DOI: 10.21917/ijsc.2022.0368

DETERMINATION OF THE OPTIMUM PERFORMANCE AI MODEL AND METHODOLOGY TO PREDICT THE COMPACTION PARAMETERS OF SOILS

Jitendra Khatti and Kamaldeep Singh Grover
Department of Department of Civil Engineering, Rajasthan Technical University, India

Abstract

This technical article helps identify the optimum performance AI model for predicting compaction parameters of soil. A comparative study is mapped between regression analysis (RA), Gaussian process regression (GPR), decision tree (DT), support vector machine (SVM), and artificial neural networks (ANNs) approaches using 59 soil datasets. The soil dataset consists of soil properties such as gravel content, silt content, sand content, specific gravity, clay content, plasticity index, and liquid limit. The soil properties are used as input parameters to develop the AI model to predict soil optimum moisture content and maximum dry density. The RA, GPR, SVM, DT, and ANN models are designated as MLR_X, GPR_X, SVM_X, DT_X, ANN_X, where the X is OMC and MDD. The performance of MLR_OMC, GPR_OMC, SVM_OMC, DT_OMC, LMNN_OMC, and GDANN_OMC is 0.9714, 0.9867, 0.9689, 0.9832, 0.9435, and 0.9520, respectively. Similarly, the performance of MLR_MDD, GPR_MDD, SVM_MDD, DT_MDD, LMNN_MDD, and GDANN_MDD is 0.9512, 0.9854, 0.9482, 0.9199, 0.8679, and 0.9395, respectively. Based on the performance of AI models, the GPR_OMC and GPR_MDD models are identified as the optimum performance model to predict the soil maximum dry density (MDD) and optimum moisture content (OMC). The predicted OMC and MDD are compared with laboratory OMC and MDD, and it is found that the GPR_OMC and GPR_MDD model has the potential to predict soil compaction parameters.

Keywords:

Regression Analysis, Gaussian Process Regression, Support Vector Machine, Artificial Neural Network, Compaction Parameters of Soil

1. INTRODUCTION

Every soil has a different index, compaction, and strength parameters. The gravel content, sand content, fine content, and consistency limits are index parameters of soil. The consistency limits are determined experimentally using Casagrande or Cone penetration apparatus. Furthermore, the moisture content, dry density, and CBR are the compaction parameters of soil. The standard and modified proctors are light and heavy compaction apparatus used to determine soil moisture content and dry density [8]. The compaction parameters are affected by the size and shape of the soil particles. The laboratory procedures for determining MDD and OMC are a time-consuming and cumbersome task. Therefore, various investigators, researchers, and scientists developed and introduced different methods and methodologies to compute the compaction parameters of soil.

Han-Lin Wang et al. [12] evolved the multi expression programming (MEP) AI models to predict the OMC and MDD of soil. The authors reported that the performance of the MEP model of OMC and MDD was 0.916 (R = 0.9571) and 0.872 (R = 0.9338), respectively. The authors concluded that the evolved model can predict the compaction parameters. It was also concluded that the four physical parameters of soil, such as fine content, liquid limit, plastic limit, and compaction energy, play an important role in predicting compaction parameters.

Hasnat et al. [7] proposed the support vector machine AI models to estimate the compaction parameters of the soil. The authors concluded the following point, (i) the maximum dry density is inversely proportional to LL and LL is directly proportional to OMC, (ii) the plastic limit is not much strongly related with OMC and MDD in the comparison of LL, (iii) the liquid limit and plastic limit give better results for both OMC and MDD. The performance of OMC and MDD model of SVM was 0.86 (R = 0.9274) and 0.91 (R = 0.9539), respectively.

Verma et al. [11] wrote a review article on predicting the compaction parameters of soil. It was suggested that improve the dataset for better performance of existing models. The study concluded that soft computing is the trustworthy methodology and outperformed the statistical techniques in predicting compaction parameters and solving the other geotechnical problems.

Gunaydin et al. [18] suggested an SVM regression model to predict the compaction parameters of soil. The performance of the SVM model of OMC and MDD was 0.917 and 0.892, respectively. The authors concluded that the proposed model may be useful for the preliminary prediction of OMC and MDD of soil.

Salahudeen et al. [1] evolved artificial neural network models to predict the compaction characteristics of stabilized BCS with cement kiln dust. The performance of the 10-5-1 OMC and 10-7-1 MDD models was 0.8855 and 0.9754, respectively. The authors concluded that models' performance is satisfactory and strongly correlated with actual and predicted OMC and MDD values.

Ardakani et al. [4] proposed a GMDH type neural network and genetic algorithm AI models to predict the compaction parameters of soil. The GMDH is a type of neural network which stands for Group Method of Data Handling. The authors conducted the study using 212 datasets in the published research work. The performance of the GMDH model of OMC and MDD was 0.96 and 0.93, respectively. The authors finally concluded that the proposed model may be used to predict soil compaction parameters.

George et al. [6] reported that the maximum dry density may be predicted using genetic algorithms. The authors used 200 case histories from different sources in Kerala. It was concluded that the genetic algorithm gives a reliable result, and it can be used to predict the maximum dry density of soil. The performance of genetic algorithm-based models was 0.9197 (1000; 1000), 0.5311 (1000; 500), and 0.3803 (1000; 100).

Suman et al. [20] suggested different AI approaches to predict the maximum dry density and UCS of cement stabilized soil. The BRNN, DENN, LMNN, SVM, FN, MARS, and MLR models were developed and predicted MDD with the performance of 0.84, 0.88, 0.76, 0.93, 0.92, 0.94, and 0.75, respectively. The maximum performance was 0.94 determined for MARS models. The authors concluded that the model can be used for the initial trial of the different mixtures.

Ranasinghe et al. (2016) and A. K. Shrivastava et al. [2] reported that the artificial intelligence approaches have the potential to predict the compaction parameters.

Jayan et al. [14] evolved the AI models to predict soil compaction parameters. The authors reported that the performance of the OMC and MDD model of ANN was 0.91 and 0.92, respectively. The study was carried out using 180 plus laboratory test data. The authors concluded that the ANN models have the ability to predict the OMC and MDD of the soil.

Khuntia et al. [22] developed a MARS model to predict the compaction parameters of soil. The performance of MARS models of OMC and MDD was 0.88 and 0.81, respectively.

Smith et al. [21] evolved a multilinear regression model to investigate the relationship between soil compaction parameters and index properties. The authors suggested that MDD was well correlated with OMC. The OMC and MDD were best correlated with PI compared to LL and PL. The author concluded that the MLR model may be used to predict soil compaction parameters.

Majidi et al. [3] developed multilayer perceptron class-based neural network models to predict marl soil compaction parameters. The performance of model OM-2H was 0.97 (R = 0.9848) reported, comparatively higher than other models.

Sivrikaya et al. [19] evolved the MLR and GEP model to predict coarse and fine soil compaction parameters. The performance of MLR-SMP-1, MLR-SMP-2, MLR-eq7, GEP-1, GEP-2, MLR-SMP-3, MLR-SMP-4, MLR-eq11, GEP-3, and GEP-4 was 0.92, 0.94, 0.94, 0.93, 0.95, -0.44, 0.94, 0.97, 0.98, and 0.95, respectively.

Khattab et al. [9] proposed ANN models to predict the compaction parameters of soil. The performance of MDD and OMC ANN models was 0.905 and 0.932, respectively. The author concluded that the ANN is able to predict the compaction parameters, but the prediction accuracy is also affected by the range of the data. Hossein et al. [5] also concluded that the artificial intelligence approaches may be used to predict the compaction parameters of soil.

1.1 PROBLEM STATEMENT

The literature survey shows that artificial intelligence is a powerful tool for engineering. It helps to predict or compute the preliminary value of engineering properties of soil with high accuracy. But the best method and methodology are still doubtable. Therefore, the present study has been conducted to determine the optimum performance AI model and methodology to predict soil compaction parameters.

2. DETAILS OF AI MODELS

The compaction parameters of soil have been predicted using regression analysis, Gaussian process regression, support vector machine, decision tree, and artificial neural network AI approaches in the present study. The details of AI models are given below.

2.1 REGRESSION MODEL

Regression analysis is the most traditional method of prediction or forecasting of data. The regression analysis is the two types, i.e., Simple Regression and Multiple Regression. The regression analysis may be carried out using linear and non-linear methods. The most popular regression analysis is linear regression analysis. The multilinear regression analysis model has been developed to predict soil compaction parameters. The index properties and specific gravity have been used as input parameters. An equation has been derived to predict the soil OMC and MDD by MLR_OMC and MLR_MDD models. The equations for predicting the OMC and MDD are:

$$MLR\_OMC = 9.63-0.008*G-0.092*S+0.0117*M-0.011*C-0.181*SG + 0.2837*LL + 0.0364*PI \quad (1)$$

$$MLR\_MDD=2.01+0.0008*G + 0.0034*S-0.0003*M+0.0002*C-0.081*SG-0.007*LL + 0.0035*PI \quad (2)$$

where G is gravel content (in %), S is sand content (in %), M is silt content (in %), C is clay content (in %), SG is specific gravity, LL is the liquid limit (in %), and PI is the plasticity index (in %). Eq.(1) and Eq.(2) have been used to predict soil optimum moisture content and maximum dry density of soils. The MLR_OMC and MLR_MDD models have been developed using the Data Analysis Tool of Microsoft Excel 2019.

2.2 GAUSSIAN PROCESS REGRESSION MODELS

Gaussian process regression is a stochastic process and Bayesian approach to regression in machine learning. The advantage of GPR is it works well on small datasets having good prediction accuracy. The Gaussian process is based on the kernel function, i.e., linear, squared exponential, exponential, Matern, periodic, and rational quadratic kernels. The kernels are the mathematical formulation or algorithm, and the mathematical formula is shown below:

$$\text{Linear: } K(x,x') = x^T x' \quad (3)$$

$$\text{Squared Exponential: } K(x,x') = \exp\left(-\frac{|d|^2}{2l^2}\right) \quad (4)$$

$$\text{Exponential } K(x,x') = \exp\left(-\frac{|d|}{l}\right) \quad (5)$$

$$\text{Matern } K(x,x') = \frac{2^{1-v}}{\Gamma(v)}\left(\frac{\sqrt{2v}|d|}{l}\right)^{v} k_v\left(\frac{\sqrt{2v}|d|}{l}\right) \quad (6)$$

$$\text{Periodic } K(x,x') = \exp\left(\frac{2\sin(0.5d)}{l^2}\right) \quad (7)$$

$$\text{Rational Quadratic } K(x,x') = \left(1+|d|^2\right)^{-\alpha}, \alpha \geq 0 \quad (8)$$

The Gaussian process regression hyperparameters to optimize the model and optimizer configurations are given in Table.1 and Table.2, respectively.

Table.1. GPR Hyperparameters to Optimize

Parameters	Value/Condition
Basic function	Auto
Kernel function	Auto
Kernel scale	Auto
Signal standard deviation	3.5431484
Sigma	Auto
Standardize	Enable
Optimize numeric parameters	Enable

Table.2. GPR Optimizer Configurations

Parameters	Condition
Optimizer	Bayesian Optimization
Acquisition function	Expected Improved per second plus
Iterations	30
Maximum training time (s)	300
Number of grid division	10

The GPR_OMC and GPR_MDD models have been developed using the Regression Learner Tool of MATLAB R2020a.

2.3 SUPPORT VECTOR MACHINE MODELS

The support vector machine was introduced to analyze data for regression analysis and classification problems by Vladimir Vapnik. The support vector machine is categorized under supervised learning. SVM is based on kernel functions, namely Gaussian, linear, quadratic, and cubic. The mathematical formulation of the kernel functions is:

$$K(\bar{x}) = \begin{cases} 1 & if \ \|\bar{x}\| \le 1 \\ 0 & otherwise \end{cases} \tag{9}$$

$$\text{Polynomial Kernel } K(x_i, x_j) = (x_i \cdot x_j + 1)^d \tag{10}$$

$$\text{Gaussian Kernel } K(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{11}$$

$$\text{Gaussian RBF } K(x_i, x_j) = \exp\left(-\gamma \|x_i, x_j\|^2\right) \tag{12}$$

$$\text{Laplace RBF Kernel } K(x, y) = \exp\left(-\frac{\|x-y\|}{\sigma}\right) \tag{13}$$

$$\text{Hyperbolic Tangent Kernel } K(x_i, x_j) = \tanh(kx_i \cdot x_j + c) \tag{14}$$

$$\text{Sigmoid Kernel: } K(x, y) = \tanh(\alpha x^T y + c) \tag{15}$$

The support vector machine hyperparameters to optimize the model and optimizer configurations are given in Table.3 and Table.4.

Table.3. SVM Hyperparameters to Optimize

Parameters	Value/ Condition
Kernel function	Auto
Box constraint	Auto (4.763)
Kernel scale	Auto
Epsilon	Auto (0.476)
Standardize data	Enable

Table.4. SVM Optimizer Configurations

Parameters	Condition
Optimizer	Bayesian Optimization
Acquisition function	Expected Improved/per second plus
Iterations	30
Maximum training time (s)	300
Number of grid division	10

The Regression learner tool of MATLAB R2020a has been used to develop SVM_OMC and SVM_MDD models.

2.4 DECISION TREE MODELS

A flowchart-like structure having nodes, branches, and leaves to solve the operations research and operations management problem is known as a decision tree. The flowchart presents the classification rules from root to leaf. The decision, chance, and end nodes are the main components of any decision tree [15].

The decision tree (DT) hyperparameters to optimize the model and optimizer options are given in Table.5 and Table.6.

Table.5. DT Hyperparameters to Optimize

Parameters	Value/Condition
Minimum leaf size	Auto
Surrogate decision splits	Off
Maximum surrogate per node	10

Table.6. DT Optimizer Configurations

Parameters	Condition
Optimizer	Bayesian Optimization
Acquisition function	Expected Improved per second plus
Iterations	30
Maximum training time (s)	300
Number of grid division	10

The Regression Learner Tool of MATLAB R2020a has been used to develop DT_OMC and DT_MDD models.

2.5 ARTIFICIAL NEURAL NETWORK MODELS

A computational model was created for the neural networks [16] by Warren McCullouch and Walter Pitts [17] in 1943. In the late 1940s, Donald Hebb [13] introduced Hebbian learning in 1949. For the simulation of the Hebbian network, a computational machine was introduced by Farley and Clark [10]. With time, neural networking was improved, and applications were developed in different fields. The artificial neural network is a network of layers. These layers are the input layer, hidden layers, and output layer. These layers are interconnected with neurons, and each neuron has its weight.

The neural network has two processes, i.e., feed-forward and backpropagation. In the feed-forward process, the information transfers from the input layer to the output layer, and in backpropagation, the information transfers from the output layer to the input layer. First, the error is calculated at the output layer after the feed-forward process. Then, the calculated error is distributed to the neurons, and the weight of neurons is updated in the backpropagation process. For the development of the artificial neural network, the selected hyperparameters of the artificial neural network model are shown in Table.7.

Table.7. ANN Hyperparameters to Optimize

Parameters	Value/ Condition
Hidden layer(s)	Single hidden layer
Neurons	10
Backpropagation algorithms	LM, GDA
Normalizing function	Min-max, Log function
Activation function	Sigmoid, linear
Train: Validation ratio	70%: 30%
Epochs	1000
Network type	Feed-forward backpropagation
Network class	MLP
Mu	0.001
Max fail	6
Min gradient	$10e^{-7}$

The LMNN_OMC, GDANN_OMC, LMNN_MDD, and GDANN_MDD models have been developed in MATLAB R2020a. The MLR_OMC, GPR_OMC, SVM_OMC, DT_OMC, LMNN_OMC, and GDANN_OMC models have been used to predict the optimum moisture content. Thus, the MLR_MDD, GPR_MDD, GPR_MDD, DT_MDD, LMNN_MDD, and GDANN_MDD models have been used to predict the maximum dry density of the soil.

3. DATA STATISTICS

Fifty-nine soil datasets including specific gravity (SG), liquid limit (LL), plasticity index (PI), optimum moisture content (OMC), maximum dry density (MDD), gravel content (G), sand content (S), silt content (M), and clay content (C) have been used to develop the AI models in this study. The statistics of the soil datasets are given in Table.8.

Table.8. Statistics of Soil Datasets

	Min	Max	Mean	Median	St Dev	CL
G (%)	0.0	74.0	12.1	1.7	17.8	4.64
S (%)	0.0	100	54.5	60.9	31.1	8.10
M (%)	0.0	82.0	10.7	7.0	15.5	4.0
C (%)	0.0	87.0	17.8	8.4	21.4	5.6
SG	2.5	2.8	2.6	2.69	0.05	0.01
LL (%)	25.6	54.2	33.1	28.9	8.7	2.3
PI (%)	9.0	30.9	15.1	13.9	5.0	1.3
OMC (%)	7.6	24.7	13.9	12.1	5.2	1.4
MDD (gm/cc)	1.5	2.0	1.83	1.8	0.1	0.0

The correlation coefficient more than 0.8, between 0.8-0.2, and less than 0.2 show the strong, good, and weak relationship between independent and dependent parameters. The correlation between the input and output parameters has been identified by drawing the Pearson correlation matrix, as given in Table.9. The correlation between input parameters such as LL, PI, G, S, M, C, SG, and output parameters such as OMC and MDD is shown in Fig.1 and Fig.2, respectively.

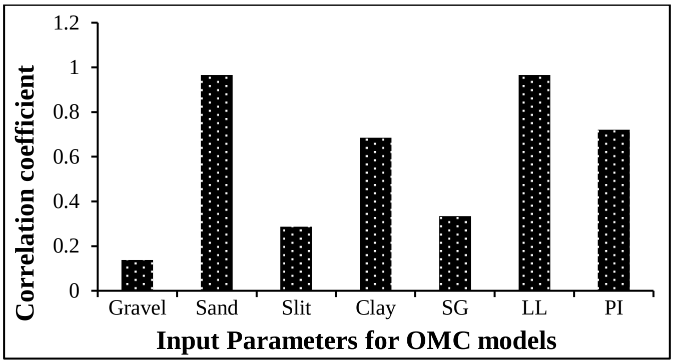

Fig.1. Correlation Coefficient for OMC

Table.9. Pearson correlation matrix for soil properties

Para-meters	G (%)	S (%)	M (%)	C (%)	SG (%)	LL (%)	PI (%)	OMC (%)	MDD (gm/cc)
G (%)	1.000	-	-	-	-	-	-	-	-
S (%)	0.267	1.000	-	-	-	-	-	-	-
M (%)	0.228	0.252	1.000	-	-	-	-	-	-
C (%)	0.169	0.709	0.107	1.000	-	-	-	-	-
SG	0.041	0.334	0.300	0.250	1.000	-	-	-	-
LL (%)	0.079	0.916	0.242	0.664	0.292	1.000	-	-	-
PI (%)	0.635	0.755	0.028	0.318	0.139	0.729	1.000	-	-
OMC (%)	0.138	0.966	0.287	0.685	0.334	0.966	0.721	1.000	-
MDD (gm/cc)	0.024	0.932	0.327	0.738	0.340	0.932	0.582	0.972	1.000

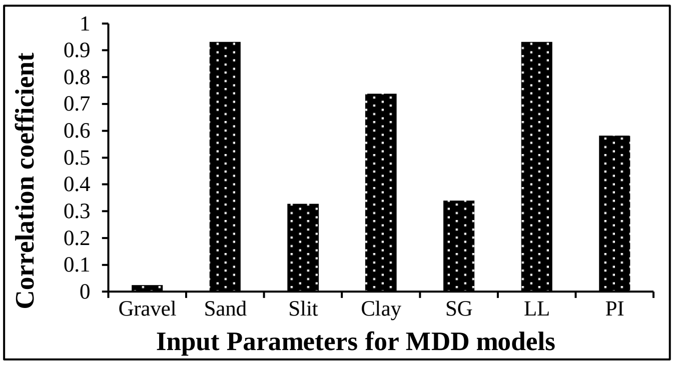

Fig.2. Correlation Coefficient for MDD

The Fig.1 and Fig.2 depict the relationship between OMC and MDD with input parameters. The sand and liquid limit are strongly correlated with compaction parameters. The silt, clay, specific gravity, and plasticity index correlate with compaction parameters, but the gravel content has a weak relationship.

3.1 TRAINING AND TESTING DATASETS

The multilinear, Gaussian process, decision tree, support vector machine, and artificial neural network models have been developed to predict soil OMC and MDD. The MLR, GPR, SVM, DT are trained and tested by 41 and 18 datasets of soil.

The 41 datasets of soil have been sub-divided into 28 and 13 datasets in artificial neural networks to train and validate the artificial neural network models. Eighteen datasets of soil have also tested the artificial neural network models.

4. RESULTS AND DISCUSSIONS

The performance and predicted results of OMC and MDD using multilinear regression, Gaussian process regression, support vector machine, decision tree, and artificial neural network models have been discussed in this section. In addition, MLR_OMC and MLR_MDD multiple regression models have been developed to predict OMC and MDD of soil. The performance of the MLR_OMC and MLR_MDD models is given in Table.10.

Table.10. Performance of MLR_OMC and MLR_MDD models

Parameters	Regression Models	
	MLR_OMC	MLR_MDD
Train-RMSE	0.4461	0.0213
Train-R	0.9959	0.9881
Train-MAE	0.3703	0.0155
Test-RMSE	0.9242	0.1507
Test-R	0.9714	0.9512
Test-MAE	1.8330	0.0023

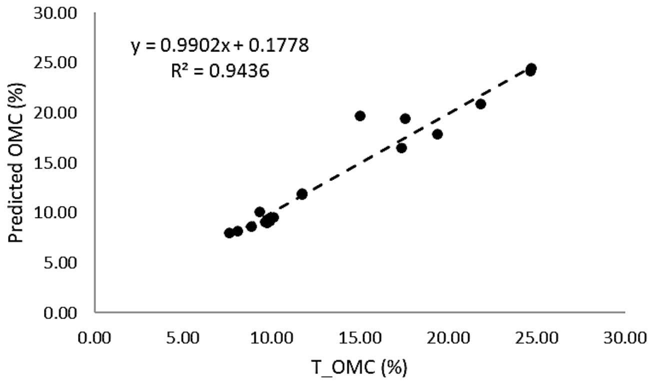

Fig.3. Actual vs Predicted Plot of OMC using MLR_OMC Model

The Table.10 shows that the performance of the MLR_OMC and MLR_MDD model is 0.9714 and 0.9512, respectively. The MLR_OMC model has predicted the OMC with the RMSE and MAE performance of 0.9242 and 1.8330, respectively. Similarly, the MLR_MDD model has predicted the MDD with RMSE and MAE performance of 0.1507 and 0.0023, respectively. The actual vs predicted plot of optimum moisture content (OMC) and maximum dry density (MDD) is drawn and shown in Fig.3 and Fig.4, respectively.

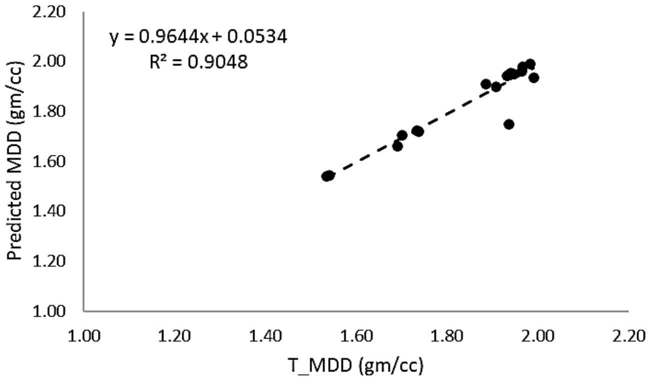

Fig.4. Actual vs Predicted Plot of MDD using MLR_MDD Model

The Fig.3 and Fig.4 show that the MLR_OMC and MLR_MDD models have predicted OMC and MDD with COD of 0.9436 and 0.9048. The Gaussian process regression approach has been applied to predict the OMC and MDD. The training performance curve for GPR_OMC and GPR_MDD model is shown in Fig.5 and Fig.6.

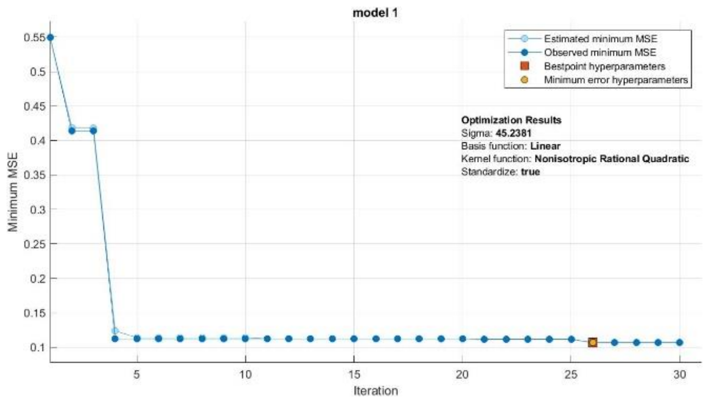

Fig.5. Performance Curve of GPR_OMC Model

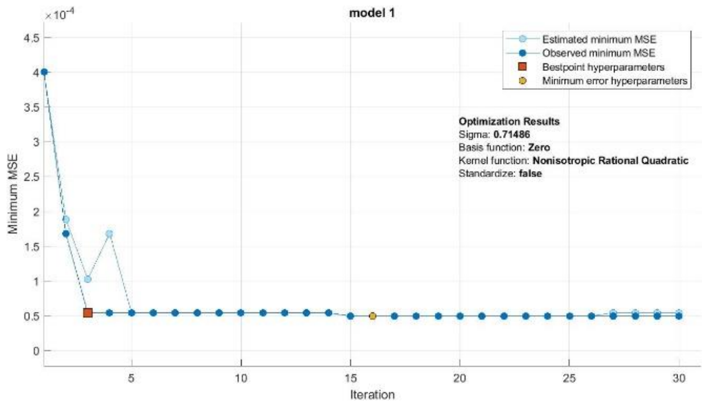

Fig.6. Performance Curve of GPR_MDD Model

The optimization results of GPR_OMC show that a better prediction of optimum moisture content may be obtained by developing the GPR_OMC model having a basic linear function, the Nonisotropic rational quadratic function. On the other hand, the optimization results of GPR_MDD show that a better prediction of optimum moisture content may be obtained by developing the GPR_OMC model having a zero-basic function, the Nonisotropic rational quadratic function.

From the performance curve of GPR_OMC and GPR_MDD, it has been observed that the value of maximum dry density is comparatively significantly less. Hence, the dataset of soil has been standardized in the GPR_OMC model. The performance of GPR_OMC and GPR_MDD models is given in Table.11.

Table.11. Performance of GPR_OMC and GPR_MDD models

Parameters	**Gaussian Regression Models**	
	GPR_OMC	**GPR_MDD**
Train-RMSE	0.3266	0.0074
Train-R	1.0000	1.0000
Train-MAE	0.2271	0.0028
Test-RMSE	0.6295	0.0913
Test-R	0.9867	0.9854
Test-MAE	0.8245	0.0007

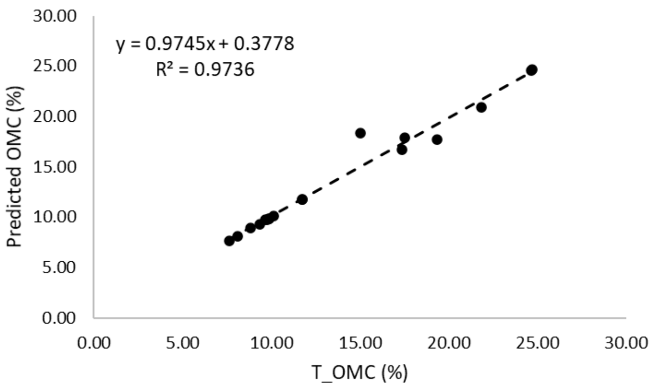

Fig.7. Actual vs Predicted Plot of OMC using GPR_OMC Model

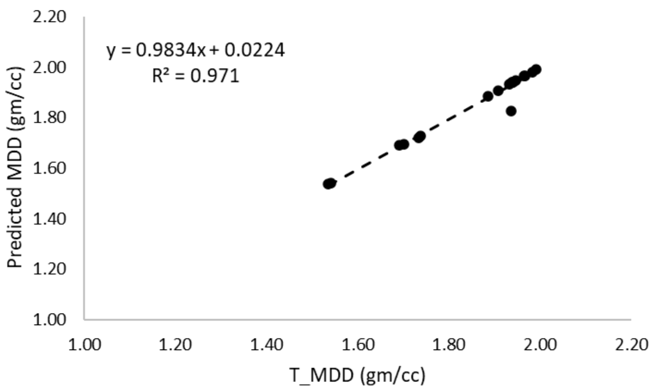

Fig.8. Actual vs Predicted Plot of MDD using GPR_MDD Model

The Table.11 shows that the GPR_OMC and GPR_MDD model has the testing performance of 0.9867 and 0.9854, respectively. The GPR_OMC model has predicted the OMC with RMSE and MAE performance of 0.6295 and 0.8245, respectively. Thus, the GPR_MDD model has predicted the MDD with RMSE and MAE performance of 0.0913 and 0.0007, respectively. The actual vs predicted plot of OMC and MDD is drawn and shown in Fig.7 and Fig.8, respectively.

The Fig.7 and Fig.8 show that the GPR_OMC and GPR_MDD models have predicted OMC and MDD with COD of 0.9736 and 0.971, respectively. The coefficient of determination indicates that the proposed GPR models have the capability to predict the OMC and MDD with the least prediction error.

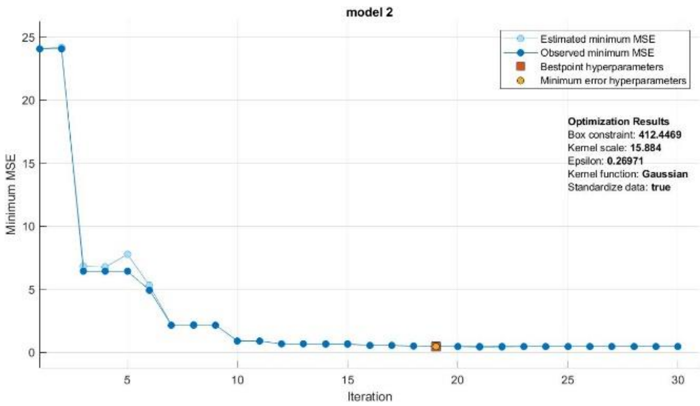

Fig.9. Performance Curve of SVM_OMC Model

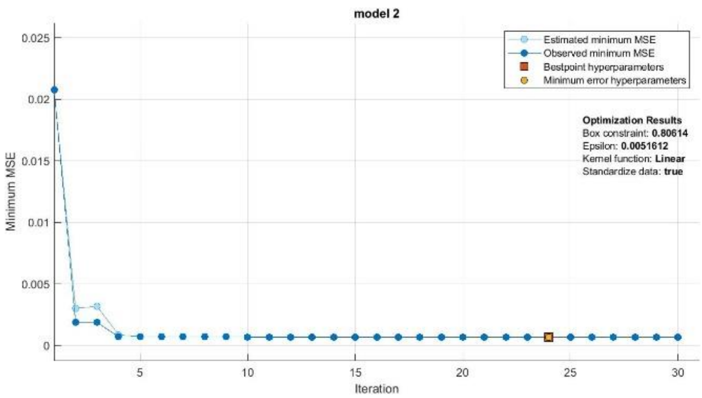

Fig.10. Performance Curve of SVM_MDD Model

Table.12. Performance of SVM_OMC and SVM_MDD models

Parameters	**Support Vector Machine Models**	
	SVM_OMC	**SVM_MDD**
Train-RMSE	0.6910	0.0256
Train-R	0.9899	0.9849
Train-MAE	0.8205	0.0185
Test-RMSE	0.9174	0.1488
Test-R	0.9689	0.9482
Test-MAE	1.9683	0.0025

Artificial intelligence support vector machine approach has also been used to predict soil OMC and MDD. Fig.9 and Fig.10 show the training performance curve of the SVM_OMC and SVM_MDD models. The optimization results of SVM_OMC show that a better prediction of optimum moisture content may be obtained by developing the SVM_OMC model having the Gaussian kernel function. The optimization results of SVM_MDD show that a better prediction of optimum moisture content may be obtained by developing the SVM_OMC model having a Linear kernel function. The performance curve of SVM_OMC and SVM_MDD shows that the optimum moisture content (OMC) and maximum dry density (MDD) are standardized during the training of SVM_OMC and SVM_MDD models. Therefore, the performance of SVM_OMC and SVM_MDD models is given in Table.12.

The Table.12 shows that the performance of the SVM_OMC and SVM_MDD model is 0.9689 and 0.9482, respectively. Therefore, the SVM_OMC model has predicted the OMC of the soil with 0.9174 RMSE and 1.9683 MAE. Similarly, the SVM_MDD model has predicted the MDD of the soil with 0.1488 RMSE and 0.0025 MAE. Finally, the actual vs predicted plot of OMC and MDD is drawn and shown in Fig.11 and Fig.12, respectively.

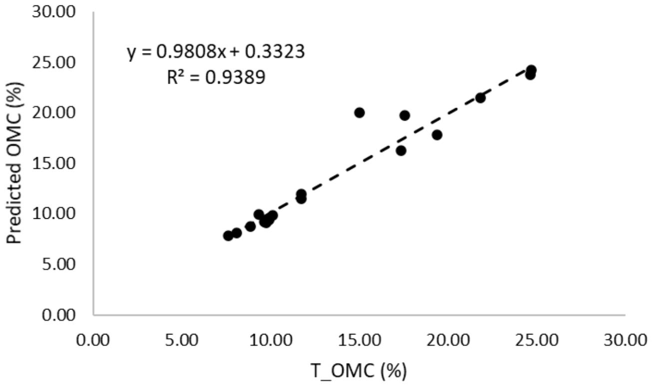

Fig.11. Actual vs Predicted Plot of OMC using SVM_OMC Model

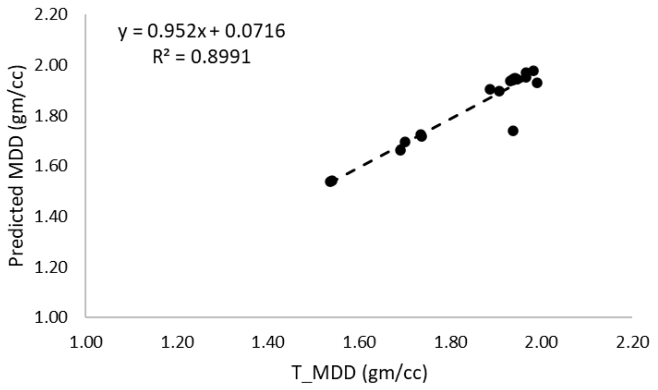

Fig.12. Actual vs Predicted Plot of MDD using SVM_MDD Model

The Fig.11 and Fig.12 show that the SVM_OMC and SVM_MDD models have COD of 0.9389 and 0.8991, respectively. Another AI approach, named decision tree, has also been used in this study. The training performance curve of the DT_OMC and DT_MDD model is shown in Figures 13 and 14, respectively. The optimization results showed that the DT_OMC and DT_MDD models can perform better by developing tree models with three and two leaves. The performance of DT_OMC and DT_MDD models is given in Table.13.

Table.13. Performance of DT_OMC and DT_MDD models

Parameters	Decision Tree Models	
	DT_OMC	DT_MDD
Train-RMSE	1.2306	0.0387
Train-R	0.9695	0.9644
Train-MAE	0.9115	0.0311
Test-RMSE	0.8535	0.1850
Test-R	0.9832	0.9199
Test-MAE	1.0674	0.0036

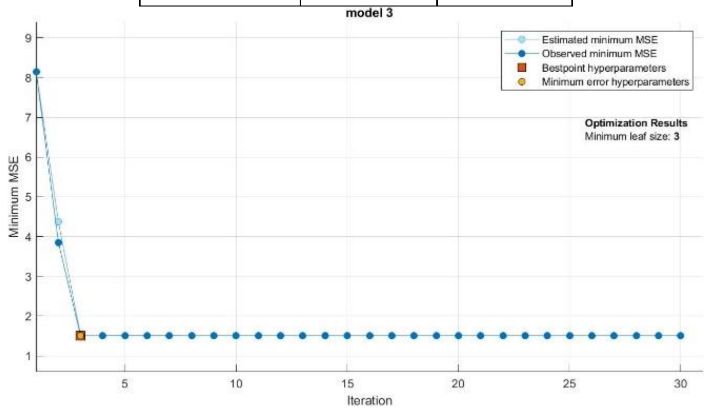

Fig.13. Performance Curve of DT_OMC Model

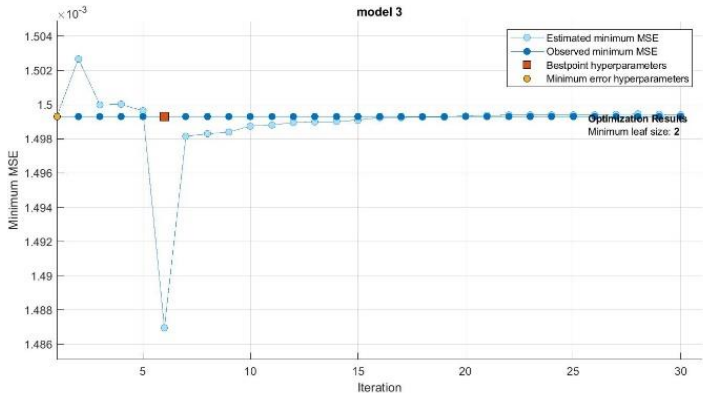

Fig.14. Performance Curve of DT_MDD Model

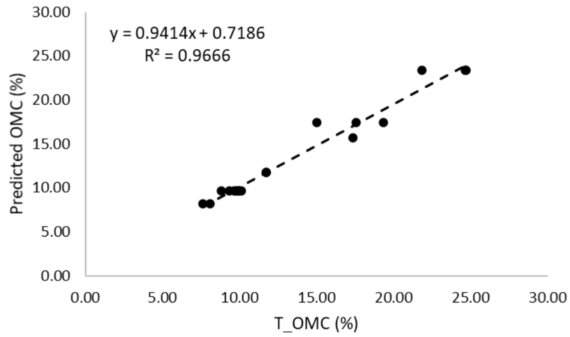

Fig.15. Actual vs Predicted Plot of OMC using DT_OMC Model

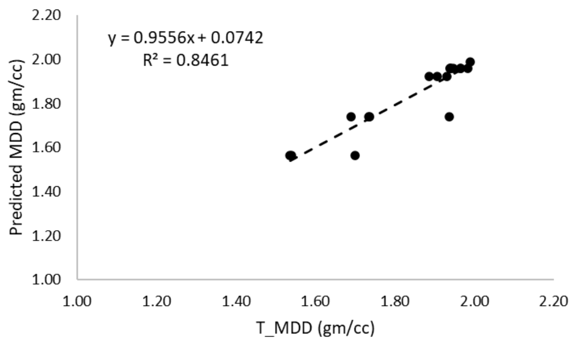

Fig.16. Actual vs Predicted Plot of MDD using DT_MDD Model

The Table.13 shows that the testing performance of the DT_OMC and DT_MDD model is 0.9832 and 0.9199, respectively. Finally, the actual vs predicted plot of OMC and MDD is drawn and shown in Fig.15 and Fig.16, respectively.

The Fig.15 and Fig.16 show that the DT_OMC and DT_MDD models have COD of 0.9666 and 0.8461, respectively. An artificial neural network approach has also been used to predict OMC and MDD. The training performance plot of LMNN_OMC and LMNN_MDD is shown in Fig.17 and Fig.18.

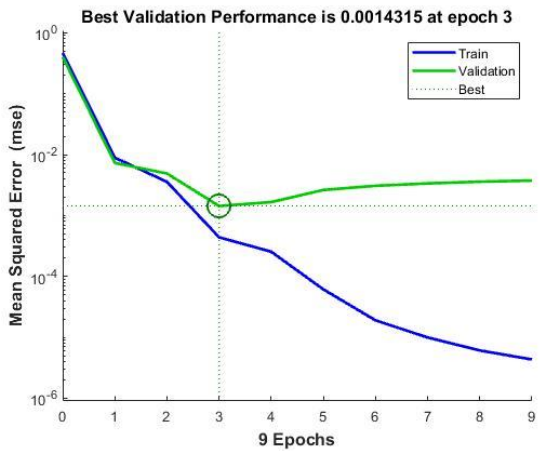

Fig.17. Performance Curve of LMNN_OMC Model

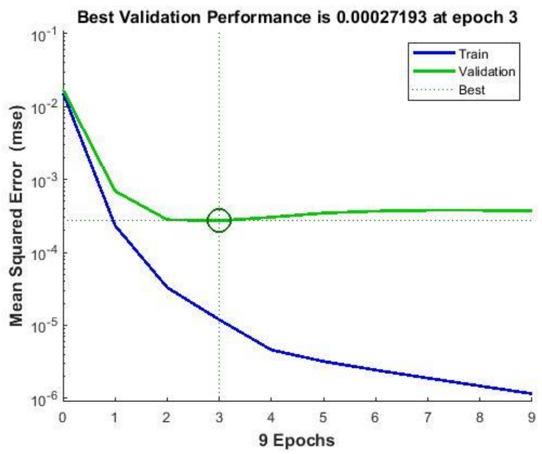

Fig.18. Performance Curve of LMNN_MDD Model

Table.14. Performance of LMNN_OMC and LMNN_MDD models

Parameters	LM Neural Network Models	
	LMNN_OMC	LMNN_MDD
Train-RMSE	0.0210	0.0035
Train-R	0.9976	0.9977
Train-MAE	0.0479	0.0016
Test-RMSE	1.0512	0.2161
Test-R	0.9435	0.8679
Test-MAE	3.5910	0.0061

The Table.14 shows that the testing performance of the LMNN_OMC and LMNN_MDD model is 0.9435 and 0.8679, respectively. The LMNN_OMC model has predicted the OMC with RMSE and MAE performance of 1.0512 and 3.5910, respectively. Thus, the LMNN_MDD model has predicted the MDD with RMSE and MAE performance of 0.2161 and 0.0061, respectively. The actual vs predicted plot of OMC and MDD is drawn and shown in Fig.19 and Fig.20, respectively.

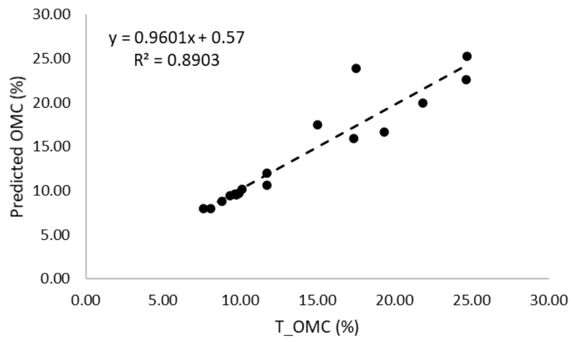

Fig.19. Actual vs Predicted Plot of OMC using LMNN_OMC Model

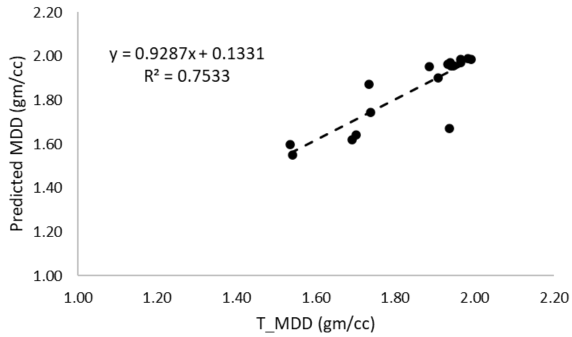

Fig.20. Actual vs Predicted Plot of MDD using LMNN_MDD Model

The Fig.19 and Fig.20 show that the LMNN_OMC and LMNN_MDD model can predict optimum moisture content and maximum dry density with COD of 0.8903 and 0.7533. Other artificial neural network models of OMC and MDD have been developed using gradient-descent with an adaptive learning algorithm. The performance of GDANN_OMC and GDANN_MDD models is given in Table.15.

Table.15. Performance of LMNN_OMC and LMNN_MDD models

Parameters	GDA Neural Network Models	
	GDANN_OMC	GDANN_MDD
Train-RMSE	0.0478	0.0088
Train-R	0.9889	0.9859
Train-MAE	0.0490	0.0135
Test-RMSE	1.1931	0.1854
Test-R	0.9520	0.9395
Test-MAE	3.6301	0.0031

The training performance curve of the GDANN_OMC and GDANN_MDD model is shown in Fig.21 and Fig.22, respectively.

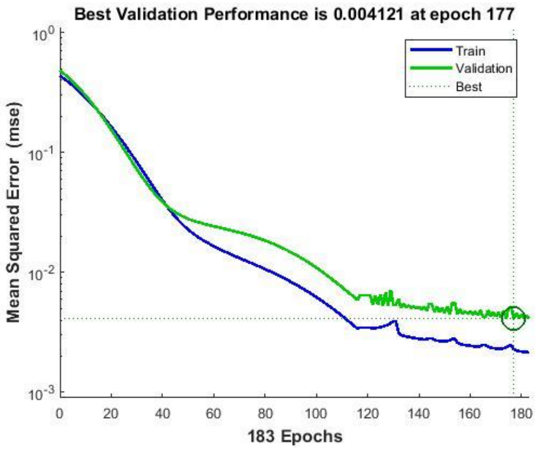

Fig.21. Performance Curve of GDANN_OMC Model

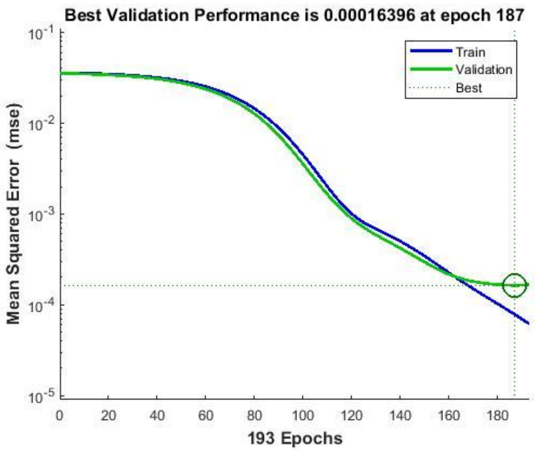

Fig.22. Performance Curve of GDANN_MDD Model

The Table.15 shows that the testing performance of the GDANN_OMC and GDANN_MDD model is 0.9520 and 0.9395, respectively. Finally, the actual vs predicted plot of OMC and MDD is drawn and shown in Figures 23 and 24, respectively.

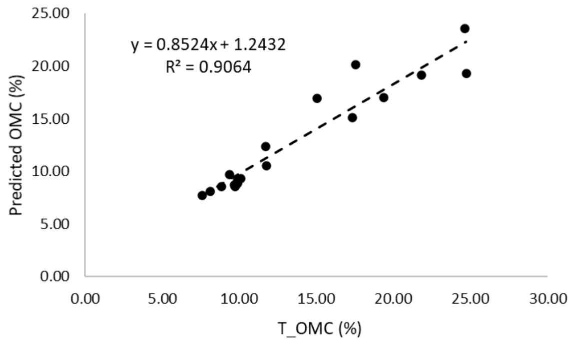

Fig.23. Actual vs Predicted Plot of OMC using GDANN_OMC Model

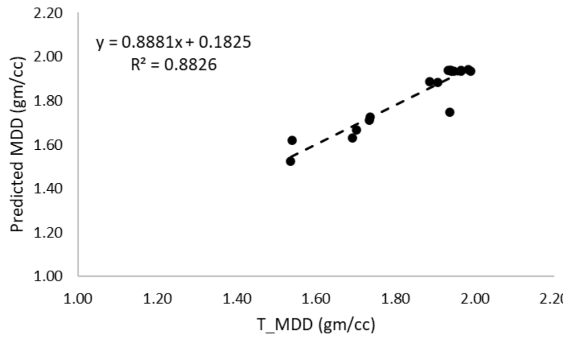

Fig.24. Actual vs Predicted Plot of MDD using GDANN_MDD Model

The Fig.23 and Fig.24 show that the GDANN_OMC and GDANN_MDD model has COD of 0.9064 and 0.8826. The optimum performance model and methodology are determined by comparing the performance of proposed OMC and MDD models, as shown in Fig.25 and Fig.26, respectively.

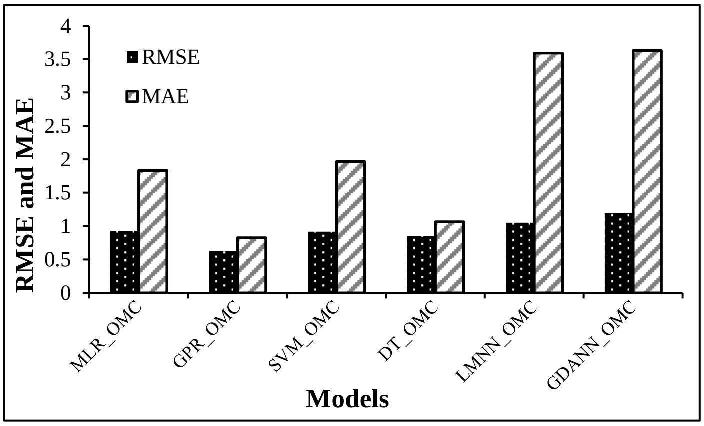

Fig.25. Comparison of RMSE and MAE Performance of AI Models of OMC

The Fig.25 and Fig.26 show that the GPR_OMC model has predicted the optimum moisture content (OMC) with high performance and significantly less prediction error (RMSE and MAE). Therefore, the GPR_OMC model has been identified as an optimum performance model for predicting OMC. Thus, the performance comparison of the AI model of MDD is shown in Fig.27 and Fig.28, respectively.

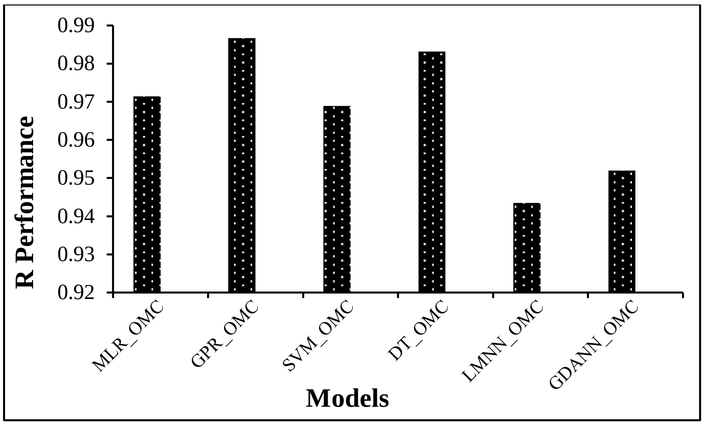

Fig.26. Comparison of R Performance of AI Models of OMC

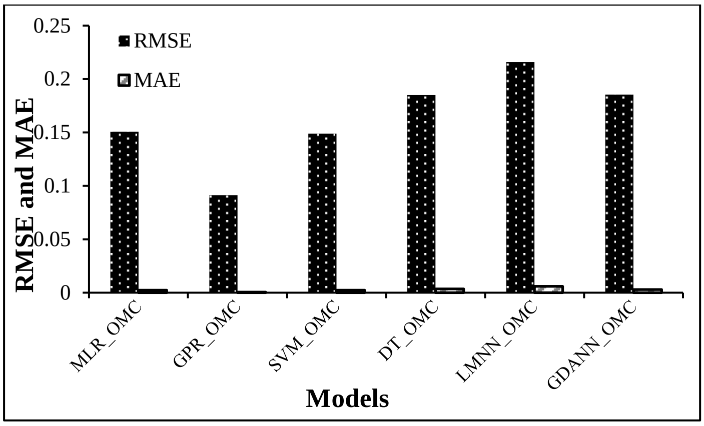

Fig.27. Comparison of RMSE and MAE Performance of AI Models of MDD

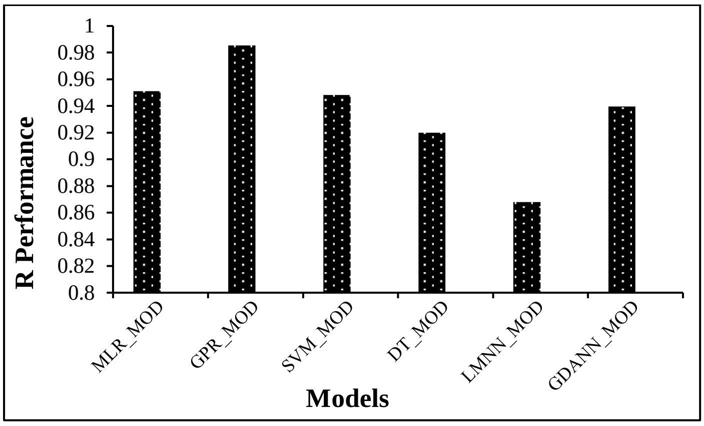

Fig.28. Comparison of R Performance of AI Models of MDD

The Fig.27 and Fig.28 show that the GPR_MDD model has predicted the maximum dry density (OMC) with high performance and significantly less prediction error (RMSE and MAE). Therefore, the GPR_MDD model has been identified as an optimum performance model for predicting MDD. Finally, the GPR_OMC and GPR_MDD models outperformed other proposed AI models in predicting OMC and MDD, respectively, with the testing performance of 0.9867 and 0.9854.

5. COMPARISON OF PERFORMANCE

The performance of GPR_OMC and GRP_MDD models has been compared with the performance of presently available AI models in the literature survey and shown in Table.16.

Table.16. Comparison of Performance for the prediction of Compaction Parameters

Author(s)	Performance	
	OMC	MDD
Present Study	0.9867	0.9854
Wang et al. [12]	0.9571	0.9338
Hasnat et al. [7]	0.9274	0.9539
Gunaydin et al. [18]	0.9170	0.8920
Salahudeen et al. [1]	0.8855	0.9754
Ardakani et al. [4]	0.9600	0.9300
Jeeja Jayan et al. [14]	0.9100	0.9200
Sunil Khuntia et al. [22]	0.8800	0.8100
Khattab et al. [9]	0.9320	0.9050

From Table.16, it has been observed that the proposed GPR model of OMC and MDD has also outperformed the previously published AI models in predicting the compaction parameters of soil. Therefore, the developed GPR_OMC and GPR_MDD models in the present study can be used to predict the OMC and MDD for published articles.

6. CONCLUSIONS

The literature survey shows that artificial intelligence approaches have the potential to predict soil properties. The multilinear regression (MLR), Gaussian process regression (GPR), support vector machine (SVM), decision tree (DT), and artificial neural network (ANN) models were developed to predict the compaction parameters of soil. The OMC and MDD models of Gaussian process regression outperformed other proposed AI models in the present study. In this study, the GPR_OMC and GPR_MDD were predicted OMC and MDD with COD of 0.9867 and 0.9854 (strong correlation between actual and prediction OMC and MDD; G.N. Smith, 1986). The performance comparison of the GPR_OMC and GPR_MDD model with literature survey models showed that the present models have the high capabilities to predict compaction parameters.

The artificial neural network models were developed using Levenberg-Marquardt (LM) and Gradient-Descent with an Adaptive (GDA) Learning backpropagation algorithm. It was observed that the GDA algorithm-based ANN models outperformed LM algorithm-based ANN models in predicting OMC and MDD of soil. An artificial neural network is an approach to deep learning, and it performs better on large datasets, but in the present research work, the datasets were limited. Finally, this research concludes that Gaussian process regression can predict the compaction parameters with high accuracy for a small dataset.

REFERENCES

[1] A.B. Salahudeen, T.S. Ijimdiya, A.O. Eberemu and K.J. Osinubi, “Artificial Neural Networks Prediction of Compaction Characteristics of Black Cotton Soil Stabilised with Cement Kiln Dust”, *Journal of Soft Computing in Civil Engineering*, Vol. 2, No. 3, pp. 50-71, 2018.

[2] A.K. Shrivastava and P.K. Jain, “Prediction of Compaction Parameters using Regression and ANN Tools”, *International Journal for Scientific Research and Development*, Vol. 3, No. 11, pp. 697-702, 2016.

[3] G. Majidi, M. Noruzi Nashlaji andd Y. Firouzei, “Estimating Compaction Parameters of Marl Soils using Mult-layer Perceptron Neural Network”, Vol. 20, No. 2, pp. 170-198, 2014.

[4] Alireza Ardakani and Afshin Kordnaeij, “Soil Compaction Parameters Prediction using GMDH-type Neural Network and Genetic Algorithm”, *European Journal of Environmental and Civil Engineering*, Vol. 23, No. 4, pp. 449-462, 2017.

[5] Amir Hossein Alavi, Amir Hossein Gandomi, Ali Mollahassani, Ali Akbar Heshmati, Azadeh Rashed, “Modelling of Maximum Dry Density and Optimum Moisture Content of Stabilized Soil using Artificial Neural Networks”, *Journal of Plant Nutrition and Soil Science*, 173, pp. 368-379, 2010.

[6] N.A. Christy Antony George and Soumya V. Krishnankutty, “Prediction of Maximum Dry Density of Soil using Genetic Algorithm”, *International Journal of Engineering Research and Technology*, Vol. 6, No. 3, pp. 550-552, 2017.

[7] Arif Hasnat, Md Mehedi Hasan, Md Rabiul Islam and Md Abdul Alim, “Prediction of Compaction Parameters of Soil using Support Vector Regression”, *Current Trend in Civil and Structural Engineering*, Vol. 4, No. 1, pp. 1-7, 2019.

[8] K.R. Arora, “*Soil Mechanics, and Foundation Engineering*”, Standard Publisher Distributors, 2004.

[9] Suhail I. Khattab and Salim T. Yousif, “Prediction of Soil’s Compaction Parameter using Artificial Neural Network”, *Al-Rafidain Engineering*, Vol. 21, No. 3, pp. 15-27, 2012.

[10] B.G. Farley and W.A. Clark, “Simulation of Self-Organizing Systems by Digital Computer”, *IRE Transactions on Information Theory*, Vol. 4, No. 4, pp. 76-84, 1954.

[11] Gaurav Verma and Brind Kumar, “Prediction of Compaction Parameters for Fine-Grained and Coarse-Grained Soils: A Review”, *International Journal of Geotechnical Engineering*, Vol. 14, No. 8, pp. 970-977, 2019.

[12] Han-Lin Wang and Zhen-Yu Yin, “High-Performance Prediction of Soil Compaction Parameters using Multi-expression Programming”, *Engineering Geology*, Vol. 276, pp. 1-14, 2020.

[13] D. Hebb, “*The Organization of Behavior*”, Wiley, 1949.

[14] Jeeja Jayan and N. Sankar, “Prediction of Compaction Parameters of Soils using Artificial Neural Network”, *Asian Journal of Engineering and Technology*, Vol. 3, No. 4, pp. 368-375, 2015.

[15] B. Kaminski and P. Szufel, “A Framework for Sensitivity Analysis of Decision Trees”, *Central European Journal of Operations Research*, Vol. 26, No. 1, pp. 135-159, 2017.

[16] S.C. Kleene, “Representation of Events in Nerve Nets and Finite Automata”, *Annals of Mathematics Studies*, Vol. 34, pp. 3-41, 1956.

[17] Warren McCulloch and Walter Pitts, “A Logical Calculus of Ideas Immanent in Nervous Activity”, *Bulletin of Mathematical Biophysics*, Vol. 5, No. 4, pp. 115-133, 1943.

[18] Osman Gunaydin, Abdurrahman Ozbeyaz and Mehmet Soylemez, “Regression Analysis of Soil Compaction Parameters using Support Vector Method”, *Celal Bayar University Journal of Science*, Vol. 14, No. 4, pp. 443-447, 2018.

[19] Osman Sivrikaya, Cafer Kayadelen and Emre Cecen, “Prediction of the Compaction Parameters for Coarse-Grained Soils with Fine Content by MLR and GEP”, *Acta Geotechnica Slovenica*, Vol. 2, pp. 29-41, 2013.

[20] Shakti Suman, Mahasakti Mahamaya and Sarat Kumar Das, “Prediction of Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilised Soil using Artificial Intelligence Techniques”, *International Journal of Geosynthetics and Ground Engineering*, Vol. 2, No. 11, pp. 1-11, 2016.

[21] G.N. Smith, “Probability and Statistics in Civil Engineering: An Introduction”, Collins, 1986.

[22] Sunil Khuntia, Hassan Mujtaba, Chittaranjan Patra, Khalid Farooq, Nagaratnam Sivakugan and Braja M. Das, “Prediction of Compaction Parameters of Coarse-Grained Soil using Multivariate Adaptive Regression Splines (MARS)”, *International Journal of Geotechnical Engineering*, Vol. 9, No. 1, pp. 79-88, 2015.