International Journal of Engineering and Technology Innovation, vol. 13, no. 3, 2023, pp. 175-188

Determination of the California Bearing Ratio of the Subgrade and Granular Base Using Artificial Neural Networks

Jose Manuel Palomino Ojeda[1], Billy Alexis Cayatopa Calderon[2], Lenin Quiñones Huatangari[1,*], Wilmer Rojas Pintado[2]

[1]Instituto de Ciencia de Datos, Universidad Nacional de Jaen, Jaen, Peru

[2]Instituto de Investigación en Sismológica y Construcción, Universidad Nacional de Jaen, Jaen, Peru

Received 26 October 2022; received in revised form 27 February 2023; accepted 04 March 2023

DOI: https://doi.org/10.46604/ijeti.2023.11053

Abstract

The objective of the research is to estimate the value of the California bearing ratio (CBR) through the application of ANN. The methodology consists of creating a database with soil index and CBR variables of the subgrades and granular base of pavements in Jaen, Peru, carried out in the soil mechanics laboratories of the city and the National University of Jaen. In addition, the Python library Seaborn is for variable selection and relevance, and the scikit-learn and Keras libraries were used for the learning, training, and validation stage. Five ANN are proposed to estimate the CBR value, obtaining an error of 4.47% in the validation stage. It can be concluded that this method is effective and valid to determine the CBR value in subgrades and granular bases of any pavement for its evaluation or design.

Keywords: CBR, subgrade, soil, prediction, model

1. Introduction

Peru has paved 17% of the 168,953.9 km of roads in the country, because in 2021 only Huanuco, Ancash, Junin, Piura, and San Martin experienced an increase of 10 km in their paving compared to Apurimac, Cajamarca, Pasco, and Moquegua that had a slight increase of 5.4 km. The government is investing in road works in the highlands, coast, and jungle, allowing connectivity between rural and urban areas, which boosts trade and the extraction of products to different parts of the country. By 2022, investments in pavements will exceed US$325 million [1].

In pavement construction, quality control is essential, and the bearing capacity of the subgrade, subbase, base, CBR values, and compaction characteristics should be properly evaluated [2]. The pavement is composed of layers of different thicknesses, and quality, which are supported on the subgrade. The layers that form the pavement structure are the subbase, base, and asphalt binder for flexible pavements, and the subbase and hydraulic concrete slab for rigid pavements [1]. However, in any type of pavement, the properties of the subgrade are determinant in the design and behavior of the structure. The subgrade can be constituted by the soil in its natural state or improved by mechanical, and chemical processes or using geosynthetics [3], being the important part of the pavement structure, which must be adequately compacted to maximize its strength while supporting loads of the previous pavement layers, as well as loads of moving traffic [4].

Pavement quality is tested by the California bearing ratio (CBR) test, which is a field/laboratory test used to determine the bearing strength of soil layers [5], is plotted using empirical curves, to establish the thickness of pavement and its component layers [6]. In the construction of pavements, earth dams, retaining wall backfills, and bridge abutments [7], civil

* Corresponding author. E-mail address: lenin.quinones@unj.edu.pe

English language proofreader: Chih-Wen Teng

engineers always find it difficult to obtain a representative CBR value due to time and cost [8]. The CBR test is performed in the laboratory using the American Association of State Highway and Transportation Officials (AASHTO) T193 method, on unaltered or compacted specimens, in a mold of 152 mm diameter and 178 mm height, packing the specimen to a height of 119 mm in five layers using a stress of 2700 kN-m/m^3 [9].

$$\text{CBR}=\frac{\text{Measured force}}{\text{Standard force}}\times 100\% \quad (1)$$

Afterward, the specimen is left to soak for four days to evaluate the saturated CBR value, its value is expressed in percentage, being the ratio between the effort required to penetrate a soil mass with a 50 mm diameter plunger at a speed of 1.25 mm/min [10], between the measured force and the standard force for a penetration given in Eq. (1). The standard forces corresponding to penetrations of 2.5 mm and 5.0 mm are 13.24 kN and 19.96 kN respectively [11].

According to Nalawade and Jadhao [6] and Alam et al. [12], CBR is expensive because it requires a specialized machine to perform the test, and it is time-consuming since it is necessary to evaluate several samples of subgrade and granular base for pavement design, it requires specialized labor since it must be performed according to the AASHTO T193 standard under the supervision of a specialist in soil mechanics, consequently, this generates serious delays in long-term projects. Faced with this problem, researchers have proposed the use of linear regression models and artificial neural networks (ANN).

The novelty and contribution of this study include the following aspects: (i) The use of the freely available Python programming language for the configuration, training, and validation of the ANN. (ii) ANN is proposed to estimate the CBR in the subgrade and granular base of the pavements of the city of Jaen, structured in an input layer, four hidden layers, and an output layer. (iii) Using the proposed method, the tested and trained results are highly accurate with a coefficient of determination (R^2) of 0.96 with fast processing time and simple implementation.

The objective of this study is to estimate the CBR value using ANN, using five models, developed in Python, to find the appropriate ANN architecture and the most influential parameters on the CBR in the subgrade and granular base, since the technical behavior of soils varies according to the area and with time. These models will allow the evaluation of the CBR in less time and at a low cost, predicting parameters in real time for the pavement design and evaluation phase.

The remaining sections of this study are organized as follows. Section 2 presents the literature review. Section 3 describes the materials and methods used. Section 4 presents the results with a detailed description. Section 5 discusses the prediction results of the proposed ANN. Finally, Section 6 presents the conclusions.

2. Literature Review

The authors have addressed the problem using linear regression and ANN models. Table 1 shows a review of the different linear regression and ANN models used to estimate CBR. Teklehaymanot and Alene [8] propose a linear model correlating the CBR with the different variables of percent fines, percent sand, PG, PL, PI, MX, OWC, SG, and WC, obtaining different equations for each parameter, as do [3, 8]. Other authors [2, 4, 13-18] have developed multiple linear regression models using different soil index variables such as uniformity coefficient, curvature coefficient, SG, fine soil size, standard proctor MX, LL, PI, sand percentage, PG, PL among others.

Given the limitations of the previous methods, ANN models are excellent for solving complex problems with multiple inputs, Janjua and Chand [13] and S. K. Das and A. K. Sabat [19] propose an ANN of 7-3-1 topology (seven input layers, three hidden layers, and one output layer) using the sigmoid function, with clay, ignition loss, and PG being the most influential variables in CBR estimation, obtaining a R^2 of 0.87. Zabielska-Adamska and Sulewska [20] propose an ANN of 8-5-1 topology with moisture content and MX being the most influential parameters for CBR prediction, obtaining a $R^2 = 0.95$. Taskiran [21]

creates a 7-4-1 topology ANN capable of learning from the relationship between CBR and soil properties by performing a sensitivity analysis, where the most influential parameter in the CBR is the MX, among others such as PI, LL, optimum moisture content, G size, and sand, obtaining a $R^2 = 0.91$.

Table 1 Linear regression and ANN models

Model	Architecture	Input variables	Activation function	R^2	Reference
Model linear	-	PI, IBI, UCS, MX, GV, OWC	-	0.77	[3]
Model linear	-	PF, PS, PG, PL, PI, MX, OWC, SG, WC, SC	-	0.89	[8]
Multiple linear regression	-	CU, AC, SG, FS	-	0.83	[2]
Multiple linear regression	-	MX, OWC, LL, PL, PI	-	0.98	[4]
Multiple linear regression	-	MX, LL, OWC	-	0.84	[13]
Multiple linear regression	-	DLPI, OWC	-	0.85	[14]
Multiple linear regression	-	OWC, LL, CC, CU, D30, D50	-	0.85	[15]
Multiple linear regression	-	MX, OWC, LL, PL, PI	-	0.81	[16]
Multiple linear regression	-	PI, PF, OWC, MX	-	0.73	[17]
Multiple linear regression	-	LL, PI, MX, OWC, PG, PS, PF	-	0.84	[18]
ANN	15-11-1	TG, OWC, MX, LL, PI, PSO, SS, GY, OG	Hyperbolic tangent	0.78	[10]
ANN	2-3-1	OH, MX	Hyperbolic tangent	0.89	[5]
ANN	7-3-1	A, SI, FE, CA, LOI, PG, OWC	Sigmoid function	0.87	[19]
ANN	8-5-1	D50, ST, MX, WC, MX/OWC, OWC/OCH	Sigmoid function	0.90	[20]
ANN	7-4-1	LL, PI, MX, OWC, C+SC, PS, PG	Logsig–logsig	0.91	[21]

ANN represents a simplified model of the human brain, with a complex communication network consisting of hundreds of simple processing units connected. Thus, neural networks manage information in a similar way to the human brain. A neural network is composed of numerous interconnected neurons working simultaneously to solve a specific problem; therefore, neural networks cannot be configured to perform an exact task but learn by examples that must be carefully selected to train the network correctly [10].

3. Materials and Methods

3.1. California bearing ratio

It is the measure of the shear strength of soil under known moisture and density conditions. The CBR evaluation method is standardized in the AASHTO T193 and the American Society for Testing and Materials (ASTM) D1883. It is an indicator of the suitability of the natural subgrade soil as a construction material. If the CBR value of the subgrade is high, it means that the subgrade is firm and as a result, the pavement thickness design can be reduced, but if it is low, the opposite is true [17].

The test was performed by AASHTO T193, which states that when more than 75% by weight of the sample passes through the 19.1 mm (3/4") sieve, the material passing through this sieve shall be used for the test. When the fraction of the sample retained on the 19.1 mm (3/4") sieve is greater than 25% by weight, the material retained on that sieve shall be separated and replaced by an equal proportion of material between the 19.1 mm (3/4") and 4.75 mm (No. 4) sieves, obtained by sieving another portion of the sample.

Then, three 5 kg samples were selected for each CBR mold, then the mold was weighed with its base and the spacer disc was placed with the filter paper, then the specimens were prepared in 5 layers with different compaction energies 55, 26, and 12 blows per layer and with optimum moisture content obtained from the Proctor test, once the compaction was finished, the collar was removed, the weight of the specimen was noted, then the specimen was reassembled inverted, without the spacer disc, placing a filter paper between the mold and the base.

Later, a perforated plate with a shank is placed on the surface of the inverted specimen, and on it, the necessary rings are placed to complete an overburden such that it produces a pressure equivalent to that originated by all the layers of materials that will go on top of the tested soil, the approximation will be within 2.27 kg (5.5 lb); the first reading is taken to measure the swelling and then, the mold is immersed in the tank with the overburden placed leaving free access to water in the lower and upper part of the specimen. The specimens are kept in these conditions for 96 hours (4 days). At the end of the immersion period, the strain gauge is read again to measure swelling; afterward, the specimens were transferred to the CBR machine, where an overburden was applied at a uniform penetration rate of 1.27 mm (0.05") per minute (see Fig. 1). All data obtained in the test were recorded and the CBR value was calculated with the stress-penetration values for the 2.54 mm and 5.08 mm by dividing the corrected stresses by the reference stresses 6.9 MPa (1000 lb/plg^2) and 10.3 MPa (1500 lb/plg^2) respectively, multiplying by 100.

Fig. 1 Application of the overburden on the soil sample

3.2. Artificial neural networks

ANN is a mathematical tool obtained from the biological structure of neurons, it is composed of a set of neurons interconnected with each other, and these neurons employ the backpropagation algorithm, used especially for data prediction and modeling. This means that the signal is sent forward, and errors are propagated backward. This minimizes the total quadratic error of the output calculated by the network. Fig. 2 shows the typical layout and process of a feed-forward neural network with a backpropagation algorithm. A neuron with a single R number of the input element is shown below [12].

$$P=\{P_1 P_2 \cdots P_R\} \tag{2}$$

$$W=\{W_{1,1} W_{1,2} \cdots W_{1,R}\} \tag{3}$$

$$n=\{W_{1,1}P_1+W_{1,2}P_2+\cdots+W_{1,R}P_R+B\} \tag{4}$$

These inputs are multiplied by the weights, as shown in Eq. (3). WP is the dot product of the matrices W and vector P. The neuron has a bias called B, which is summed with the weighted input WP to form the new output n, which is given by Eq. (4).

$$f(x)=\frac{1}{1+e^{-x}} \tag{5}$$

$$f(x)=\max(0,X)=\begin{Bmatrix}0 \text{ for } x<0\\0 \text{ for } x\geq 0\end{Bmatrix} \tag{6}$$

Subsequently, the net input n is processed through a transfer function f to produce the nodal output A, which is reweighted and passed to the next layer processing unit. The transfer function can be linear, log sigmoid, tan sigmoid, etc. The activation functions used in the different models were sigmoid and rectified linear unit (ReLU), Eq. (5) describes the sigmoid activation function and Eq. (6) describes the ReLU activation function.

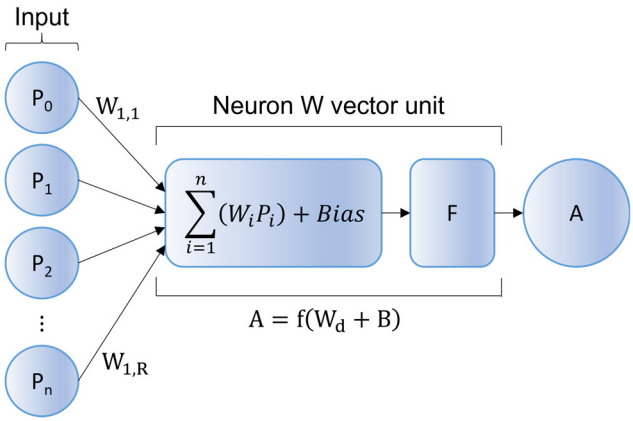

Fig. 2 Structure of a feed-forward neuron

3.3. *Database for the training of ANN*

The database was created from the information obtained from standardized laboratory tests: granulometry, LL, PL, PI, proctor, and CBR (see Table 2), of samples taken from pavements in the province of Jaen, tested in the soil mechanics laboratory of the Professional School of Civil Engineering of the National University of Jaen and soil mechanics laboratories of the city of Jaen, evaluated between the years 2018-2022. AASHTO and ASTM standards were used to perform the tests.

Table 2 Standardized laboratory tests

Essay	Method		Purpose of the test
	AASHTO	ASTM	
Granulometry	T88	D422	Determine the quantitative distribution of soil particle sizes.
Liquid limit	T89	D4318	Find the moisture content between the liquid and plastic states.
Plastic limit	T90	D4318	To find the moisture content between the plastic and semi-solid states.
Plasticity index	T90	D4318	Find the range of moisture content above which the soil is in its plastic state.
Proctor	T180	D1557	Find the maximum dry density of the soil at the optimum compaction moisture.
CBR	T193	D1883	Determine the bearing capacity. It allows inferring the resilient modulus.

The data matrix collected consisted of 521 CBR records and eight variables (see Fig. 3). The variables collected were G, sand, fine, LL, PL, PI, MX, and optimum moisture content (see Table 3), presenting a bimodal, symmetric, right-symmetric, and left-symmetric distribution.

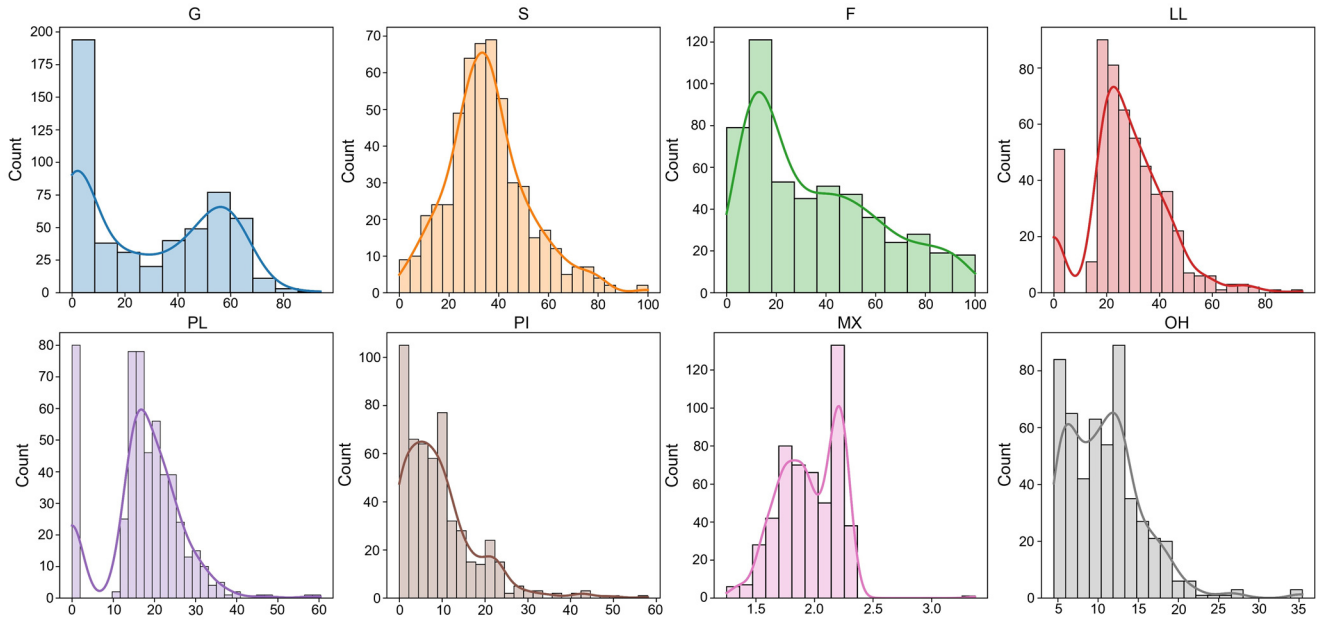

Fig. 3 Database histogram

Table 3 Parameters and characteristics to determine the CBR

Group	Concept	Variable	Unit	Symbol	ANN representation
Granulometry	Classification	Gravel	%	G	Real numeric
		Sand	%	S	Real numeric
		Fine	%	F	Real numeric
Atterberg limits	Classification	Liquid limit	%	LL	Real numeric
		Plastic limit	%	PL	Real numeric
		Plasticity index	%	PI	Real numeric
Proctor	Density	Maximum dry density	g/cm^3	MX	Real numeric
		Optimum moisture content	%	OH	Real numeric
CBR	Support capacity	CBR	%	CBR	Real numeric

The statistical analysis of the input and output variables is shown in Table 4, the maximum standard deviation was 26.17 for the fine variable and a minimum of 0.25 for the MX variable, and the minimum and maximum values were between 0 and 100 respectively.

Table 4 Statistical information on variables

	Parameter used	Mean	Std. deviation	Minimum	Maximum
Input parameters	G	28.61	25.52	0.00	94.00
	S	35.71	16.40	0.00	100.00
	F	35.56	26.17	0.00	100.00
	LL	27.76	14.76	0.00	94.00
	PL	17.31	9.47	0.00	60.40
	PI	9.27	8.70	0.00	58.00
	MX	1.95	0.25	1.20	3.37
	OH	10.86	4.75	4.50	35.4
Output parameter	CBR	35.93	31.56	0.40	100

For the selection of the input variables and the conformation of the information vectors, the conditions of dependence on the CBR value were taken into account. The eight variables collected in the database were grouped: granulometry, Atterberg limits, proctor, and CBR. The latter contains the output variable CBR value. Likewise, the variables that most influence the CBR value was selected, for which a statistical correlation analysis was performed using the Python Seaborn library (see Fig. 4) [22].

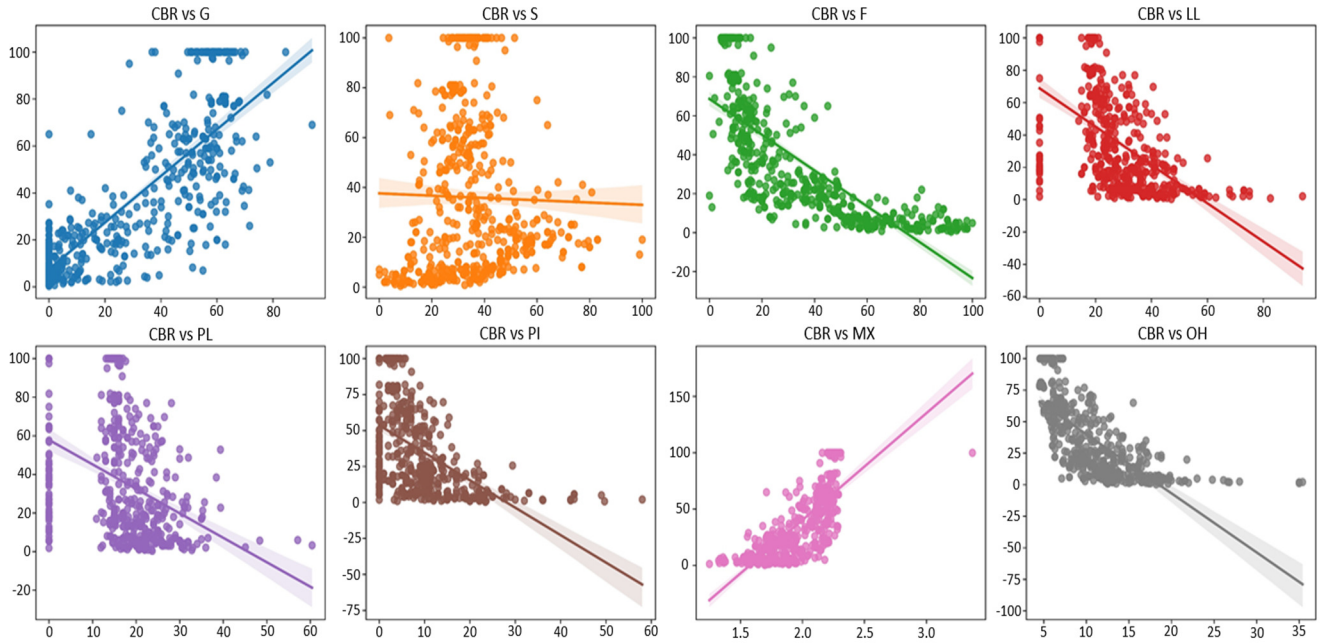

Fig. 4 Linear correlation between input variables and CBR

3.4. Database for validation of ANN

A database was constructed for the validation process of the ANN considering eight variables with the greatest impact on the CBR value (see Table 5). This source was obtained experimentally by performing granulometry, LL, PL, PI, proctor, and CBR tests on 70 samples of subgrade and granular base (see Tables 6-7), according to ASTM D422, D4318, D4318, D4318, D1557, and D1883. To observe repeatability, each test was repeated three times until the results obtained varied within ± 0.5% in the laboratory [23].

Table 5 Variables selected for training

Variables	Unit	Symbol	R
Gravel	%	G	0.8
Sand	%	S	-0.024
Fine	%	F	-0.76
Liquid limit	%	LL	-0.55
Plastic limit	%	PL	-0.38
Plasticity index	%	PI	-0.53
Maximum dry density	g/cm^3	MX	0.76
Optimum moisture content	%	OH	-0.70

Table 6 Results of laboratory tests performed on soil samples from group 1 for validation

Index	Gravel	Sand	Fine	Liquid limit	Plastic limit	Plasticity index	Maximum dry density	Optimum moisture content
1	0	33	67	33	20	13	1.9	12
2	5	20	75	41	20	21	1.95	12.5
3	1	11	88	44	21	23	1.88	13
4	3	18	79	43	20	23	1.92	12.5
5	2	17	81	43	21	22	1.88	13.2
6	8	16	76	43	21	22	1.94	12.3
7	0	2	98	37	22	15	1.78	16.5
8	0	11	89	32	22	10	1.76	13
9	0	33	67	30	22	8	1.76	14
10	63	27	10	17	15	2	2.22	4.6
11	62	29	9	19	15	4	2.22	4.6
12	17	52	31	16	15	1	2.1	8.5
13	55	34	11	27	13	14	2.13	6
14	50	34	16	24	20	4	2.08	9
15	0	52	48	48	25	23	1.87	12
16	0	63	37	38	35	3	1.62	12
17	0	41	59	51	29	22	1.87	13.4
18	0	9.5	90.5	57	24	33	1.609	11.81
19	0	6.3	93.7	53	30	23	1.612	9.62
20	0	18.9	81.1	73	30	43	1.595	11.48
21	58.58	34.19	7.23	18.64	15.79	2.85	2.24	6.06
22	69.08	26.38	4.54	19.55	17.6	1.95	2.23	5.94
23	54.37	37.8	7.83	17.76	15.11	2.65	2.25	5.99
24	60.44	32.88	6.68	17.48	15.36	2.12	2.3	5.5
25	59.36	31.52	9.12	18.87	15.54	3.33	2.31	5.61
26	60.61	34.02	5.37	0	0	0	2.23	4.84
27	58.13	36.83	5.04	0	0	0	2.23	5.89
28	50.74	41.19	8.07	0	0	0	2.26	5.19
29	62.68	32.04	5.28	0	0	0	2.24	5.16
30	56.75	37.6	5.65	18.82	14.95	3.87	2.3	5.62
31	57.06	34.99	7.95	18.92	13.54	5.38	2.29	5.25
32	63.46	32.06	4.48	0	0	0	2.27	4.81
33	64.34	29.72	5.94	18.49	16.09	2.4	2.31	5.92
34	61.3	32	6.7	19.85	16.01	3.84	2.26	5.97
35	53.01	36.1	10.89	21.45	16.35	5.1	2.26	6.09

Table 7 Results of laboratory tests performed on soil samples from group 2 for validation

Index	Gravel	Sand	Fine	Liquid limit	Plastic limit	Plasticity index	Maximum dry density	Optimum moisture content
36	84.55	3.64	11.81	17.81	15.45	2.36	2.26	5.27
37	46.22	36.96	16.82	23.91	16.8	7.11	2.14	6.63
38	5	20	75	41	20	21	1.95	12.5
39	1	11	88	44	21	23	1.88	13
40	3	18	79	43	20	23	1.92	12.5
41	2	17	81	43	21	22	1.88	13.2
42	8	16	76	43	21	22	1.94	12.3
43	0	2	98	37	22	15	1.78	16.5
44	0	11	89	32	22	10	1.76	13
45	0	33	67	30	22	8	1.76	14
46	60	25	15	33	17	16	2.18	9.2
47	5.02	34.58	60.4	58.7	16.48	42.22	1.64	17
48	0	5.6	94.4	60.6	30.6	30	1.51	35
49	0	8.5	91.5	63	31	32	1.25	35
50	0	5.6	94.4	72.7	30.5	42.2	1.85	21.2
51	2	32.5	65.5	82.5	32.8	49.7	1.58	19.6
52	0	11	89	94	36	58	1.33	35.4
53	0	11	89	75	32	43	1.47	28
54	0	9	91	48	26	22	1.61	20
55	2	27	71	49	24	25	1.69	19.6
56	9	39	52	59	34	25	1.69	19
57	0	11.7	88.3	20.8	17.5	3.3	1.814	14.6
58	0	59.6	40.4	22.8	18.6	4.2	1.958	10.1
59	55	0	45	28	17	11	2	12
60	0	31	69	26	17	9	1.86	12.2
61	0	64	36	22	15	7	1.95	10.2
62	0	20	80	42	22	20	1.89	12.5
63	0	60	40	21	16	5	1.92	10.5
64	10	36	54	30	22	8	1.71	15
65	0	37	63	31	16	15	1.82	13
66	0	58	42	24	14	10	1.71	15.5
67	0	80	20	15	14	1	1.7	12
68	12	46	42	30	24	6	1.7	16
69	0	77	23	19	14	5	1.67	12
70	56	30	14	21	17	4	2.24	5.8

3.5. Development, training, and validation of ANN

For data preparation, the Python Keras library was used, which is a high-level neural network API capable of running on top of the TensorFlow or Theano libraries [24]. Using the JupyterLab interface, the database was imported with the Pandas library. Then, the data were preprocessed, normalizing the input variables with the scikit-learn min-max library [25] and scaling the data between 0 and 1. The neural network model was defined in Keras with a sequence of layers creating the input layer with the input dim argument, the hidden layers, and the output layer, specifying the optimizer "Adam" and the loss function the mean square error (MSE). The neural network models were trained with the "fit ()" command, which has a built-in training cycle. Running the code created seven ANN architectures (see Table 8), selecting the architecture with the highest correlation coefficient. The ANN architecture described as [k1 k2 k3 k4] refers to the hidden layers respectively.

Table 8 Selection of the number of neurons in the hidden layer

ANN architecture	R^2
[08 15 10 05]	0.86
[08 06 07 05]	0.89
[08 20 15 25]	0.83
[25 30 12 15]	0.82
[05 10 08 04]	0.85
[12 08 05 10]	0.81
[08 04 07 04]	0.80

Subsequently, five neural networks ANN_1, ANN_2, ANN_3, ANN_4, and ANN_5 trained, and composed of different input variables, with multilayer typology structured by an input layer, four hidden layers with eight, six, seven, and five neurons respectively, and an output layer in each neural network, corresponding to the CBR value. The input and output variables of each neural network developed and trained are shown in Table 9. In addition, the G variable was used in all models because it has a high correlation value concerning the other variables.

Table 9 ANN conformation and input variables for CBR value estimation

Neural network	Entries	Output
ANN_01	G, F, PI, MX, OH	CBR
ANN_02	G, S, LL, PL, PI, MX	
ANN_03	G, S, F, PI, MX, OH	
ANN_04	G, S, F, PL, MX	
ANN_05	G, F, LL, MX, OH	

The training, validation, and testing phase was performed with the Keras Python library, with which several configurations were designed for the five ANN consisting of different learning algorithms and parameter variations. In the validation phase, each network performed the processing of the database considered to validate the models. Considering the relationships learned in the training process, the synaptic weights were saved and stored in vector form. The CBR value will be the output of the network. By distinguishing this prediction from the actual value, we will obtain evidence of the predictive capacity of the model.

3.6. ANN performance evaluation

The following statistics were used:

$$\text{RMSE} = \sqrt{\sum_{t=1}^{T} \frac{(Yt - Pt)^2}{T}} \tag{7}$$

$$\text{SSE} = \sum_{t=1}^{T} \left(P_t - \bar{P}_t\right)^2 \tag{8}$$

$$\text{R}^2 = 1 - \frac{\sum_{t=1}^{T} \left(Y_t - P_t\right)^2}{\sum_{t=1}^{T} \left(Y_t - \bar{P}_t\right)^2} \tag{9}$$

$$\text{RE} = \left| \frac{R_{real} - R_{simul}}{R_{real}} \right| \times 100 \tag{10}$$

root mean square error (RMSE) (Eq. (7)), total squares of error (SSE) (Eq. (8)), coefficient of determination (R^2) (Eq. (9)), and relative error (RE) (Eq. (10)). Where Y_t is the desired output, P_t is the obtained output, $\bar{P}_t$ is the average of the obtained outputs, and T is the number of records taken in each phase (learning and validation). The error found for each simulated data was established by Eq. (10) where R_{real} and R_{simul} are the real and simulated CBR values, respectively.

4. Results

After designing the five proposed ANN: ANN_1, ANN_2, ANN_3, ANN_4, and ANN_5, in Python; configured with different input variables and trained with 521 CBR records, the CBR value of 70 experimentally obtained data was predicted and compared (see Tables 6-7) to validate the model. Fig. 5 shows the actual CBR value obtained experimentally in the laboratory of the faculty of civil engineering versus the CBR predicted by the different ANN. The ANN that was closest to the actual CBR value were models ANN_1, ANN_2, and ANN_3.

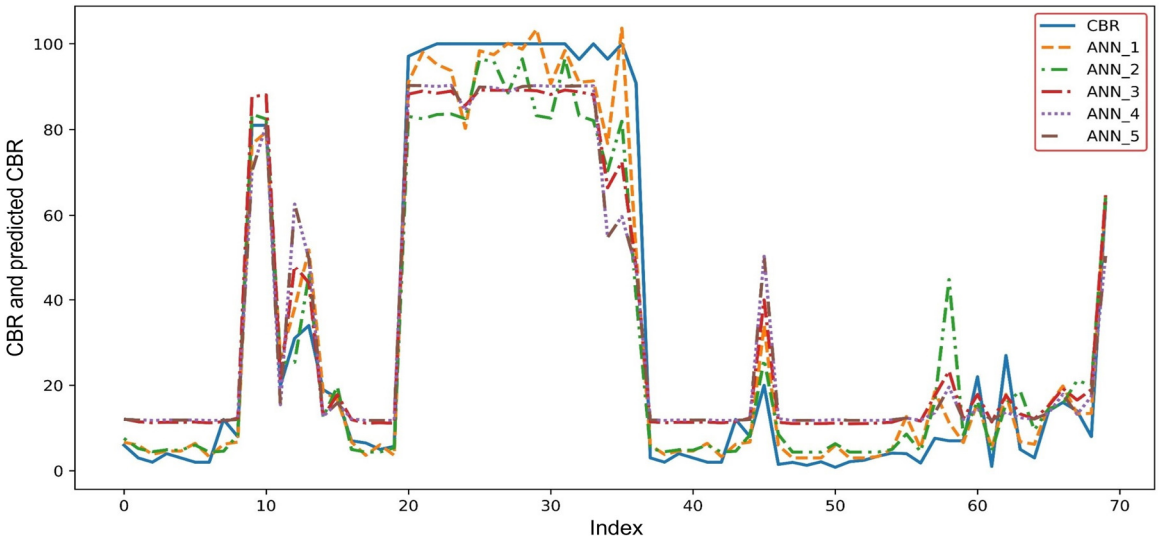

Fig. 5 CBR value estimation

The structure of the ANN_1, ANN_2, and ANN_3 networks and their statistics are shown in Figs. 6-11; the ANN_4 and ANN_5 networks were discarded for having a low MSE during the training and validation process as visualized in Fig. 5; where these models are far from the real CBR values. The three ANNs: ANN_1, ANN_2, and ANN_3, were evaluated considering the experimentally obtained database (validation database) in terms of the correlation coefficient, see Table 10.

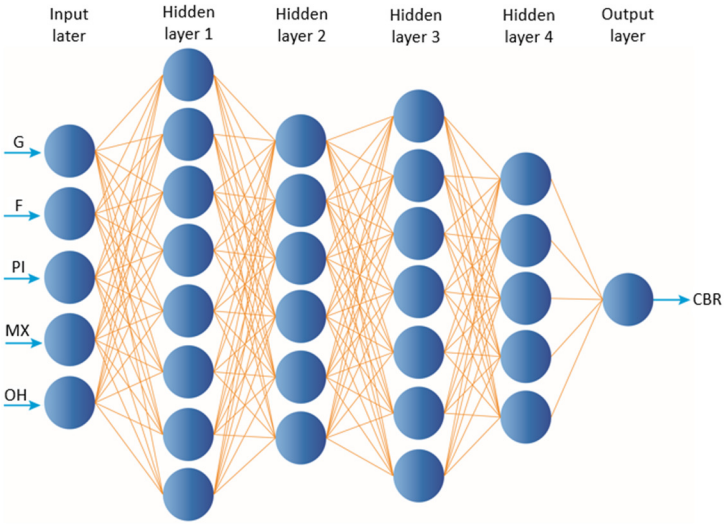

Fig. 6 Structure of the ANN_1

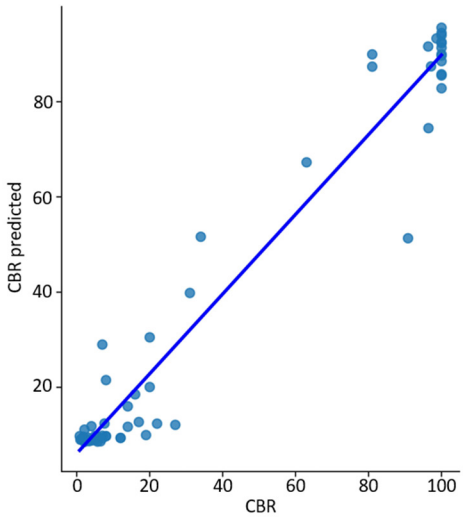

Fig. 7 Predicted and actual values of the ANN_1

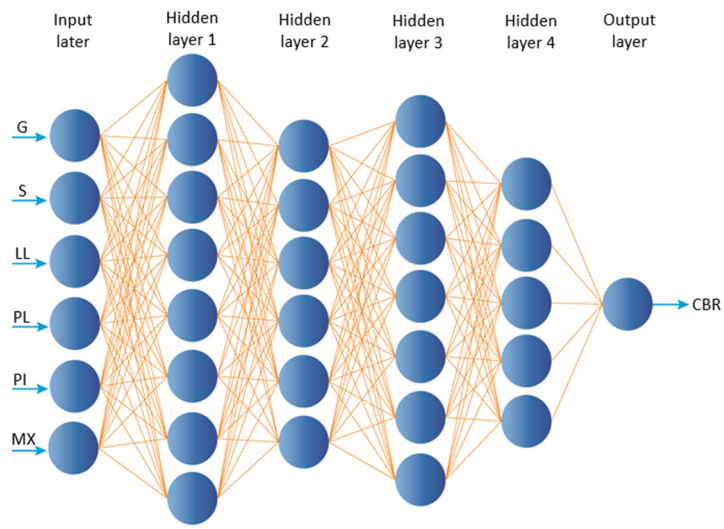

Fig. 8 Structure of the ANN_2

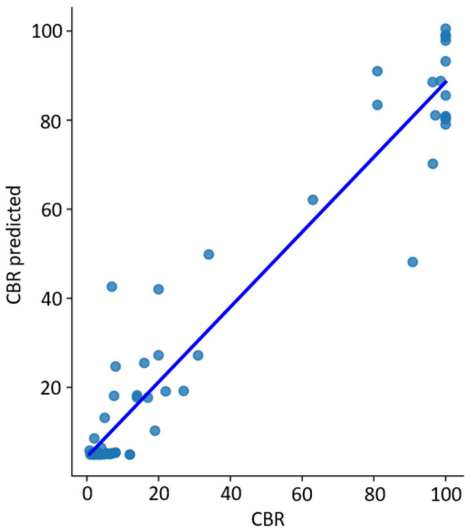

Fig. 9 Predicted and actual values of the ANN_2

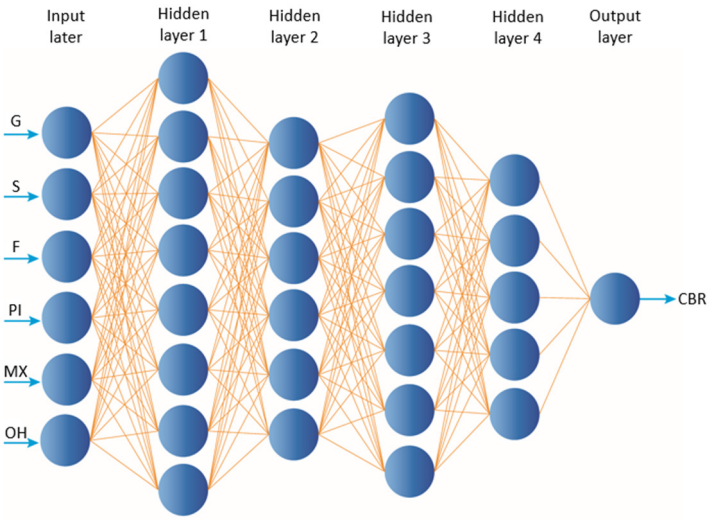

Fig. 10 Structure of the ANN_3

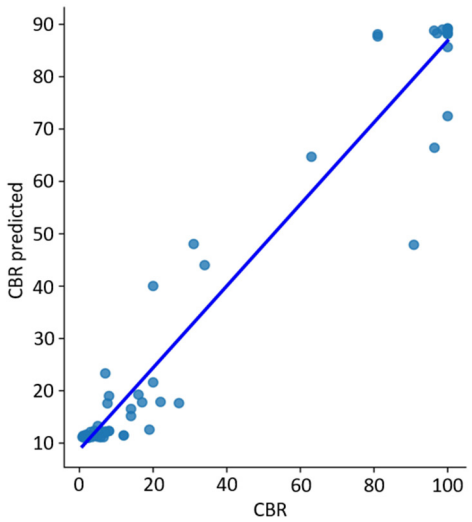

Fig. 11 Predicted and actual values of the ANN_3

Table 10 Performance evaluation of neural networks ANN_01, ANN_02, and ANN_03

ANN	Relative error	RMSE	R	R^2
ANN_1	4.47%	7.48	0.98	0.96
ANN_2	7.15%	11.20	0.97	0.94
ANN_3	9.35%	11.34	0.98	0.95

5. Discussion

The research was carried out in the province of Jaen, Peru because pavements in various areas need repair and maintenance, for which the subgrade and granular base must be evaluated using the CBR value. The database for the construction of the model consisted of 512 records and 8 variables, which represent the characteristics of the soils used to determine the CBR. In the input layer of the proposed neural networks, eight variables that have a greater incidence on the CBR value were considered: granulometry, LL, PL, PI, MX, and optimum moisture content, collected from the laboratory of the professional school of civil engineering and soil mechanics of the city of Jaen, being similar to the database collected by various researchers.

Abdella et al. [17] perform laboratory tests to make a database: Atterberg limit, compaction test, granulometric analysis, and CBR tests based on the AASHTO standard, Al-Busultan et al. [10] collect test data on granulometry, OWC, MX, LL, PI and the percentages of SO3 from the Karbala construction laboratory, Alam et al. [12] make a database according to soil index properties such as SG, coefficient of uniformity, CC, LL, PL, PI, optimum moisture content and MX for alluvial soils in West Bengal, India. The most influential variables for determining the CBR value in Jaen, Peru, are the PG and fines, PI, MX, and optimum moisture content, different from those obtained by Rehman et al. [15] MX, optimum moisture content, the uniformity coefficient, the particle size corresponding to 30% passing (D30) and the average particle size (D50) in Pakistan pavements, Park and Kim [26] the percentage passing the No. 200 sieve, the LL and compaction characteristics performed in Yeongdong construction zone. According to Teklehaymanot and Alene [8], soil type is not the only parameter that affects the CBR value but also varies with the different soil properties of each area.

Five neural networks were designed and trained with the Python Keras library, from which those with the lowest MSE were selected, concerning the performance evaluation of the estimation of the CBR value and the graphical behavior between the actual and estimated values of the proposed neural model is consistent with the behavior presented in the neural model proposed by several authors [10, 19, 21, 27] where the results obtained in the performance indicator, the coefficient of determination is R^2 = 0. 97, R^2 = 0.88, R^2 = 0.78 and R^2 = 0.74 respectively, and when compared with the coefficient of determination obtained in the research R^2 = 0.96 is acceptable since this value is higher than that obtained by three of the

mentioned researches and is within the range to estimate with great precision the value of the CBR with a significance of 4.5%. The R^2 takes values between 0 and 1, see Eq. (9), therefore, if the value obtained is close to one, it implies that the model fit of the CBR value we are estimating will be better.

6. Conclusions

A database was created with soil index variables, extracted from the soil mechanics laboratories of the city of Jaen, certified by the national quality institute. Afterward, five ANN with feedforward typology, backpropagation learning, and multilayer architecture were elaborated and developed to estimate the CBR value of pavements in Jaen, Peru. The following conclusions were obtained:

(1) The database consisted of 521 CBR records and eight soil index variables: PG, sands, and fines, LL, PL, PI, MX, and optimum moisture content, obtained from standardized laboratory tests; performed on different samples taken from pavements in the city of Jaen between the years 2018-2022.

(2) An experimental database of 70 samples of subgrade and granular bases of pavements from the city of Jaen was created to validate the proposed ANN models, performing tests of Granulometry, LL, PL, PI, proctor, and CBR, in the soil mechanics laboratory of the professional school of civil engineering of the National University of Jaen.

(3) Five ANNs (ANN_1, ANN_2, ANN_3, ANN_4, and ANN_5) were created, because they are excellent classifiers/pattern recognizers and are used where traditional techniques do not work, allowing complex problems to be solved. They were programmed using the Python Keras library, with multilayer typology structured in one input layer, four hidden layers, and one output layer. ANN_4 and ANN_5 were eliminated for presenting an R^2 value lower than 0.8 during training.

(4) The predicted values of the proposed neural networks were evaluated with the CBR value of the subgrade and granular base samples using the mean error, RMSE, and R^2 statistics. Being ANN_1 is the one that best satisfies the model since it estimates the CBR value with an error of 4.47% and when evaluated with the $R^2 = 0.96$, it gives us a significance of 4.5%, allowing the possibility of automating the training processes and the estimation processes of the CBR value with precision.

Nomenclature

A	Al_2O_3		GV	Volumetric swelling		R	Correlation coefficient
AASHTO	American Association of State Highway and Transportation Officials		GY	Gypsum		R^2	Coefficient of determination
ANN	Artificial neural networks		IBI	Immediate bearing index		RE	Relative error
ASTM	American Society for Testing and Materials		LL	Liquid limit		ReLU	Rectified linear unit
C	Clay		LOI	Loss on ignition		RMSE	Root mean square error
C+SC	Clay + Silt content		MSE	Mean square error		S	Sand
CA	CaO		MX	Maximum dry density		SC	Silt content
CBR	California bearing ratio		OCH	Optimum standard proctor dry moisture content		SG	Specific gravity
CC	Coefficient of curvature		OG	Organic		SI	SiO_2
CU	Coefficient uniformity		OH	Optimum moisture content		SS	Soluble salt
D30	Grain size corresponding to 30% passing		OWC	Optimum water content		SSE	Squares of error
D50	Grain size corresponding to 50% passing		PF	Percentage of fines		ST	Saturation
DLPI	Dynamic lightweight cone penetration index		PG	Percentage of gravels		TG	Tested for gradation
F	Fine		PI	Plasticity index		UCS	Unconfined compressive strength
FE	Fe_2O_3		PL	Plastic limit		WC	Water content
FS	Fine soil		PS	Percentage of sand			
G	Gravel		PSO	Percentages of SO3			

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] "Investment in Transport Infrastructure Exceeds USD 325 Million in 2022," https://www.gob.pe/institucion/ositran/noticias/643296-inversion-en-infraestructuras-de-transporte-supera-los-usd-325-millones-en-el-2022, August 22, 2022. (In Spanish)

[2] D. Kuttah, "Strong Correlation Between the Laboratory Dynamic CBR and the Compaction Characteristics of Sandy Soil," International Journal of Geo-Engineering, vol. 10, no. 1, article no. 7, December 2019.

[3] B. Celauro, A. Bevilacqua, D. L. Bosco, and C. Celauro, "Design Procedures for Soil-Lime Stabilization for Road and Railway Embankments. Part 1-Review of Design Methods," Procedia-Social and Behavioral Sciences, vol. 53, pp. 754-763, October 2012.

[4] F. Iqbal, A. Kumar, and A. Murtaza, "Co-Relationship between California Bearing Ratio and Index Properties of Jamshoro Soil," Mehran University Research Journal of Engineering & Technology, vol. 37, no. 1, pp. 177-190, January 2018.

[5] S. Bhatt and P. K. Jain, "Prediction of California Bearing Ratio of Soils Using Artificial Neural Network," American International Journal of Research in Science, Technology, Engineering & Mathematics, vol. 8, no. 2, pp. 156-161, November 2014.

[6] R. D. Nalawade and P. D. Jadhao, "Prediction of CBR Value of Stabilized Black Cotton Soil Use for Road Construction," International Journal of Engineering and Advanced Technology, vol. 9, no. 3, pp. 2317-2322, February 2020.

[7] E. A. Sandoval-Vallejo and W. A. Rivera-Mena, "Correlation between CBR and Resistance to Unconfined Compression," Ciencia e Ingeniería Neogranadina, vol. 29, no. 1, pp. 135-152, June 2019. (In Spanish)

[8] F. G. Teklehaymanot and E. Alene, "CBR Predictive Model Development from Soil Index and Compaction Properties in Case of Fine-Grained Soils of Debre–Tabor City, Ethiopia," International Journal of Advanced Science and Engineering, vol. 8, no. 2, pp. 2224-2234, November 2021.

[9] G. Norouznejad, I. Shooshpasha, S. M. Mirhosseini, and M. Afzalirad. "Effect of Zeolite on the Compaction Properties and California Bearing Ratio (CBR) of Cemented Sand," International Journal of Engineering and Technology Innovation, vol. 11, no. 3, pp. 229-239, June 2021.

[10] S. Al-Busultan, G. K. Aswed, R. R. A. Almuhanna, and S. E. Rasheed, "Application of Artificial Neural Networks in Predicting Subbase CBR Values Using Soil Indices Data," IOP Conference Series: Materials Science and Engineering, vol. 671, article no. 012106, 2020.

[11] M. M. Nujid, J. Idrus, N. A. Azam, D. A. Tholibon, and D. Jamaluddin, "Correlation Between California Bearing Ratio (CBR) with Plasticity Index of Marine Stabilizes Soil with Cockle Shell Powder," Journal of Physics: Conference Series, vol. 1349, article no. 012036, 2019.

[12] S. K. Alam, A. Mondal, and A. Shiuly, "Prediction of CBR Value of Fine Grained Soils of Bengal Basin by Genetic Expression Programming, Artificial Neural Network and Krigging Method," Journal of the Geological Society of India, vol. 95, no. 2, pp.190-196, February 2020.

[13] Z. S. Janjua and J. Chand, "Correlation of CBR with Index Properties of Soil," International Journal of Civil Engineering and Technology, vol. 7, no. 5, pp. 57-62, September-October 2016.

[14] B. T. Nguyen and A. Mohajerani, "Determination of CBR for Fine-Grained Soils Using a Dynamic Lightweight Cone Penetrometer," International Journal of Pavement Engineering, vol. 16, no. 2, pp. 180-189, 2014.

[15] A. Rehman, K. Farooq, and H. Mujtaba, "Prediction of California Bearing Ratio (CBR) and Compaction Characteristics of Granular Soil," Acta Geotechnica Slovenica, vol. 14, no. 1, pp. 63-72, 2017.

[16] K. A. Rashed, N. B. Salih, and T. A. Abdalla, "Prediction of California Bearing Ratio from Consistency and Compaction Characteristics of Fine-grained Soils," Al-Nahrain Journal for Engineering Sciences, vol. 24, no. 2, pp. 123-129, December 2021.

[17] D. Abdella, T. Abebe, and E. T. Quezon, "Regression Analysis of Index Properties of Soil as Strength Determinant for California Bearing Ratio (CBR)," Global Scientific Journals, vol. 5, no. 6, pp. 1-12, June 2017.

[18] V. Y. Katte, S. M. Mfoyet, B. Manefouet, A. S. L. Wouatong, and L. A. Bezeng, "Correlation of California Bearing Ratio (CBR) Value with Soil Properties of Road Subgrade Soil," Geotechnical and Geological Engineering, vol. 37, no. 1, pp. 217-234, January 2019.
[19] S. K. Das and A. K. Sabat, "Using Neural Networks for Prediction of Some Properties of Fly Ash," Electronic Journal of Geotechnical Engineering, vol. 13, pp. 1-14, 2008.
[20] K. Zabielska-Adamska and M. J. Sulewska, "Neural Modelling of CBR Values for Compacted Fly Ash," Proceedings of the 17th International Conference on Soil Mechanics and Geotechnical Engineering, vol. 1, 2, 3, and 4, pp. 781-784, 2009.
[21] T. Taskiran, "Prediction of California Bearing Ratio (CBR) of Fine Grained Soils by AI Methods," Advances in Engineering Software, vol. 41, no. 6, pp. 886-892, June 2010.
[22] E. Bisong, "Matplotlib and Seaborn," Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners, pp. 151-165, 2019.
[23] S. Mukherjee and P. Ghosh, "Soil Behavior and Characterization: Effect of Improvement in CBR Characteristics of Soil Subgrade on Design of Bituminous Pavements," Indian Geotechnical Journal, vol. 51, no. 3, pp. 567-582, June 2021.
[24] T. B. Arnold, "KerasR: R Interface to the Keras Deep Learning Library," Journal of Open Source Software, vol. 2, no. 14, article no. 296, 2017.
[25] J. Hao and T. K. Ho, "Machine Learning Made Easy: A Review of Scikit-Learn Package in Python Programming Language," Journal of Educational and Behavioral Statistics, vol. 44, no. 3, pp. 348-361, June 2019.
[26] H. G. Park and K. R. Kim, "A Study of Correlation Between Soil Characteristic and CBR Value by Experimented Method," Journal of the Korean GEO-environmental Society, vol. 4, no. 1, pp. 41-48, 2003.
[27] T. F. Kurnaz and Y. Kaya, "Prediction of the California Bearing Ratio (CBR) of Compacted Soils by Using GMDH-Type Neural Network," The European Physical Journal Plus, vol. 134, no. 7, article no. 326, July 2019.