datasets, as shown in Fig. 7. In these diagrams, the observed values serve as a reference point. In

contrast, the position of each model simultaneously represents three key characteristics, including the correlation coefficient, standard deviation, and centered RMSE, thereby enabling an integrated assessment of model performance (Fig. 7). For the

DNN model, the representative point is located much closer to the observed reference and lies in a region with a very high correlation coefficient (approximately 0.99). In contrast, its standard deviation is close to that of the observed data. This indicates that the DNN model not only captures the overall trend of the data but also effectively reproduces the magnitude of variability, resulting in small, stable prediction errors. In contrast, the RNN

model is farther from the observed point, with a lower correlation coefficient and a more pronounced deviation in standard deviation, particularly evident on the test dataset. The greater distance between the RNN model and the observed reference indicates a higher RMSE, suggesting reduced prediction accuracy and weaker stability when the model is applied to unseen data.

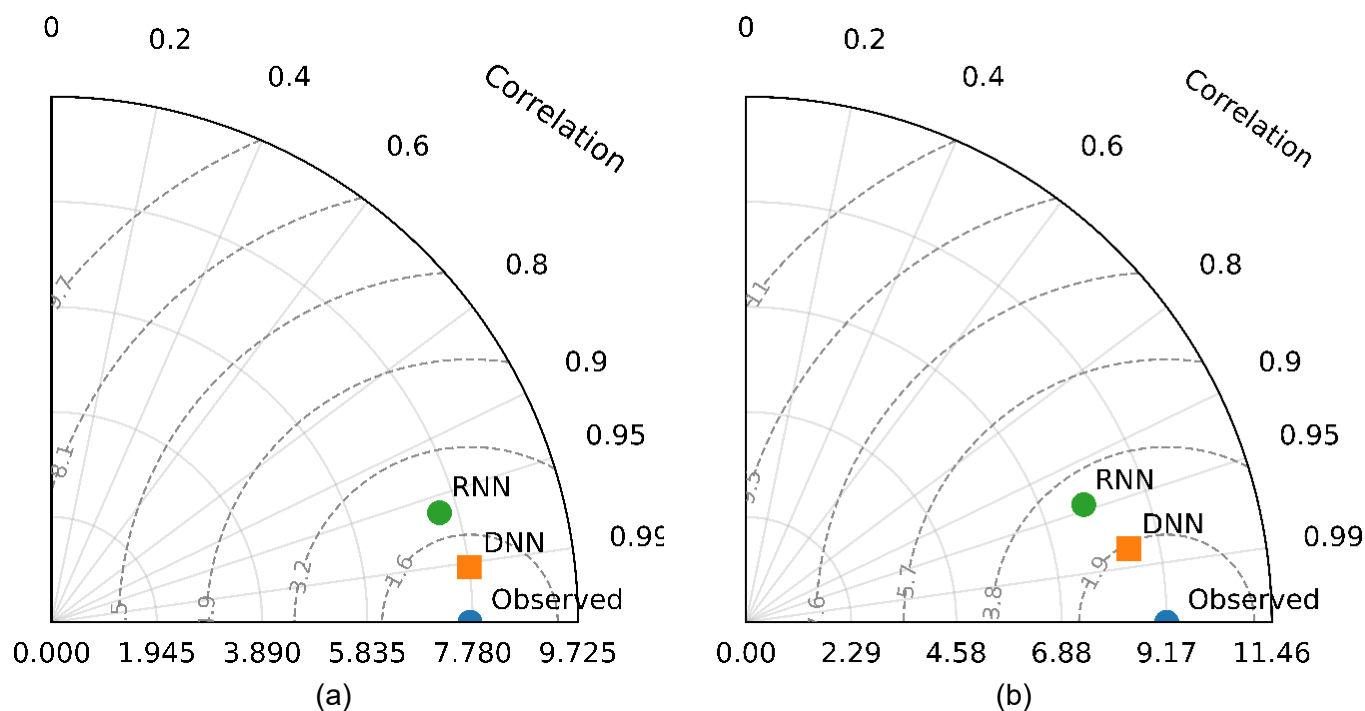


Fig. 7. Taylor diagrams for model performance evaluation: (a) training dataset and (b) testing dataset

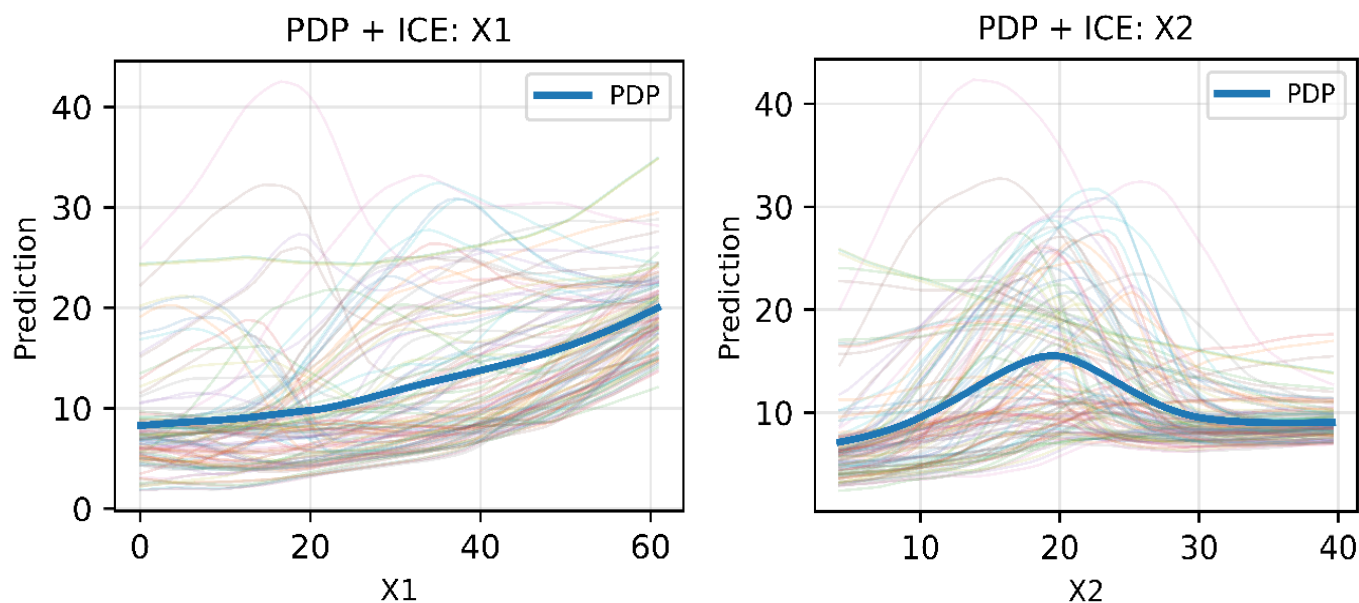


Fig. 8. Partial Dependence Plots (PDP) and Individual Conditional Expectation (ICE) plots for input variables derived from the DNN model

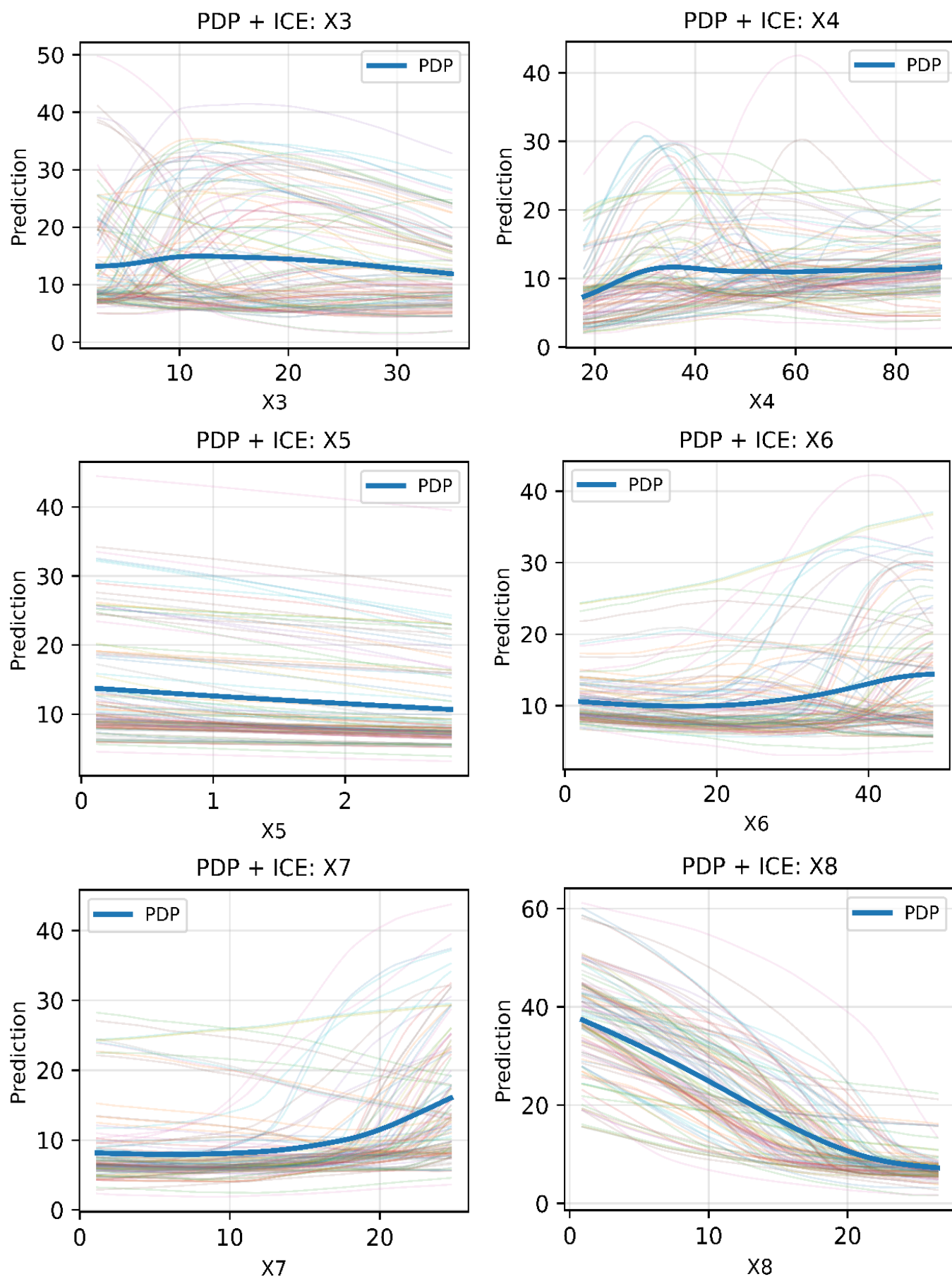
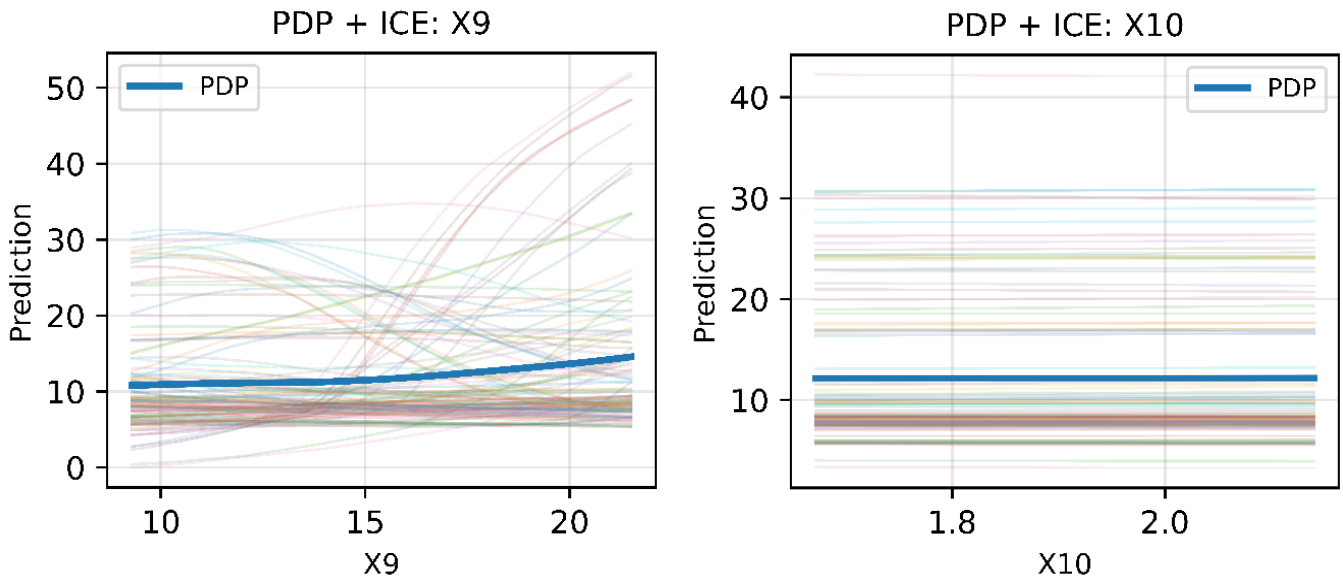


Fig. 8. (continued)

**Fig. 8.** (continued)

The analysis of the combined PDP and ICE plots derived from the DNN model for the input variables X1–X10 reveals distinct levels and patterns of influence on the model predictions (Fig. 8). Among the examined variables, X8 exhibits the strongest impact, as evidenced by a pronounced and nearly monotonic decrease in the PDP as X8 increases, together with the largest variation in predicted values compared to the other variables, indicating a strong negative and nonlinear effect. This finding aligns with plasticity's essential role in determining subgrade strength, as elevated plasticity correlates with increased fines activity and moisture sensitivity, leading to diminished penetration resistance and lower CBR values [3, 4]. In addition, X7 and X6 play important roles, showing clear increasing trends in their PDPs at medium to high value ranges and substantial dispersion in the ICE curves, suggesting the presence of complex interactions with other input variables. The variables X1, X2, X3, and X4 demonstrate moderate influence, characterized by mild nonlinear relationships or threshold effects with smaller changes in predictive magnitude. In contrast, X10 exhibits the weakest influence, with an almost flat PDP, indicating that variations in maximum dry density have only a limited effect on the predicted CBR within the present dataset. Although maximum dry density is generally

considered an important geotechnical parameter associated with soil compaction state and bearing capacity, the range of X10 varies relatively narrowly compared with some other input variables. This limited variation may reduce the discriminatory contribution of X10 during model training. Therefore, the low PDP response should not be interpreted as evidence that maximum dry density is geotechnically unimportant; rather, it suggests that, within the range and characteristics of the analyzed dataset, the predicted CBR is more strongly influenced by plasticity-related properties than by variations in maximum dry density. This indicates that, in the analyzed dataset, density-related impacts are predominantly represented indirectly via plasticity- and moisture-related characteristics. X5 follows, with a weak, slightly decreasing influence.

Overall, the comprehensive evaluation results indicate that the DNN model demonstrates superior predictive capability and reliability in estimating the CBR value compared to the RNN model. The consistently higher R^2 values, along with lower MAE and RMSE on both the training and test datasets, indicate that the DNN model is more effective at capturing the complex nonlinear relationships between geotechnical input variables and CBR. This is particularly meaningful from a geotechnical perspective, as CBR is governed by

the combined effects of soil gradation, compaction state, and plasticity characteristics, which interact through nonlinear mechanisms that are difficult to represent using conventional approaches. The closer agreement between predicted and observed values, the more concentrated residual distributions, and the faster and more stable convergence of the loss function further confirm the strong generalization ability of the DNN model and its low sensitivity to overfitting, thereby reflecting its effectiveness in modeling the key geotechnical factors controlling the bearing capacity of subgrade soils.

5. Conclusions

The results of this study demonstrate that applying deep learning models, particularly the Deep Neural Network (DNN), provides clear advantages for predicting the California Bearing Ratio (CBR) of subgrade soils, achieving higher accuracy and stability than the Recurrent Neural Network (RNN). The evaluation metrics and supplementary analyses indicate that the DNN model not only delivers reliable predictions but also reasonably represents the geotechnical mechanisms governing soil bearing capacity, in which CBR is strongly influenced by the compaction state and plasticity-related characteristics through complex nonlinear interactions. The close agreement between predicted and observed values, together with stable residual distributions, suggests strong generalization capability and low sensitivity to data variability. However, it should be noted that the number of samples used in this study remains limited and may not fully reflect the diversity of geological conditions and soil properties encountered in practice, potentially affecting the model's representativeness and generalizability. Expanding the dataset to include a wider range of soil types and geological conditions would help further enhance the reliability and practical applicability of the proposed CBR prediction approach.

Declaration of Generative AI

Authors declare that no generative Artificial Intelligence (AI) tools or AI-assisted technologies were used in the preparation of this manuscript.

References

- [1]. D.Q. Vu, D.D. Nguyen, Q.-A.T. Bui, D.K. Trong, I. Prakash, B.T. Pham. (2021). Estimation of California bearing ratio of soils using random forest based machine learning. *Journal of Science and Transport Technology*, 1(1), 45-58. <https://doi.org/10.58845/jstt.utt.2021.en.1.1.45-58>
- [2]. V.Y. Katte, S.M. Mfoyet, B. Manefouet, A.S.L. Wouatong, L.A. Bezeng. (2019). Correlation of California bearing ratio (CBR) value with soil properties of road subgrade soil. *Geotechnical and Geological Engineering*, 37(1), 217–234. <https://doi.org/10.1007/s10706-018-0604-x>
- [3]. A.A.M. Elhassasn, A. Qabur, K. Aati, A.E. Hassaballa, A. Abbadi, R.W. Bazuhair. (2025). Optimized Deep Learning Framework for Reliable Prediction of Pavement Subgrade CBR. *Transportation Engineering*, 21, 100382. <https://doi.org/10.1016/j.treng.2025.100382>
- [4]. AASHTO. (2022). AASHTO T193-22, Standard Method of Test for the California Bearing Ratio. *American Association of State Highway and Transportation Officials, USA*.
- [5]. B.B. Pule, J.A. Yendaw. (2024). The effect of geotechnical soil properties on CBR value: review. *AI in Civil Engineering*, 3(1), 19. <https://doi.org/10.1007/s43503-024-00039-1>
- [6]. S. Tamassoki, N.N.N. Daud, S. Wang, M.J. Roshan. (2023). CBR of stabilized and reinforced residual soils using experimental, numerical, and machine-learning approaches. *Transportation Geotechnics*, 42, 101080. <https://doi.org/10.1016/j.trgeo.2023.101080>
- [7]. W. Gao. (2024). The application of machine learning in geotechnical engineering. *Applied Sciences*, 14(11), 4712. <https://doi.org/10.3390/app14114712>
- [8]. C.C. Ikeagwuani. (2021). Estimation of modified expansive soil CBR with multivariate

- adaptive regression splines, random forest and gradient boosting machine. *Innovative Infrastructure Solutions*, 6(4), 199. <https://doi.org/10.1007/s41062-021-00568-z>
- [9]. I. Thapa, S. Ghani. (2025). Explainable AI-enhanced machine learning for CBR prediction in stabilized and unstabilized subgrade soils. *Applied Soft Computing*, 178, 113275. <https://doi.org/10.1016/j.asoc.2025.113275>
- [10]. A.R. Tenpe, A. Patel. (2020). Application of genetic expression programming and artificial neural network for prediction of CBR. *Road Materials and Pavement Design*, 21(5), 1183-1200. <https://doi.org/10.1080/14680629.2018.1544924>
- [11]. N. Tado, S. Medhajit, D. Pal. (2025). Forecasting California bearing ratio (CBR) of soil using machine learning algorithms: A review. *Research on Engineering Structures and Materials*, 11(1), 383-398. <https://doi.org/10.17515/resm2025-623ml0115rv>
- [12]. J. Khatti, K.S. Grover. (2024). CBR prediction of pavement materials in unsoaked condition using LSSVM, LSTM-RNN, and ANN approaches. *International Journal of Pavement Research and Technology*, 17(3), 750-786. <https://doi.org/10.1007/s42947-022-00268-6>
- [13]. M.Q. Shaaban, D.A.-J. Al-Hamdani, N.K. Lwti, N.G. Fakhir, A. Saad. (2026). Deep learning in geotechnical engineering: predicting soil behavior for safer constructions. *Research on Engineering Structures and Materials*, 12(2), 753-764. <https://doi.org/10.17515/resm2025-1234ic1011rs>
- [14]. M.A. Al-Zubi, M. Ahmad, S. Abdullah, B.J. Khan, W. Qamar, G.M.S. Abdullah, R.A. González-Lezcano, S. Paul, N.S.A. EL-Gawaad, T. Ouahbi, M. Kashif. (2024). Long short term memory networks for predicting resilient Modulus of stabilized base material subject to wet-dry cycles. *Scientific Reports*, 14, 27928. <https://doi.org/10.1038/s41598-024-79588-5>
- [15]. A.H.Y. Habal, A. Medjnoun, L. Djerbal, R. Bahar. (2025). Comprehensive review on predicting CBR values using machine learning techniques. *Asian Journal of Civil Engineering*, 26(8), 3153-3165. <https://doi.org/10.1007/s42107-025-01369-w>
- [16]. M.A. Hasan, P.B. S, R. Bag, R. Baadiga. (2025). The State-of-the-Art review on prediction of subgrade CBR: Past and present trends. *Indian Geotechnical Journal*, 55(5), 3410-3431. <https://doi.org/10.1007/s40098-024-01113-2>
- [17]. J.H. Steiger. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2), 245-251. <https://psycnet.apa.org/doi/10.1037/0033-2909.87.2.245>
- [18]. C.C. Aggarwal. (2018). Neural networks and deep learning. *Springer*. <https://doi.org/10.1007/978-3-319-94463-0>
- [19]. G. Li, S.K.S. Hari, M. Sullivan, T. Tsai, K. Pattabiraman, J. Emer, S.W. Keckler. (2017). Understanding error propagation in deep learning neural network (DNN) accelerators and applications. *Proceedings of SC17 the International Conference for High Performance Computing, Networking, Storage and Analysis*, 8, pp. 1-12. Association for Computing Machinery. <https://doi.org/10.1145/3126908.3126964>
- [20]. M. Reyad, A.M. Sarhan, M. Arafa. (2023). A modified Adam algorithm for deep neural network optimization. *Neural Computing and Applications*, 35(23), 17095-17112. <https://doi.org/10.1007/s00521-023-08568-z>
- [21]. Z.C. Lipton, J. Berkowitz, C. Elkan. (2015). A critical review of recurrent neural networks for sequence learning. *arXiv:1506.00019*. [arXiv. https://doi.org/10.48550/arXiv.1506.00019](https://doi.org/10.48550/arXiv.1506.00019)
- [22]. B.-J. Hou, Z.-H. Zhou. (2020). Learning with interpretable structure from gated RNN. *IEEE Transactions on Neural Networks and Learning Systems*, 31(7), 2267-2279.

- <https://doi.org/10.1109/TNNLS.2020.2967051>
- [23]. A. Rius, I. Ruisánchez, M.P. Callao, F.X. Rius. (1998). Reliability of analytical systems: Use of control charts, time series models and recurrent neural networks (RNN). *Chemometrics and Intelligent Laboratory Systems*, 40(1), 1-18. [https://doi.org/10.1016/S0169-7439\(97\)00085-3](https://doi.org/10.1016/S0169-7439(97)00085-3)
- [24]. D. Chicco, M.J. Warrens, G. Jurman. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, e623. DOI: 10.7717/peerj-cs.623
- [25]. T.J. Smith, C.M. McKenna. (2013). A comparison of logistic regression pseudo R² indices. *Multiple Linear Regression Viewpoints*, 39(2), 17-26.
- [26]. C.J. Willmott, K. Matsuura. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79-82. <https://doi.org/10.3354/CR030079>
- [27]. M.A. Ghorbani, R. Jani, F. Mohajeri, F.A. Daneshvar, E. Shabani, M. Khafagy. (2025). The Taylor Diagram with Distance: A New Way to Compare the Performance of Models. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 49, 305-321. <https://doi.org/10.1007/s40996-024-01477-8>
- [28]. Z.U. Rehman, Z. Aziz, U. Khalid, N. Ijaz, S.U. Rehman, Z. Ijaz. (2025). Artificial intelligence-driven enhanced CBR modeling of sandy soils considering broad grain size variability. *Journal of Rock Mechanics and Geotechnical Engineering*, 17(5), 3161–3179. <https://doi.org/10.1016/j.jrmge.2024.05.048>
- [29]. Y. Wang. (2024). A comparative analysis of model agnostic techniques for explainable artificial intelligence. *Research Reports on Computer Science*, 3(2), 25–33. <https://doi.org/10.37256/rrcs.3220244750>
- [30]. P. Smolensky. (1986). Neural and conceptual interpretation of PDP models. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, 22, pp. 390-431. MIT Press.
- [31]. A. Goldstein, A. Kapelner, J. Bleich, E. Pitkin. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics*, 24(1), 44-65. <https://doi.org/10.1080/10618600.2014.907095>
- [32]. T. Taskiran. (2010). Prediction of California bearing ratio (CBR) of fine grained soils by AI methods. *Advances in Engineering Software*, 41(6), 886-892. <https://doi.org/10.1016/j.advengsoft.2010.01.003>
- [33]. M. Shantal, Z. Othman, A.A. Bakar. (2023). A novel approach for data feature weighting using correlation coefficients and min–max normalization. *Symmetry*, 15(12), 2185. <https://doi.org/10.3390/sym15122185>
- [34]. J. Tan, J. Yang, S. Wu, G. Chen, J. Zhao. (2021). A critical look at the current train/test split in machine learning. *arXiv:2106.04525*. <https://doi.org/10.48550/arXiv.2106.04525>
- [35]. R. Jaiswal, B. Singh. (2024). A Comparative Study of Loss Functions for Deep Neural Networks in Time Series Analysis. *International Conference on Big Data, Machine Learning, and Applications*, pp. 147–163. Springer. https://doi.org/10.1007/978-981-99-3481-2_12