



# Machine Learning Techniques to Predict Rock Strength Parameters

Arsalan Mahmoodzadeh<sup>1</sup> · Mokhtar Mohammadi<sup>2</sup> · Sirwan Ghafoor Salim<sup>3</sup> · Hunar Farid Hama Ali<sup>1</sup> · Hawkar Hashim Ibrahim<sup>4</sup> · Sazan Nariman Abdulhamid<sup>4</sup> · Hamid Reza Nejati<sup>5</sup> · Shima Rashidi<sup>6</sup>

Received: 14 July 2021 / Accepted: 6 December 2021 / Published online: 6 January 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

## Abstract

To accurately estimate the rock shear strength parameters of cohesion ( $C$ ) and friction angle ( $\varphi$ ), triaxial tests must be carried out at different stress levels so that a failure envelope can be obtained to be linearized. However, this involves a higher budget and time requirements that are often unavailable at the early stage of a project. To address this problem, faster and more inexpensive indirect techniques such as artificial intelligence algorithms are under development. This paper first aims to utilize four machine learning techniques of Gaussian process regression (GPR), support vector regression (SVR), decision trees (DT), and long-short term memory (LSTM) to develop a predictive model to estimate parameters  $C$  and  $\varphi$ . To this aim, 244 datasets are available in the RockData software for intact Sandstone, including three input parameters of uniaxial compressive strength (UCS), uniaxial tensile strength (UTS), and confining stress ( $\sigma_3$ ) are employed in the models. The dropout technique is used to overcome the overfitting problem in LSTM-based models. A comprehensive evaluation is adopted for the performance indices of the prediction models. In this step, the most accurate results are produced by the LSTM model ( $C$ :  $R^2 = 0.9842$ ; RMSE = 1.295; MAPE = 0.009/ $\varphi$ :  $R^2 = 0.8543$ ; RMSE = 1.857; MAPE = 1.4301). In the second step, we improve the performance of the proposed LSTM model by fine-tuning the LSTM hyper-parameters, using six metaheuristic algorithms of grey wolf optimization (GWO), particle swarm optimization (PSO), social spider optimization (SSO), sine cosine algorithm (SCA), multiverse optimization (MVO), and moth flame optimization (MFO). The developed models' prediction performance for predicting parameter  $C$  from high to low was PSO-LSTM, GWO-LSTM, MVO-LSTM, MFO-LSTM, SCA-LSTM SSO-LSTM, and LSTM with ranking scores of 34, 29, 24, 21, 14, 12, and 5, respectively. Also, the models' prediction performance for predicting parameter  $\varphi$  from high to low was PSO-LSTM, GWO-LSTM, MVO-LSTM, MFO-LSTM, SCA-LSTM SSO-LSTM, and LSTM with ranking scores of 34, 31, 23, 18, 15, 14, and 5, respectively. However, the most robust results are produced by the PSO-LSTM model. Finally, the results indicate that applying a metaheuristic algorithm to tune the hyper-parameters of the LSTM model can significantly improve the prediction results. In the last step, the mutual information test method is applied to sensitivity analysis of the input parameters to predict parameters  $C$  and  $\varphi$ . Finally, it is revealed that parameters  $\sigma_3$  and UCS have the highest and lowest impact on the parameters  $C$  and  $\varphi$ , respectively.

## Highlights

- Employing a large dataset consists of 244 data.
- Using six ML algorithms that most of them had not been tested before for this issue.
- Applying 5-fold CV to validate the results.
- Application of feature selection to find the most effective parameters on the water inflow into tunnels.
- Recognition of the best prediction method.

**Keywords** Rock shear strength parameters · Cohesion · Friction angle · Machine learning · Metaheuristic optimization

## List of symbols

$\sigma_1$	Maximum principal stress
$\sigma_3$	Minimum principal stress
$k$	An intermediate auxiliary parameter
$\sigma_{c_i}$ <sub>fitted</sub>	Fitted UCS value from regression analysis

✉ Mokhtar Mohammadi  
mukhtar@lfu.edu.krd

Extended author information available on the last page of the article

$N$	Number of tests
$y_i$	Actual value
$\hat{y}_i$	Predicted value
$\bar{y}_i$	Mean of actual value
$\bar{\hat{y}}_i$	Mean of predicted value
$\mu(x_i)$	Mean
$k(x_i, x_i)$	Kernel
$S$	A group of samples that is not separated yet
$S_t$	A group of separated samples with true result
$S_f$	A group of separated samples with a false result

## 1 Introduction

In geotechnical and geological engineering, one of the most crucial parameters that can evaluate a rock's mechanical behavior is the strength of the intact rock with the discontinuities' presence (Grima and Babuška 1999; Gokceoglu 2002). Even though it is clear that the strength envelopes of an intact rock are a nonlinear function of the level of stress, because of its simplicity, still the linear model of Mohr–Coulomb for the shear strength of rocks is used in the actual engineering applications (Shen and Jimenez 2018).

The criterion of Mohr–Coulomb includes two parameters of cohesion ( $c$ ) and friction angle ( $\varphi$ ). Parameter  $c$  is related to the bond between crystals or particles of rock, and parameter  $\varphi$  is related to the friction internally created along the shear surface (Singh et al. 2020). Before the employment of the Mohr–Coulomb criterion in practice, parameters  $c$  and  $\varphi$  should be estimated (Adrien et al. 2020).

The popularity of triaxial tests conducted on rocks at different confining pressures to evaluate the Mohr–Coulomb parameters of  $c$  and  $\varphi$  is quite obvious. However, because of the high time and cost associated with the triaxial tests, the need for alternative methods to achieve the Mohr–Coulomb parameters is fully felt (Ulusay et al. 1994; Kahraman et al. 2009; Cai 2010; Beiki et al. 2013; Shen and Jimenez 2018). To this end, many attempts have been made to develop faster and cheaper indirect tests to estimate the uniaxial compressive strength (UCS) of rocks, such as Schmidt hammer (Mohammed et al. 2020; Howarth and Rowlands 1986), point load index (Şahin et al. 2020), impact strength (Jing et al. 2020), sound velocity (Kurtulus et al. 2018), and Los Angeles abrasion (Teymen 2019). Other researchers have conducted studies on the achievement of rock shear strength parameters with the help of UCS and uniaxial tensile strength (UTS) when the triaxial test data are not available (Beyhan 2008; Farah 2011; Karaman et al. 2015; Shen and Jimenez 2018).

Recently, non-traditional regression-based methods, and soft-computing artificial intelligence (AI) based techniques such as group method of data handling (GMDH)-type neural networks (NN) have been successfully used in a wide range

of geotechnical fields (Zendehboudi et al. 2018; Cevik et al. 2011; Mahmoodzadeh and Zare 2016; Yin et al. 2017; Liu et al. 2018; Mahmoodzadeh et al. 2019; Elbaz et al. 2019; Miah et al. 2020). However, AI techniques have not yet been widely used to predict the shear strength parameters of rocks. Recently, in their study, Shen and Jimenez (2018) applied genetic programming (GP) to predict the Mohr–Coulomb parameters of  $c$  and  $\varphi$  for Sandstone rocks. Their proposed model provided good forecasting performance in the absence of triaxial data. It was concluded that their model could be employed to estimate the practical strength of intact Sandstones at the pre-construction phase of geotechnical projects or data unavailability for the triaxial test.

Since there are many different AI algorithms, evaluating other algorithms' prediction performance can be imperative. For this purpose, this work aims to estimate the shear strength parameters ( $c$  and  $\varphi$ ) of intact rocks using three parameters of UCS, UTS, and confining stress  $\sigma_3$  by three AI methods of Gaussian process regression (GPR), support vector regression (SVR), decision trees (DT), and long-short term memory (LSTM). Parameters UCS and UTS are direct indicators of strength under uniaxial stress conditions. These parameters can be obtained in the laboratory using relatively normalized and straightforward procedures without requiring more specialized techniques. The UCS can be measured using the uniaxial compression test. The UTS can be estimated using the Brazilian test; the International Society has recommended procedures for both tests for Rock Mechanics (ISRM). In addition, to consider the non-linearity of failure envelopes and increase the reliability of predictions, we account for the influence of the stress range under which the shear failure will occur, as indicated by  $\sigma_3$ .

A database including 244 datasets previously employed by Shen and Jimenez (2018) in their research is employed in the AI models. The K-fold cross-validation (CV) method is considered to evaluate the prediction performance of the models. Finally, through analyzing the results of several statistical indices, the most accurate forecasting model is specified.

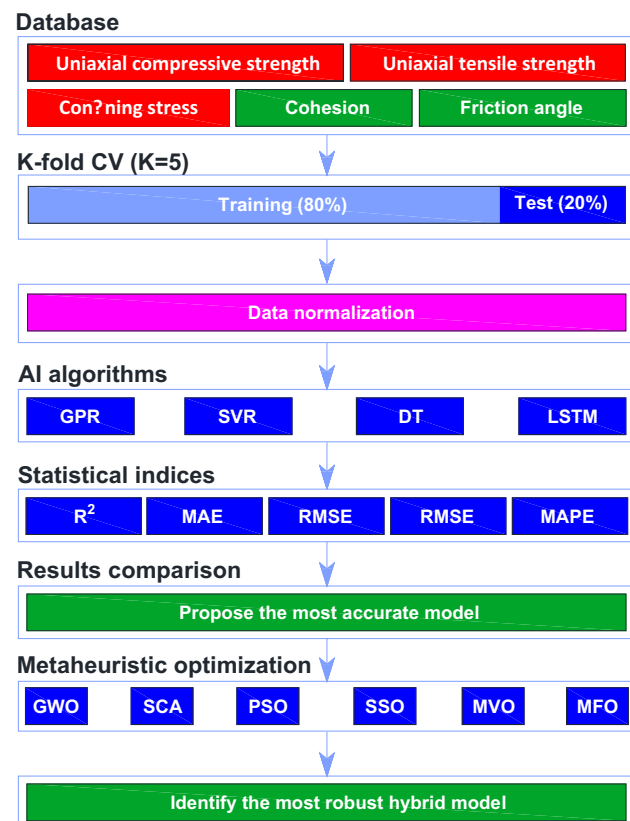
In the next step, to improve the predictions made by the proposed ML model between the four applied models, six hybrid models that are a combination of the proposed model and six metaheuristic optimization algorithms of grey wolf optimization (GWO), particle swarm optimization (PSO), social spider optimization (SSO), sine cosine algorithm (SCA), multiverse optimization (MVO), and moth flame optimization (MFO), are developed to fine-tuning of the LSTM hyper-parameters. Then, the prediction performance of the developed models for predicting parameters  $C$  and  $\varphi$  is investigated. Finally, the most robust model between the developed hybrid models is suggested.

This application demonstrates that LSTM-based metaheuristic optimization algorithms have advantages in solving the following problems: many complex parameters will affect the process and results, and the understanding of the process and results is not enough, and where there are historical or experimental data. The prediction of parameters  $C$  and  $\varphi$  is also of this type.

To determine the most influential factors between the three inputs of  $\sigma_3$ , UCS, and UTS on the parameters  $C$  and  $\varphi$ , the mutual information test method is applied.

With the above explanations, the following are the main novelties of this work to predict parameters  $C$  and  $\varphi$ .

1. Investigating four ML-based models of GPR, SVR, DT, and LSTM to predict  $C$  and  $\varphi$ . These models have not been studied for this purpose before.
2. Six metaheuristic algorithms of GWO, PSO, SSO, SCA, MVO, and MFO are developed to fine-tune the hyper-parameters of the proposed model in the prediction of parameters  $C$  and  $\varphi$ .
3. The dropout technique is used to overcome the issue of overfitting, which has not been considered in the previous ML methods for predicting  $C$  and  $\varphi$ .



**Fig. 1** Overall procedure of shear strength parameters prediction using AI techniques

4. The mutual information test is used for sensitivity analysis of the input parameters on the parameters  $C$  and  $\varphi$ .

The overall flowchart of the study is presented in Fig. 1.

## 2 Database

To predict the shear strength parameters ( $c$  and  $\varphi$ ) of intact rocks in this article, according to the literature and data availability, three effective input parameters of UCS, UTS, and  $\sigma_3$  are considered. Parameters UCS and UTS are direct indicators of strength under uniaxial stress conditions. To apply  $\sigma_3$ , a cylindrical rock specimen is placed in a specifically designed cell and the lateral pressure is applied through a liquid (usually oil) which is pumped into the cell.

A database including 244 datasets previously employed by Shen and Jimenez (2018) in their research is employed in this study. They investigated the use of linear correlations for Sandstone. To that end, they gathered extensive datasets from RocData software, presented by the company Rocscience (2012), which contains various rock properties of different rocks collected from published references. The UCS and UTS values were provided in the RocData database. They calculated the values of  $c$  and  $\varphi$  using Eqs. 1–4 with the triaxial tests available for datasets.

$$\sin\theta = \frac{k-1}{k+1} \quad (1)$$

$$c = \frac{\sigma_{ci\text{fitted}}}{2\sqrt{k}} \quad (2)$$

$$k = \frac{\sum(\sigma_3\sigma_1) - \frac{\sum\sigma_3\sum\sigma_1}{N}}{\sum\sigma_3^2 - \frac{(\sum\sigma_3)^2}{N}} \quad (3)$$

**Table 1** An overview on the database

	UCS	UTS	$\sigma_3$	$c$	$\varphi$
Count	244.00000	244.0000	244.00000	244.00000	244.00000
Mean	87.328115	5.777787	30.301516	20.949795	43.977869
Std	49.342145	4.259844	21.861209	10.225738	4.8608620
Min	26.000000	1.200000	1.470000	7.2300000	33.600000
25%	62.970000	3.157500	14.257500	14.305000	40.920000
50%	64.500000	4.050000	25.000000	17.685000	43.310000
75%	96.920000	6.190000	42.875000	24.350000	46.897500
Max	266.46000	18.10000	150.00000	55.170000	56.020000

$$\sigma_{c_i\text{-fitted}} = \frac{\sum \sigma_1}{N} - k \frac{\sum \sigma_3}{N} \quad (4)$$

where  $\sigma_1$  and  $\sigma_3$  are the maximum and minimum principal stresses,  $k$  is an intermediate auxiliary parameter,  $\sigma_{c_i\text{-fitted}}$  is the fitted UCS value from regression analysis, and  $N$  is the number of tests.

An overview of the database is presented in Table 1.

### 3 Statistical Evaluation Indices

To evaluate the accuracy of the forecasting models, some statistical evaluation indices, including coefficient of determination ( $R^2$ ), mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are taken into account. The following formulas for calculating these indices are presented (Eqs. 5–9).

$$R^2 = 1 - \frac{\text{sum squared regression (SSR)}}{\text{sum of squares total (SST)}} \quad (5)$$

$$\text{MAE} = \left(\frac{1}{n}\right) \sum_{i=1}^n |y_i - y'_i| \quad (6)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - y'_i}{y_i} \right| \times 100\% \quad (7)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2 \quad (8)$$

$$\text{RMSE} = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - y'_i)^2} \quad (9)$$

where  $y_i$  is the actual value,  $y'_i$  is the predicted value,  $\bar{y}_i$  and  $\bar{y}'_i$  are the means of actual and predicted values, and  $n$  is the number of samples.

## 4 Prediction Models of Shear Strength Parameters ( $c$ and $\varphi$ )

### 4.1 GPR

A Gaussian procedure (GP) is a gathering  $F$  of arbitrary factors  $F_{x_1}, F_{x_2}, \dots$  for which any finite subset of the factors has a joint multivariate Gaussian conveyance. The factors are listed by components  $x$  of a set  $X$ . For any finite

length vector of lists  $x = [x_1, x_2, \dots, x_n]^T$ , we have a comparing vector  $F_x = [F_{x_1}, F_{x_2}, \dots, F_{x_n}]^T$  of factors that have a multivariate Gaussian (or ordinary) distribution (Eq. 10) (Mahmoodzadeh et al. 2021a):

$$F_x \sim N\{\mu(x), k(x, x)\} \quad (10)$$

where the components of  $\mu(x)$  are given by an earlier mean capacity  $\mu(x_i)$ , and  $k$  is the portion work. The portion uses two files  $x_i$  and  $x_j$  that provides the covariance between their comparing factors  $F_{x_i}$  and  $F_{x_j}$ . Given vectors of lists  $x_i$  and  $x_j$ ,  $k$  returns the framework of covariances between all sets of factors where the first in the pair originates from  $F_{x_i}$  and the second from  $F_{x_j}$ . Each  $F_{x_i}$  is barely Gaussian, with a mean of  $\mu(x_i)$  and difference of  $k(x_i, x_i)$  (Mahmoodzadeh et al. 2021b).

Assume there has a capacity  $f(x)$  that would want to upgrade. In addition to that, suppose that  $f$  could not be watched legitimately, yet that an arbitrary variable  $F_x$  can be seen that is listed by the same space as  $f$  and whose normal esteem is  $f$ , i.e.,  $\forall x \in X, E[F_x] = f(x)$ . Notably, it is accepted that the earlier conviction about the capacity  $f$  complies with a Gaussian procedure with earlier mean  $\mu$  and part  $k$ . Furthermore, assume that  $F_x$  is a perception of  $f(x)$  that has been tainted by zero-mean, i.i.d. Gaussian clamor, i.e.,  $F_x = f(x) + \epsilon$ , where  $\epsilon \sim N(0, \sigma_\epsilon^2)$ . Consequently,  $f(x)$  is a shrouded variable whose back appropriation and could be able to derive in the wake of watching tests of  $F_x$  at different areas in the space. The following subtraction is called Gaussian procedure relapse (Mahmoodzadeh et al. 2021c, d).

Give  $x$  a chance to be the arrangement of perceptions focuses and  $F_x$  be the subsequent genuine esteemed perceptions. This required to process the back appropriation of some new point  $\hat{x} \in X$ . The appropriation will be Gaussian with mean and difference (Eqs. 11 and 12).

$$\mu(\hat{x}|x) = \mu(\hat{x}) + k(\hat{x}, x)k(x, x)^{-1}(F_x - \mu(x)) \quad (11)$$

$$\sigma^2(\hat{x}|x) = k(\hat{x}, \hat{x}) - k(\hat{x}, x)k(x, x)^{-1}k(x, \hat{x}) \quad (12)$$

The key features of the GPR model, which motivate us to use it, are as follows:

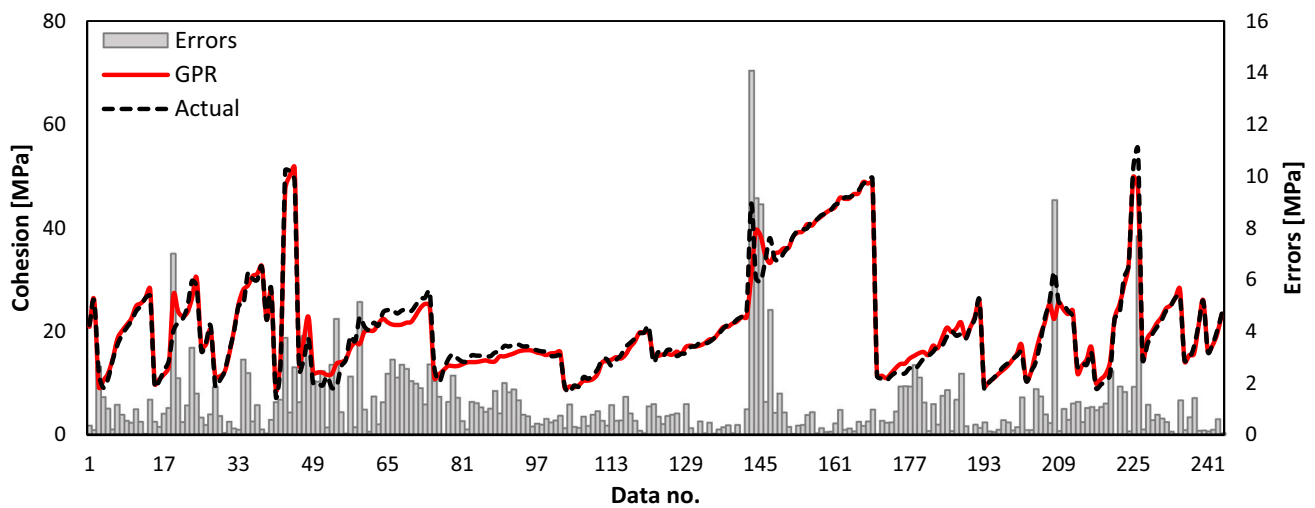
- GPR directly captures the model uncertainty. For example, in regression, GPR directly distributes the prediction value rather than just one value as the prediction. This uncertainty is not directly captured in neural networks.
- We can add prior knowledge and specifications about the model's shape by selecting different kernel functions when using GPR. For example, based on the answers to the following questions, we may choose different priors. Is the model smooth? Is it sparse? Should it be able to change drastically? Should it be differentiable? This

**Table 2** The optimized parameters of the GPR model

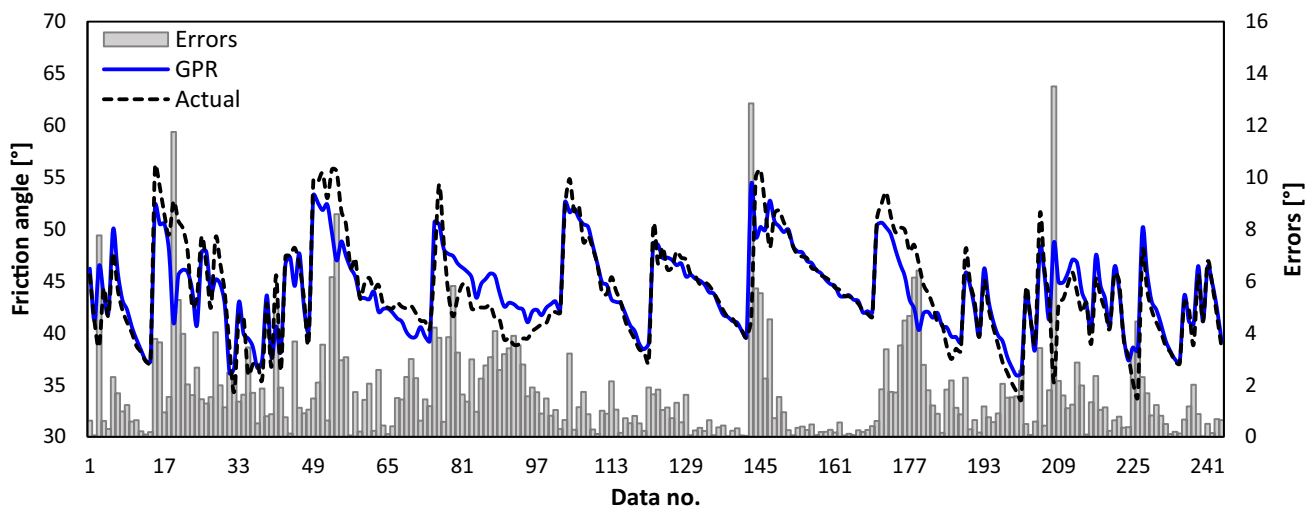
Output parameter	Parameter	Type/Value
Cohesion ( $c$ )	Kernel function	'Exponential'
	Beta	34.5763
	Fit method	Exact Gaussian process regression
	Sigma	1.1287
	Basis function	'Constant'
Friction angle ( $\varphi$ )	Kernel function	'Exponential'
	Beta	42.6303
	Fit method	Exact Gaussian process regression
	Sigma	1.9409
	Basis function	'Constant'

capability gives researchers flexible models, which can be fitted to various kinds of datasets.

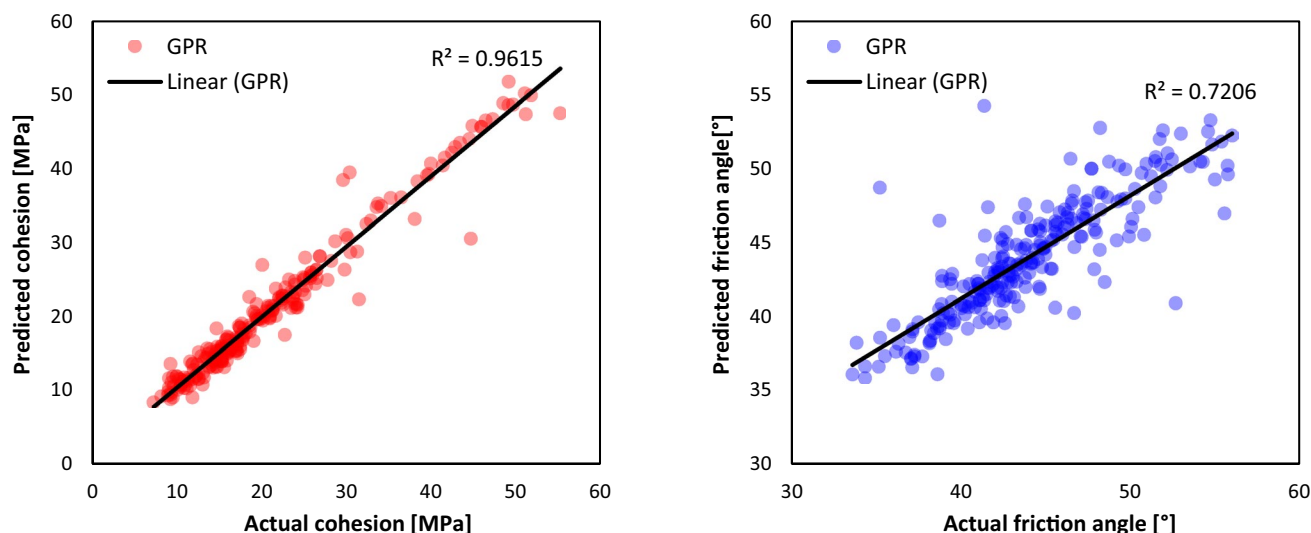
In the regression learner app embedded in MATLAB software 2018, four different GPR models are provided: squared exponential, rational quadratic, exponential, and Matern 5/2. After modeling by this program, the model type with the most accurate results is taken into account. Also, the optimization mode is considered in the app, so that the app itself optimizes the amount and type of hyper-parameters of the GPR model. The optimized type and value of the GPR hyper-parameters produced by the regression learner app are presented in Table 2.



**Fig. 2** Comparison of the cohesion results predicted by the GPR model with the actual ones



**Fig. 3** Comparison of the friction angle results predicted by the GPR model with the actual ones



**Fig. 4** Shear strength parameters results produced by the GPR model vs. the actual results (Left: Cohesion, Right: Friction angle)

**Table 3** Statistical evaluation indices results for the GPR model

Parameter	$R^2$	MAE	MSE	RMSE	MAPE [%]
Cohesion	0.9615	1.191	4.023	2.005	0.0012
Friction angle	0.7206	1.692	6.596	2.568	1.8277

The fivefold CV results of cohesion and friction angle predicted by the GPR model are shown in Figs. 2 and 3, respectively. For both actual and predicted values of parameters  $C$  and  $\varphi$  in Figs. 2 and 3, the smoothing spline method is used to fit a curve on the data. As in Figs. 2 and 3, both actual and predicted curves change together and agree well. Similar changes in these graphs and their proximity to each other indicate the proper training of the GPR model. Therefore, it can be said that the GPR model has acted correctly and accurately in predicting parameters  $C$  and  $\varphi$ . With the help of Figs. 2 and 3, it can be seen that in the high values of parameters  $C$  and  $\varphi$ , the accuracy of the predictions has decreased compared to the low values of these parameters. This may be due to the lack of data in the ranges with high values of  $C$  and  $\varphi$ . Therefore, the number of data and their range are very effective in predicting the accuracy of a ML model. The  $R^2$  results produced by the GPR model for both output parameters of  $c$  and  $\varphi$  are shown in Fig. 4. According to Fig. 4, in the GPR model, the  $R^2$  values are 0.9615 and 0.7206 for both  $c$  and  $\varphi$ , respectively. Clearly, the GPR model has shown a higher ability to predict parameter  $c$  than parameter  $\varphi$ . The other statistical indices results presented in Table 3, also confirm this assessment. The reason can be due to the type of input parameters considered in the forecast model. This means that parameters UCS, UTS, and  $\sigma_3$  have a higher effect on parameter  $c$ . However, to achieve higher

**Table 4** The optimized parameters of the SVR model

Output parameter	Parameter	Value or type
Cohesion	Kernel function	'Linear'
	Solver	'SMO'
	Epsilon	0.7569
	Bias	20.7396
Friction angle	Kernel function	'Fine Gaussian'
	Solver	'SMO'
	Epsilon	0.4496
	Bias	42.4528

accuracy in predicting the parameter  $\varphi$ , it may be necessary to consider other effective parameters on the parameter  $\varphi$ .

## 4.2 SVR

Vapnik (1995) modified his first version model ( $\epsilon$ -support vector regression, SVR) by changing the  $\epsilon$ -insensitive loss function. This modification permits the SVR model to use the margin idea in the regression process. Margin in the modified model can be described as the summation of the hyperplane's distances from the two classes' closest points. Minimizing errors between the actual training data and the hyperplane are the main target of the SVR. The kernel function idea has introduced by Vapnik (1995) for nonlinear SVR. Readers are directed to Vapnik (1995) to understand more about SVR (Mahmoodzadeh et al. 2021e).

The key features of the SVR model, which motivate us to use it, is as follows:

SVR is characterized by kernels such as linear, polynomial and Radial basis function (RBF), sparse solution, and



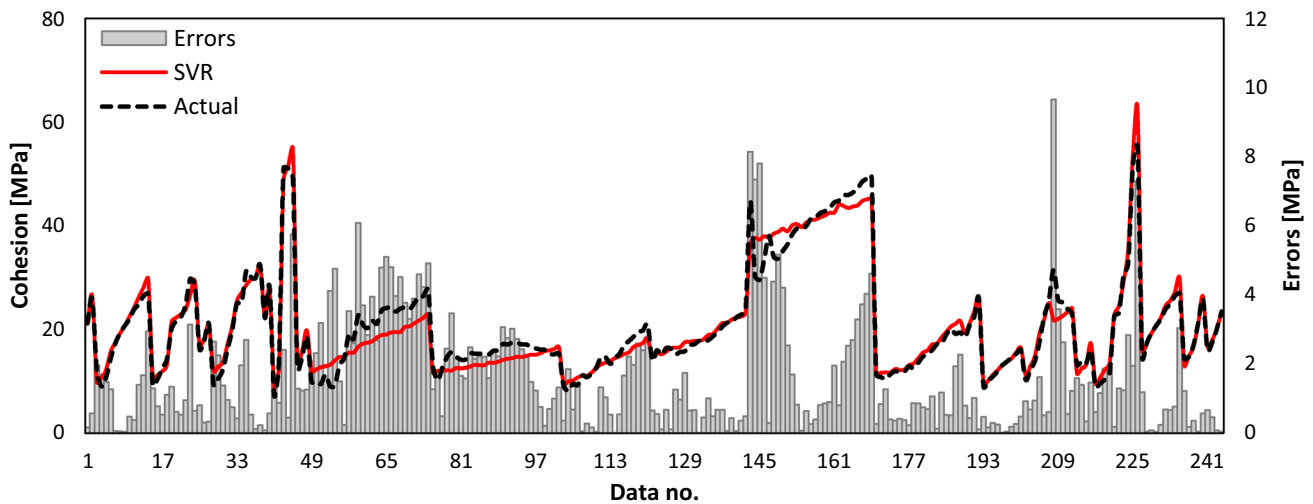


Fig. 5 Comparison of the cohesion results predicted by the SVR model with the actual ones

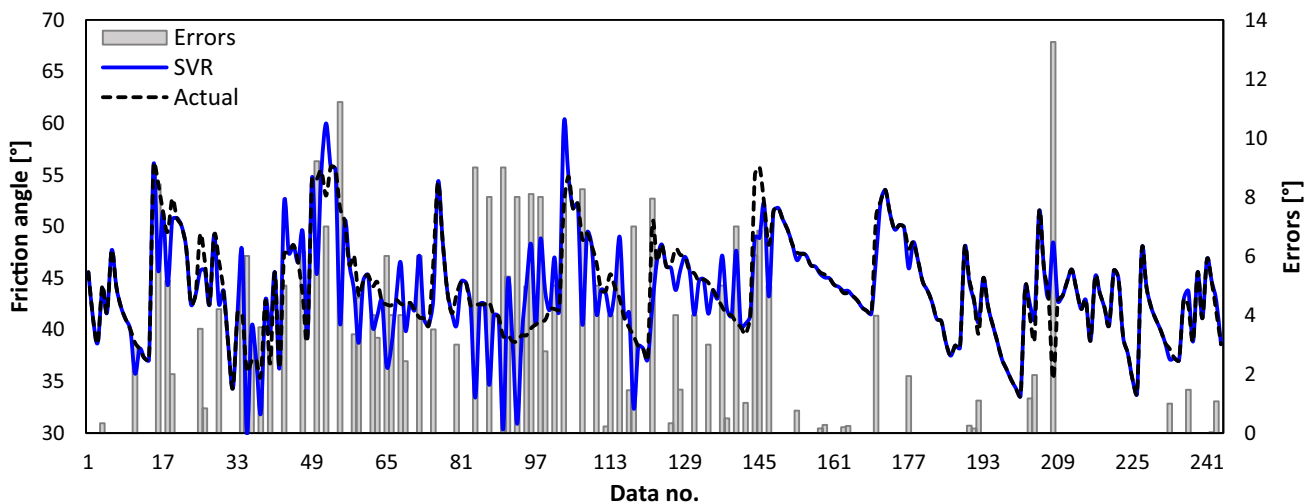


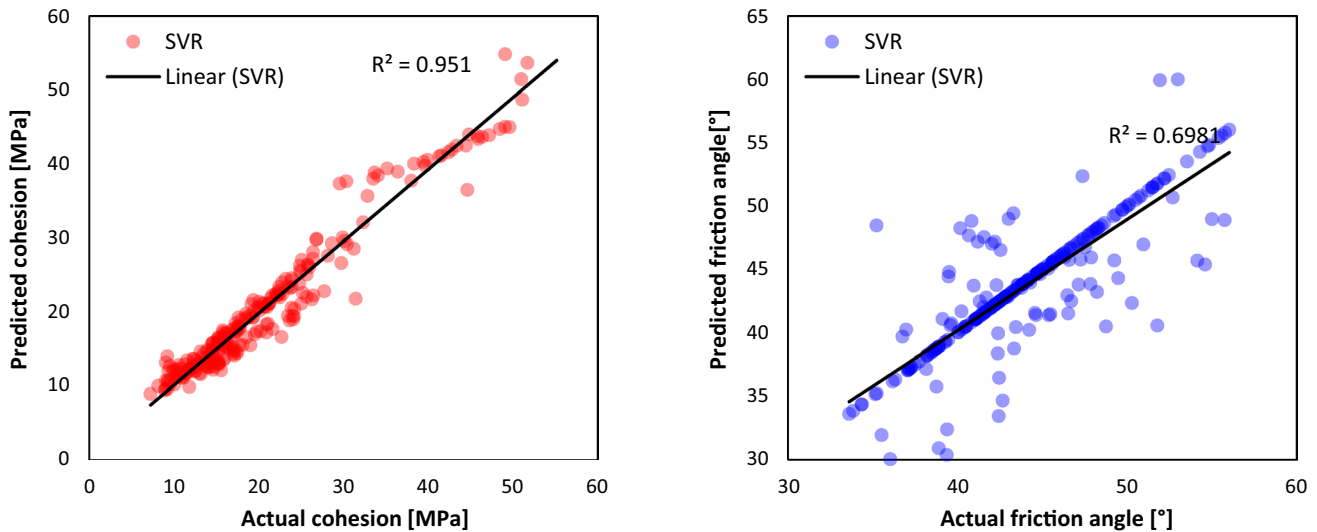
Fig. 6 Comparison of the friction angle results predicted by the SVR model with the actual ones

Vapnik–Chervonenkis (VC) control of the margin and the number of support vectors. One of the main advantages of SVR is that its computational complexity does not depend on the dimensionality of the input space. It performs lower computation compared to other regression techniques. Additionally, it has excellent generalization capability, high prediction accuracy, and is robust to outliers (Awad and Khanna 2015).

The regression learner app embedded in the Matlab 2018 software was applied to get the SVR model predictions. Six model types, including cubic, linear, medium Gaussian, quadratic, fine Gaussian, and coarse Gaussian, are provided for the SVR method in MATLAB 2018. The SVR hyper-parameters' type and values were obtained type and values

of the SVR hyper-parameters were obtained through the optimization mode in the regression learner app in Table 4.

The fivefold CV results of Mohr–Coulomb parameters of cohesion and friction angle predicted by the SVR model are shown in Figs. 5 and 6, respectively. As in Figs. 5 and 6, the actual values of parameters  $c$  and  $\varphi$  are very close to the ones predicted by the SVR model. As the GPR model, the SVR model has also presented good and acceptable predictions. As discussed for the first model, the number of data in certain ranges of the input and output parameters has shown its effect on the prediction curves. However, looking at these figures, it can be seen that the training of the SVR model is also done correctly. The  $R^2$  results produced by the SVR model for both output parameters of  $c$  and  $\varphi$  are shown in Fig. 7. According to Fig. 7, the SVR model has produced  $R^2$



**Fig. 7** Shear strength parameters results produced by the SVR model vs. the actual results (Left: Cohesion, Right: Friction angle)

**Table 5** Statistical evaluation indices results for the SVR model

Parameter	Method	$R^2$	MAE	MSE	RMSE	MAPE [%]
Cohesion	SVR	0.9510	1.599	5.182	2.276	0.0032
Friction angle	SVR	0.6981	1.316	8.281	2.877	2.9980

values of 0.9510 and 0.6981 for both  $c$  and  $\varphi$ , respectively. Clearly, as in the GPR model, the accuracy produced by the SVR model for the prediction of parameter  $c$  is higher than the prediction of parameter  $\varphi$ . This achievement can also be concluded from other statistical indices results presented in Table 5. Therefore, for the SVR model, it can be said that parameter  $c$  is more sensitive to changes in input parameters than parameter  $\varphi$ . The correlation between parameter  $c$  and inputs is greater than the correlation between parameter  $\varphi$  and model inputs.

### 4.3 DT

1. The DT is one of the classifications and regression methods based on the non-parametric survived learning technique. Furthermore, it consists of a set of if-then-else decision rules. The best prediction of the model occurs when the DT goes deeper and deeper to make the best fit with the actual data. There are several advantages of the DT. First, the distribution of explanatory variables does not require assumption. Second, strong relations among independent variables do not affect the DT outcomes. Third, various dependent variables such as

survived data, categorical and numerical can be covered by DT. Fourth, this technique comprises the influential variables and eliminates the least powerful variables which describe the dependent variable. For the DT, it is possible to predict small and large datasets well, even though this technique was initially developed to only well predict extensive data (Mahmoodzadeh et al. 2020a, b, c).

The algorithm of DT can be explained as follow:

1. First, the calculation of the targeted variance is performed.
2. Based on the various attributes, the database is divided into distinct parts, and the variance of each sectioned part is deducted from the variance before the division. This can be defined as variance reduction.

The decided node of the attribute is based on the highest VR.

Node  $N$  can be defined by the variance reduction as Eq. 13:



$$I_V(N) = \frac{1}{|S|^2} \sum_{i \in S} \sum_{j \in S} \frac{1}{2} (x_i - x_j)^2 - \left( \frac{1}{|S_t|^2} \sum_{i \in S_t} \sum_{j \in S_t} \frac{1}{2} (x_i - x_j)^2 + \frac{1}{|S_f|^2} \sum_{i \in S_f} \sum_{j \in S_f} \frac{1}{2} (x_i - x_j)^2 \right) \tag{13}$$

$S$  is a group of samples that is not separated yet,  $S_t$  is a group of separated samples with true result and  $S_f$  is a group of separated samples with a false result. Without referring to the mean, each of the summands presented above is variance estimates written in a form. Variance estimation is required in each summation term in Eq. 10 so the mean is not referred to directly.

3. Depending on the values of selected attributes, the data-sets are separated. If the variance of a part is more than zero, it is separated once more.
4. Keep another trial going until all the data is evaluated.

The key features of the DT model, which motivate us to use it, are as follows:

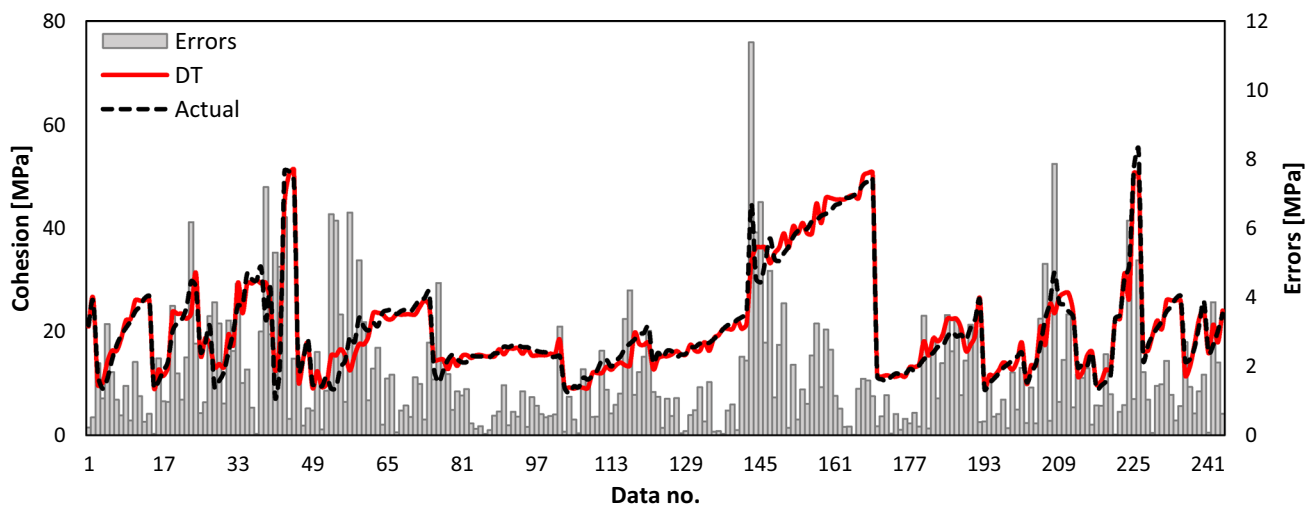
Decision trees (DTs) are supervision learning algorithms that repeatedly split the sample based on certain sample questions. These are very useful for prediction problems. They are relatively easy to understand and very effective. DTs represent several decisions followed by different chances of occurrence. This technique helps us to define the most significant variables and the relation between two or more variables. In our problem, the variables are related to each other, so we select DT to compare models. In other words, a decision tree's significant advantage is that it forces the consideration of all possible outcomes of a decision and traces each path to a conclusion. It creates a comprehensive analysis of the consequences along each branch and identifies decision nodes that need.

Key advantages:

- No preprocessing is needed on data.
- No assumptions on the distribution of data.
- Handles collinearity efficiently.
- DT can provide an understandable explanation for the prediction.

**Table 6** The optimized model type and hyper-parameters of the DT method

Output parameter	Parameter	Type/Value
Cohesion	Prune	on
	Model type	Fine tree
	Maximum number of splits	243
	Predictor selection	allsplits
	Minimum leaf size	4
	Split criterion	mse
	Minimum ParentSize	10
Friction angle	Prune	on
	Model type	Fine tree
	Maximum number of splits	243
	Predictor selection	allsplits
	Minimum leaf size	4
	Split criterion	mse
	Minimum ParentSize	10



**Fig. 8** Comparison of the cohesion results predicted by the DT model with the actual ones

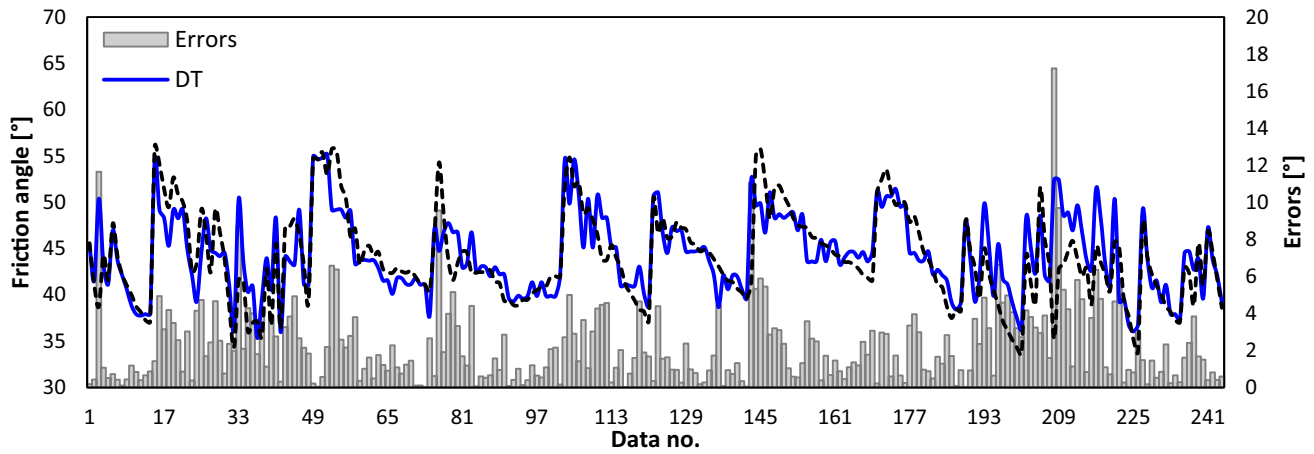


Fig. 9 Comparison of the friction angle results predicted by the DT model with the actual ones

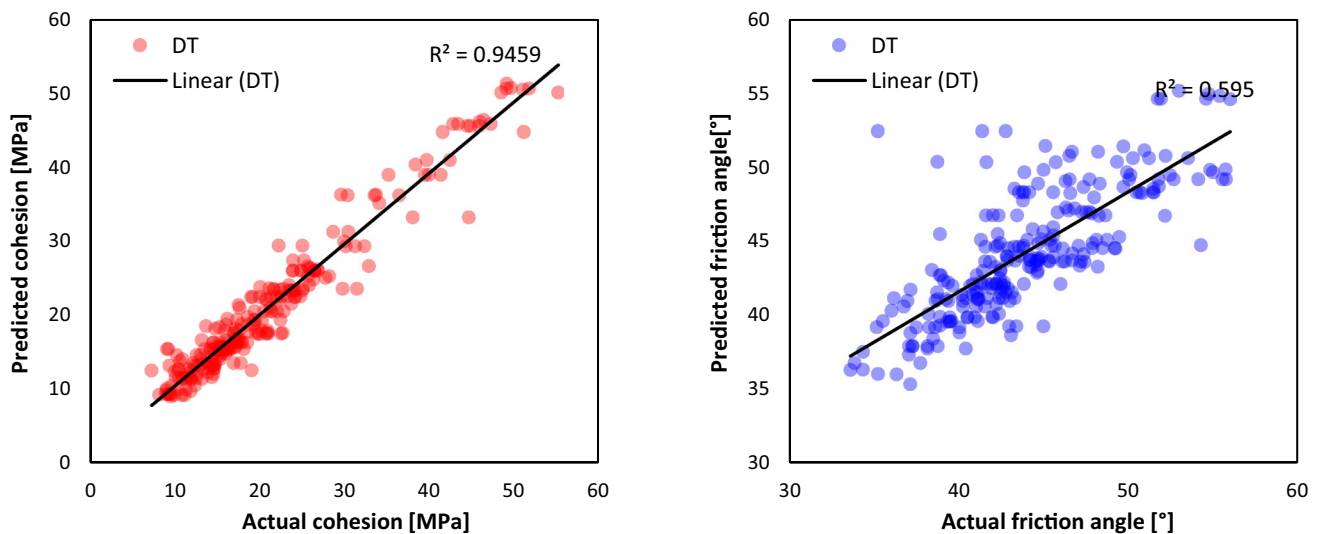


Fig. 10 Shear strength parameters results produced by the DT model vs. the actual results (Left: Cohesion, Right: Friction angle)

Table 7 Statistical evaluation indices result for the DT model

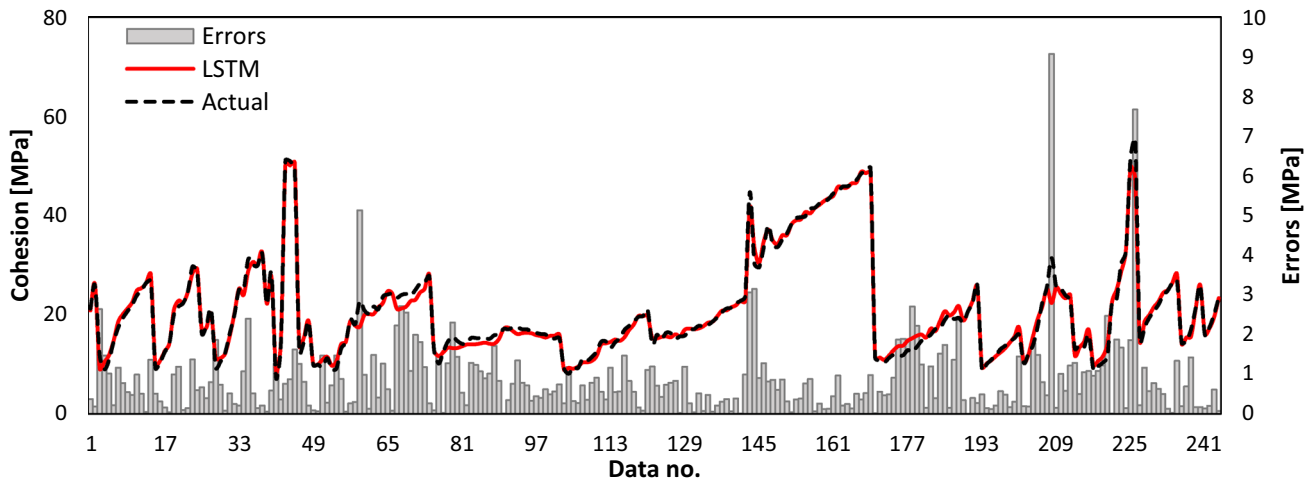
Parameter	R <sup>2</sup>	MAE	MSE	RMSE	MAPE [%]
Cohesion	0.9459	1.700	5.652	2.377	0.0043
Friction angle	0.5950	2.243	9.880	3.143	5.1387

In the DT approach, three models, medium, coarse, and fine, are embedded in the MATLAB 2018. The Mohr–Coulomb parameters predictions were performed through these three models and eventually considered the model that provided more precise results. The information about the optimized DT's hyper-parameters considered in this analysis is provided in Table 6.

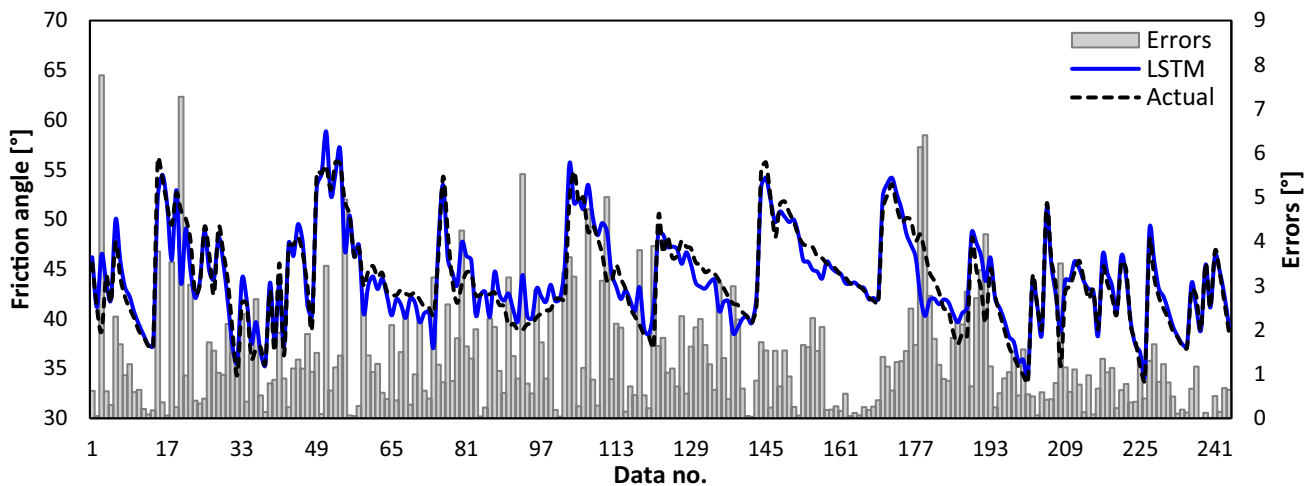
The fivefold CV results of Mohr–Coulomb parameters of cohesion and friction angle predicted by the DT model are shown in Figs. 8 and 9, respectively. As in Figs. 8 and 9, the actual values of parameters  $c$  and  $\varphi$  are in good agreement with the ones predicted by the DT model. The  $R^2$  results of the DT model for both output parameters of  $c$  and  $\varphi$  are shown in Fig. 10. According to Fig. 10, the DT model has produced  $R^2$  values of 0.9459 and 0.5950 for both parameters  $c$  and  $\varphi$ , respectively. Clearly, as in the GPR and SVR models, the accuracy produced by the DT model for the prediction of parameter  $c$  is higher than the prediction of parameter  $\varphi$ . This achievement can also be concluded from other statistical indices results presented in Table 7. Therefore, as the GPR and SVR models, for the DT model, it can be said that parameter  $c$  is more sensitive to changes

**Table 8** Type of parameters considered in the LSTM model

Parameter	Value or type
Hidden layers no	4
Batch_size	10
Activation function	'reLu'
Number of neurons	Input layer: 3, 1st hidden layer: 160, 2nd hidden layer: 80, 3rd hidden layer: 40, 4th hidden layer: 20, Output layer: 1
Loss	'mse'
Kernel_initializer	'he_normal'
Epochs	100
Validation_split	0
verbose	1
Optimizer	'Nadam'



**Fig. 11** Comparison of the cohesion results predicted by the LSTM model with the actual ones



**Fig. 12** Comparison of the friction angle results predicted by the LSTM model with the actual ones

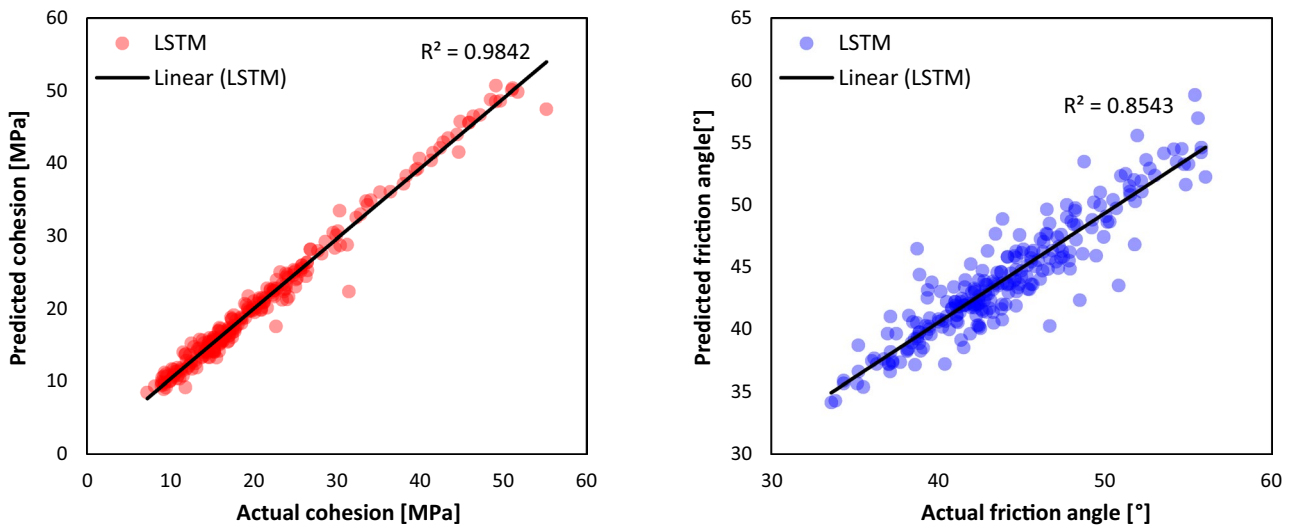


Fig. 13 Shear strength parameters results produced by the LSTM model vs. the actual results (Left: Cohesion, Right: Friction angle)

Table 9 Statistical evaluation indices result for the LSTM model

Parameter	$R^2$	MAE	MSE	RMSE	MAPE [%]
Cohesion	0.9842	0.846	1.677	1.295	0.0009
Friction angle	0.8543	1.336	3.448	1.857	1.4301

in input parameters than parameter  $\phi$ . The correlation between parameter  $c$  and inputs is greater than the correlation between parameter  $\phi$  and model inputs.

### 4.4 LSTM

The LSTM method is a form of repeat-neural network (RNN), and the final step output data is used in this step. It is utilized for the storing, prediction and classification of time series results. A standard LSTM system comprises an input gate, a cell, a forgotten gate, plus an output gate. Data are recalled by the cell through unspecified times, then the information flows into and out of the cell through the three gates. The problem of long-term RNN dependencies has been solved by LSTM, where the RNN does not expect the word kept in the long term, but it can more

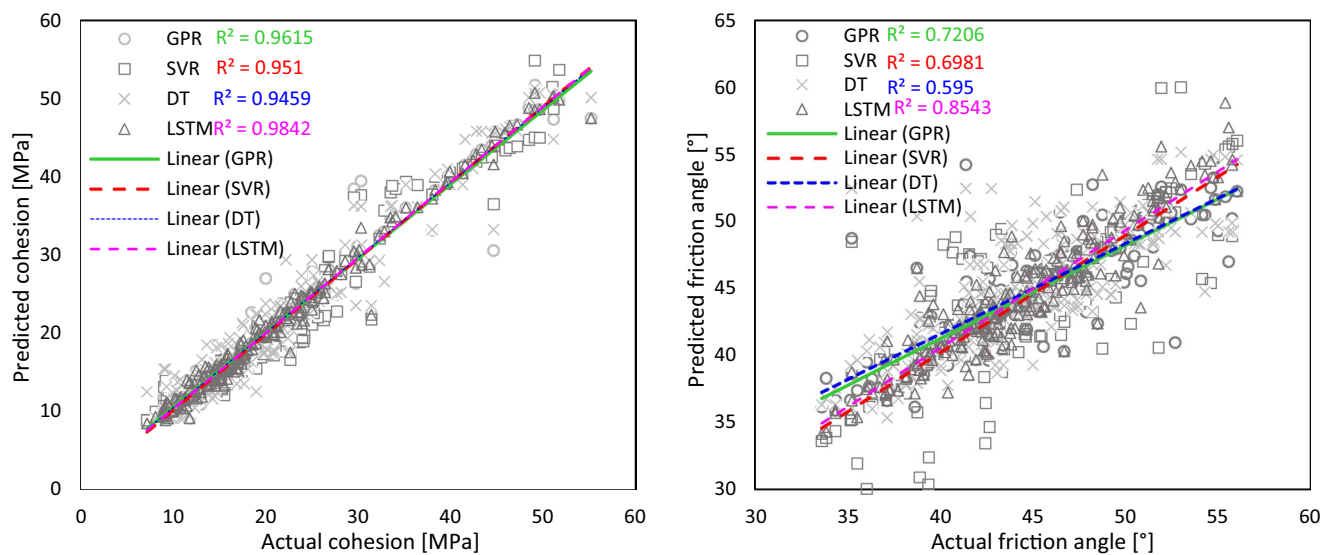
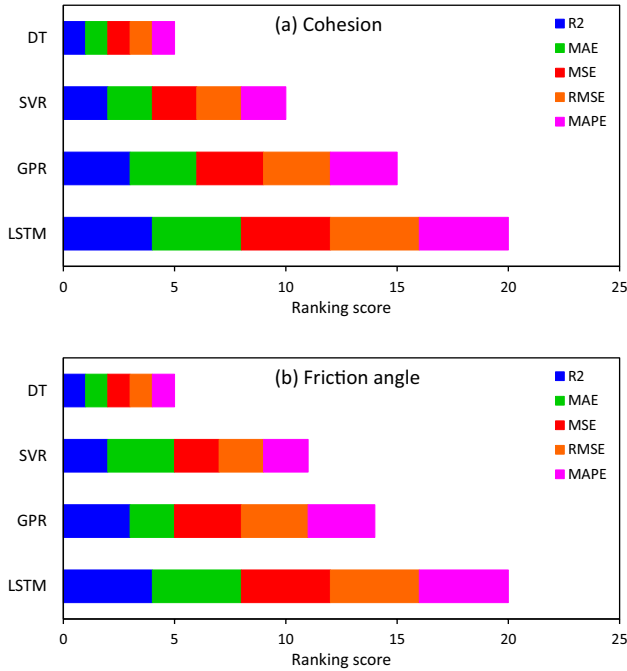


Fig. 14 Comparison between the  $R^2$  of the prediction models (Left: Cohesion, Right: Friction angle)

**Table 10** Comparison among the results produced by the ML models

Parameter	Method	$R^2$	Score	MAE	Score	MSE	Score	RMSE	Score	MAPE	Score	Rank
$C$	GPR	0.9615	3	1.191	3	4.023	3	2.005	3	0.0012	3	15
	SVR	0.9510	2	1.599	2	5.182	2	2.276	2	0.0032	2	10
	DT	0.9459	1	1.700	1	5.652	1	2.377	1	0.0043	1	5
	LSTM	0.9842	4	0.846	4	1.677	4	1.295	4	0.0009	4	20
$\varphi$	GPR	0.7206	3	1.692	2	6.596	3	2.568	3	1.8277	3	14
	SVR	0.6981	2	1.316	3	8.281	2	2.877	2	2.9980	2	9
	DT	0.5950	1	2.243	1	9.880	1	3.143	1	5.1387	1	5
	LSTM	0.8543	4	1.336	4	3.448	4	1.857	4	1.4301	4	20



**Fig. 15** Comprehensive ranking comparison of the prediction models

precisely expect the latest data. With increasing the duration gap, the RNN can not achieve good outcomes, and LSTM can retain details for a long time by default.

The type and value of the LSTM hyper-parameters considered in this research are presented in Table 8.

The fivefold CV results of cohesion and friction angle predicted by the LSTM model are shown in Figs. 11 and 12, respectively. As in Figs. 11 and 12, the actual values of parameters  $c$  and  $\varphi$  are very close to the ones predicted by the LSTM model. The  $R^2$  results of the LSTM model for both output parameters of  $c$  and  $\varphi$  are shown in Fig. 13. According to Fig. 13, the LSTM model has produced  $R^2$  values of 0.9842 and 0.8543 for parameters  $c$  and  $\varphi$ , respectively. Clearly, as in the GPR, SVR, and DT models, the accuracy produced by the LSTM model for the prediction of parameter  $c$  is higher than the prediction of parameter  $\varphi$ . This achievement can also be concluded from other

statistical indices results presented in Table 9. Therefore, as the GPR, SVR, and DT models, for the LSTM model, it can be said that parameter  $c$  is more sensitive to changes in input parameters than parameter  $\varphi$ . The correlation between parameter  $c$  and inputs is more significant than the correlation between parameter  $\varphi$  and model inputs.

### 5 Results Comparison

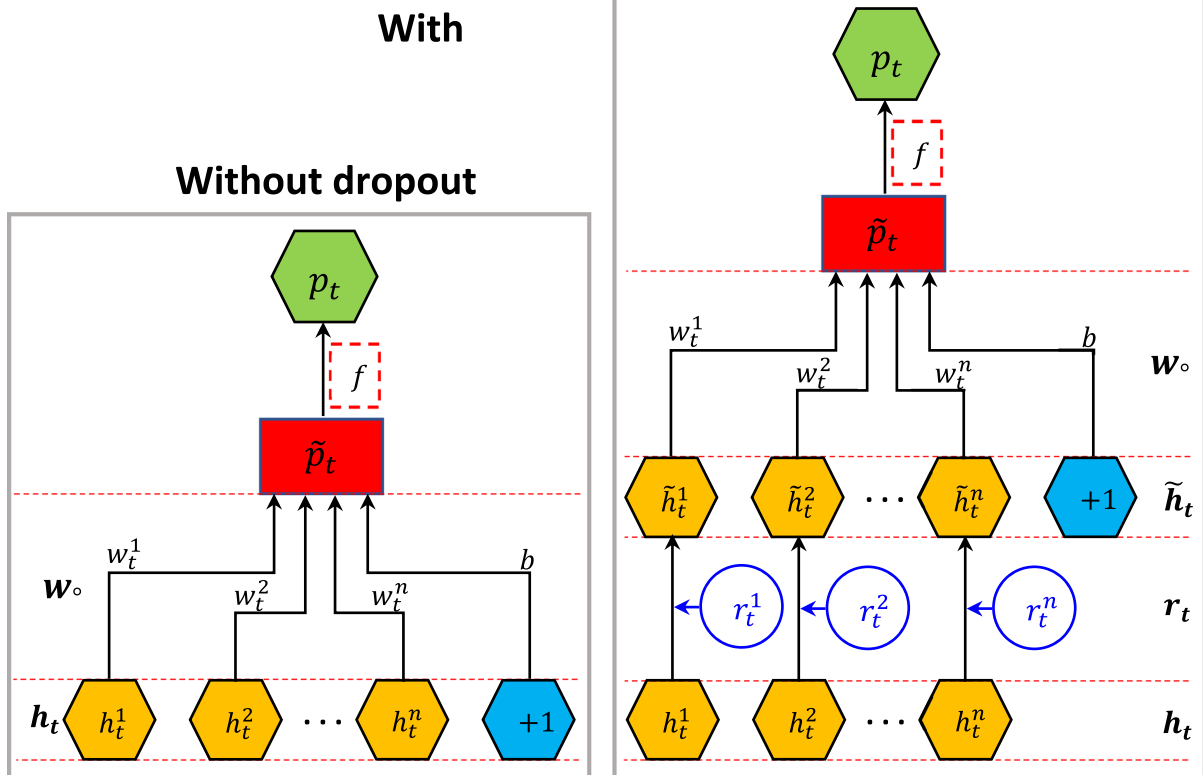
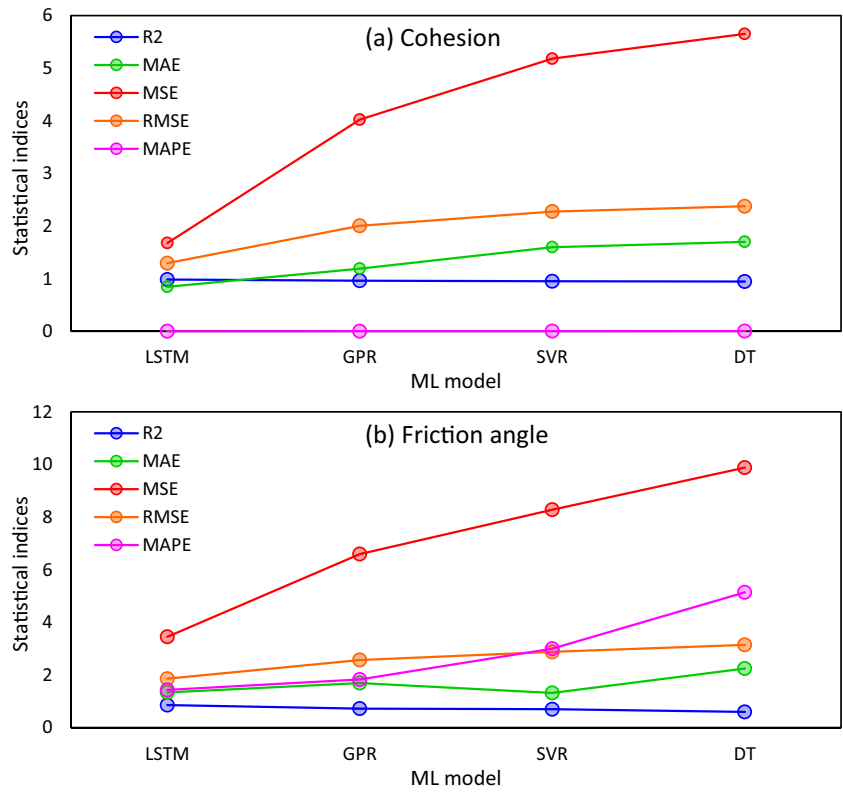
To determine the best prediction model among four ML models used in this paper to predict the shear strength parameters, in Fig. 14 and Table 10, a comparison between the results predicted by them has been made. By analyzing and comparing the values of the obtained statistical evaluation indices for each model, it can be concluded that the prediction performance of the four models for prediction of both cohesion and friction angle parameters from high to low is LSTM, GPR, SVR, and DT (see Fig. 14 and Table 10).

Table 10 illustrates the performance index results and system of ranking for ML models of LSTM, GPR, SVR, and DT in predicting rock shear strength parameters. Figure 15 shows the results of the overall ranking graphically. Figure 16 shows five evaluation indices results of the ML models. Lastly, the comprehensive ranking indicates that the LSTM model is the most robust and accurate model compared to the other three models.

### 6 Discussion

The problem of overfitting in deep learning methods such as LSTM, when the size of datasets is small, can affect the correct performance of the model. Only 244 data were applied in this study, an order of magnitude less than the data used in language and picture recognition. In ANNs, overfitting is a liable problem because of the availability of many parameters with limited training data. To obtain a reliable prediction, it is essential to avoid overfitting. The use of ANNs to predict parameters  $C$  and  $\varphi$  is more challenging due to

**Fig. 16** Evaluation indices results of the ML models



**Fig. 17** Dropout technique for overfitting: (left) without dropout; (right) with dropout



overfitting potential. The dropout approach is a useful regularization tool considered in this study for the LSTM model to alleviate the overfitting problem. The primary objective of dropout is to prevent networks from becoming overly reliant on individual neurons and minimize co-adaptability among neurons. The neurons are multiplied by a random variable that follows the Bernoulli distribution with a probability of  $p$  at each iteration throughout the training phase. The dropout rate is consistent with  $(1 - p)$ . The difference in structure between models with and without dropout is shown in Fig. 17. The corresponding formulas are as follows.

Without dropout:

$$\tilde{p}_t = w_o \cdot h_t + b_o \tag{14}$$

With dropout:

$$\tilde{p}_t = w_o \tilde{h}_t + b_o = w_o (r_t \odot h_t) + b_o \tag{15}$$

where  $\tilde{p}_t$  denotes the model output before it is processed by the active function at time  $t$ ;  $h_t$  denotes the hidden layer's output vector as described in; The weight matrix and bias that connect the hidden and output layers are represented by  $w_o$  and  $b_o$ , respectively. The output vector of the hidden layer

following dropout is represented by  $\tilde{h}_t$ ;  $r_t$  denotes a Bernoulli distribution-based random vector. The final model output is:

$$p_t = f(\tilde{p}_t) \tag{16}$$

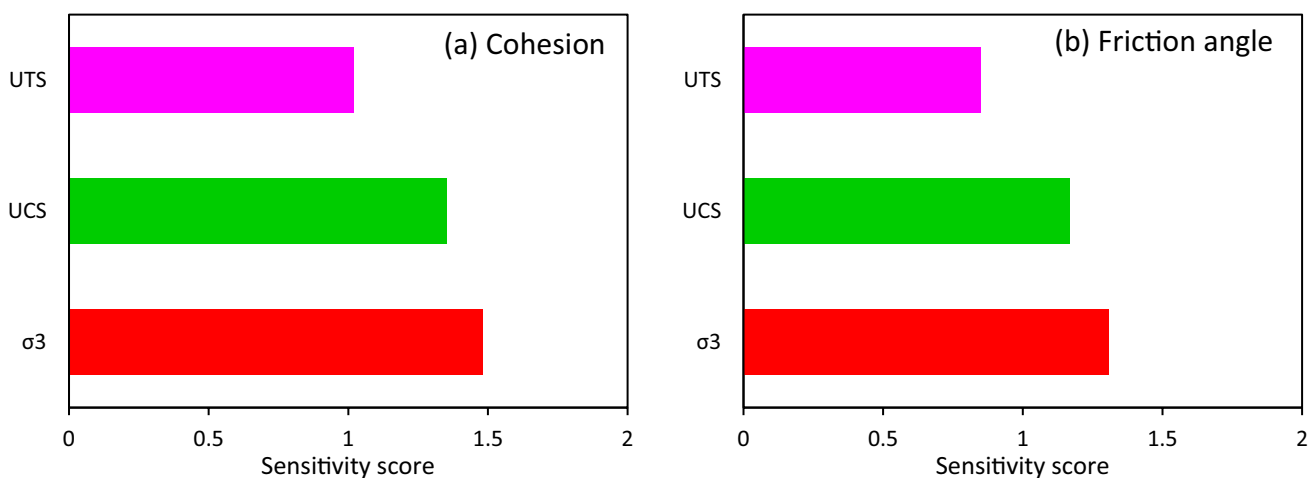
where  $p_t$  is the model output;  $f()$  represents the output layer's activation function.

It should be noted that each model predicted the cohesion parameter more accurately than the friction angle parameter. One of the reasons for this is the more significant impact of input parameters on cohesion. Other parameters may be needed to achieve higher accuracy in predicting the friction angle parameter. For example, in this paper, in addition to the three input parameters considered, the cohesion parameter was also considered the fourth input in predicting the friction angle. The values of statistical indices in predicting the friction angle using four input parameters are presented in Table 11. As in Table 11, the friction angle parameter is predicted with higher accuracy by influencing the cohesion parameter as the fourth input in the prediction models. This shows the significant impact of the type of input parameters in predicting the output parameter.

To accurately predict the shear strength parameters of  $C$  and  $\varphi$ , the impact of factors should be comprehensively studied and evaluated. In this study, three input parameters, including  $\sigma_3$ , UCS, and UTS taking into account as the effective parameters on  $C$  and  $\varphi$ . However, the sensitivity of these parameters individually is unclear on parameters  $C$  and  $\varphi$  and needs more study to reveal it. In this study, the mutual information test (MIT) proposed by Verron et al. (2008) is used to investigate the impact of the input parameters on the model output. The MIT is a filtering technique applied to capture the desired relationship

**Table 11** Statistical evaluation indices results in the prediction of friction angle parameter considering the cohesion parameter as the fourth input parameter

Method	RMSE	$R^2$	MSE	MAE
LSTM	1.1521	0.97	0.7381	0.5393
GPR	1.2798	0.93	1.6379	0.7692
SVR	2.0913	0.82	4.3736	1.0158
DT	2.7244	0.67	7.9775	2.0058



**Fig. 18** Score importance of input parameters on the prediction of parameters  $C$  and  $\varphi$

between each parameter and the label. This measure is the interdependence between parameters and shows the strength of the relationship between them. The information

gain can calculate the mutual information size between the parameters:

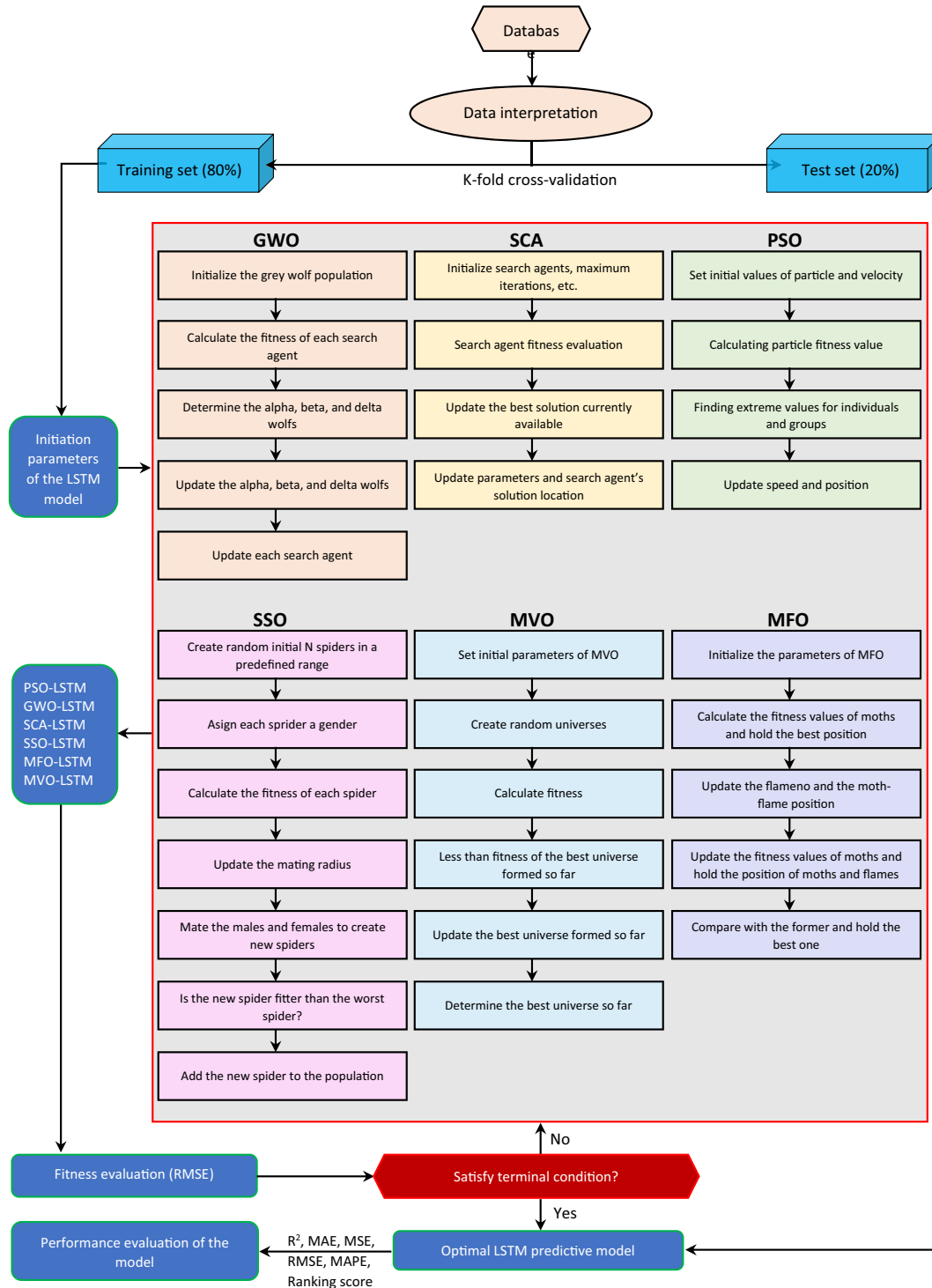


Fig. 19 The overall analysis process of hybrid intelligence models based on LSTM

$$\text{Gain}(Y, X) = \text{Ent}(Y) - \sum_{v=1}^V \frac{|Y^v|}{|Y|} \text{Ent}(Y^v) \quad (17)$$

where,  $v$  indicates the number of all possible values for  $X$ ,  $Y^v$  is the set  $Y$  related to when  $x$  takes  $x_v$ , and  $\text{Ent}(Y)$  is the entropy of the information. As  $\text{Gain}(Y, X)$  increases, the correlation between  $X$  and  $Y$  is increased.

Lastly, according to the score of the parameters in the MIT method, the importance degree of the input parameters on  $C$  and  $\varphi$  was calculated. The results obtained by the MIT method are illustrated in Fig. 18 for each input parameter. Looking at Fig. 18, it is revealed that all the three parameters of  $\sigma_3$ , UCS, and UTS with important scores of 1.48, 1.35, and 1.02 in the prediction of parameter  $C$ , and with important scores of 1.31, 1.17, and 0.85 in the prediction of parameter  $\varphi$ , respectively, have a great impact on the parameters  $C$  and  $\varphi$ . Therefore, these parameters are influential parameters that need to be considered in predicting  $C$  and  $\varphi$ . It also should be noted that, among the three input parameters,  $\sigma_3$  and UCS have the most and least impact on predicting parameters  $C$  and  $\varphi$ , respectively.

Some scholars such as Shen and Jimenez (2018), Tariq et al. (2017), and Rezaee et al. (2020) have conducted related research on rock shear strength parameters estimation and prediction through machine learning techniques in the past few years. However, achieving high accuracy in using machine learning methods is complex, and it is crucial to select the hyperparameters of the relevant model reasonably at this time.

In order to select the best model hyperparameters, this paper develops a set of hybrid prediction models to prediction of the shear strength parameters, combining LSTM and six metaheuristic optimization algorithms, including

grey wolf optimization (GWO), particle swarm optimization (PSO), social spider optimization (SSO), sine cosine algorithm (SCA), multiverse optimization (MVO), and moth flame optimization (MFO). For more information about these algorithms and to get acquainted with them, readers can refer to Zhou et al. (2021), Qiu et al. (2021), and Zendehboudi et al. (2014).

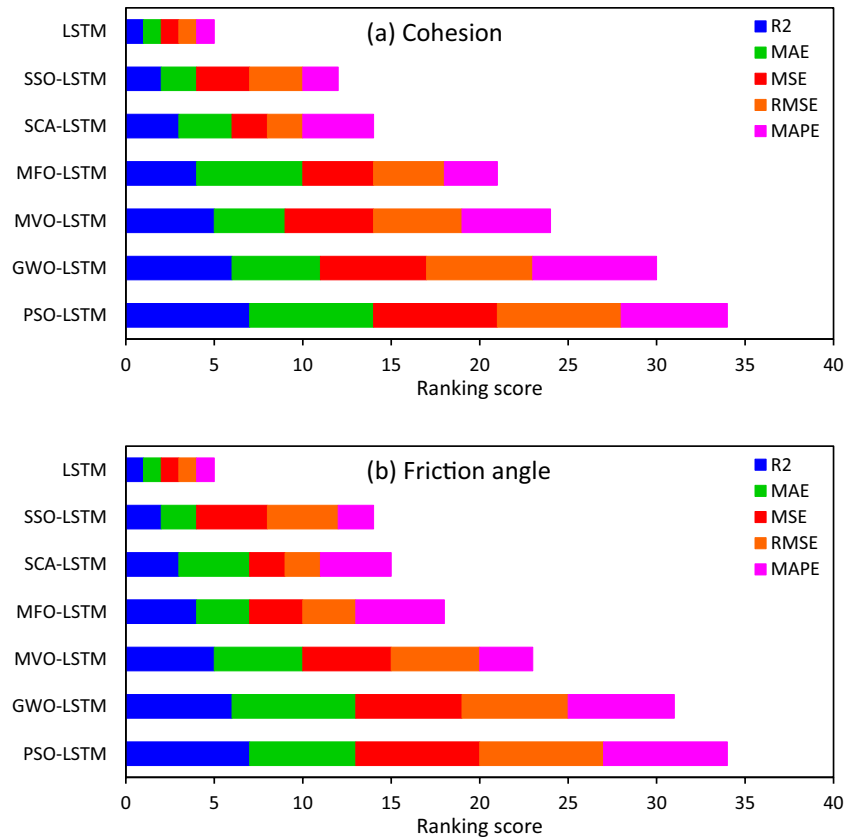
The overall analysis and hybrid modeling process is shown in Fig. 19. According to this figure, the hybrid modeling is mainly divided into four steps: (1) data set preparation; (2) model establishment; (3) model verification and evaluation; (4) result analysis. The six hybrid LSTM-based models i.e., PSO-LSTM, SCA-LSTM, SSO-LSTM, MVO-LSTM, MFO-LSTM, and GWO-LSTM are constructed to predict  $C$  and  $\varphi$ . The evaluation indices results obtained by these six hybrid models are presented in Table 12. Looking at Table 12, it can be seen that the hybrid models have increased the accuracy of the predictions compared to the non-optimized LSTM model. Also, the use of optimization algorithms to properly select the hyper-parameters has caused the predicted results related to parameter  $\varphi$  to be as accurate as the parameter  $C$ . Here, it can be said that one of the most important reasons for the low accuracy of the predictions for the parameter  $\varphi$  with the help of non-optimized previous models was that the type and amount of their hyper-parameters were not optimized. Therefore, optimizing the hyper-parameters of the ML models can be very important and increase the accuracy of forecasting models to a considerable extent.

Table 12 also illustrates the ranking system for the LSTM and six hybrid models of PSO-LSTM, SCA-LSTM, SSO-LSTM, MVO-LSTM, MFO-LSTM, and GWO-LSTM in the prediction of rock shear strength parameters. Figure 20 shows the results of the overall ranking graphically.

**Table 12** Comparison among the results produced by the LSTM model and six hybrid models

Parameter	Method	$R^2$	Score	MAE	Score	MSE	Score	RMSE	Score	MAPE	Score	Rank
C	LSTM	0.9842	1	0.846	1	1.677	1	1.295	1	0.0009	1	5
	PSO-LSTM	0.9983	7	0.482	7	0.760	7	0.872	7	0.0002	6	34
	GWO-LSTM	0.9921	6	0.546	5	0.846	6	0.920	6	0.0001	7	29
	MVO-LSTM	0.9904	5	0.551	4	0.962	5	0.981	5	0.0003	5	24
	MFO-LSTM	0.9887	4	0.542	6	1.257	4	1.121	4	0.0006	3	21
	SCA-LSTM	0.9872	3	0.694	3	1.488	2	1.220	2	0.0004	4	14
	SSO-LSTM	0.9850	2	0.715	2	1.281	3	1.132	3	0.0008	2	12
$\varphi$	LSTM	0.8543	1	1.336	1	3.448	1	1.857	1	1.4301	1	5
	PSO-LSTM	0.9831	7	0.982	6	1.423	7	1.193	7	0.9729	7	34
	GWO-LSTM	0.9769	6	0.976	7	1.479	6	1.216	6	1.0383	6	31
	MVO-LSTM	0.9721	5	1.131	5	1.491	5	1.221	5	1.3221	3	23
	MFO-LSTM	0.9616	4	1.213	3	1.938	3	1.392	3	1.286	5	18
	SCA-LSTM	0.9551	3	1.148	4	1.968	2	1.403	2	1.318	4	15
	SSO-LSTM	0.9372	2	1.280	2	1.907	4	1.381	4	1.335	2	14

**Fig. 20** Comprehensive ranking comparison of the LSTM model and six hybrid prediction models



**Fig. 21** Evaluation indices results of the LSTM model and six hybrid models

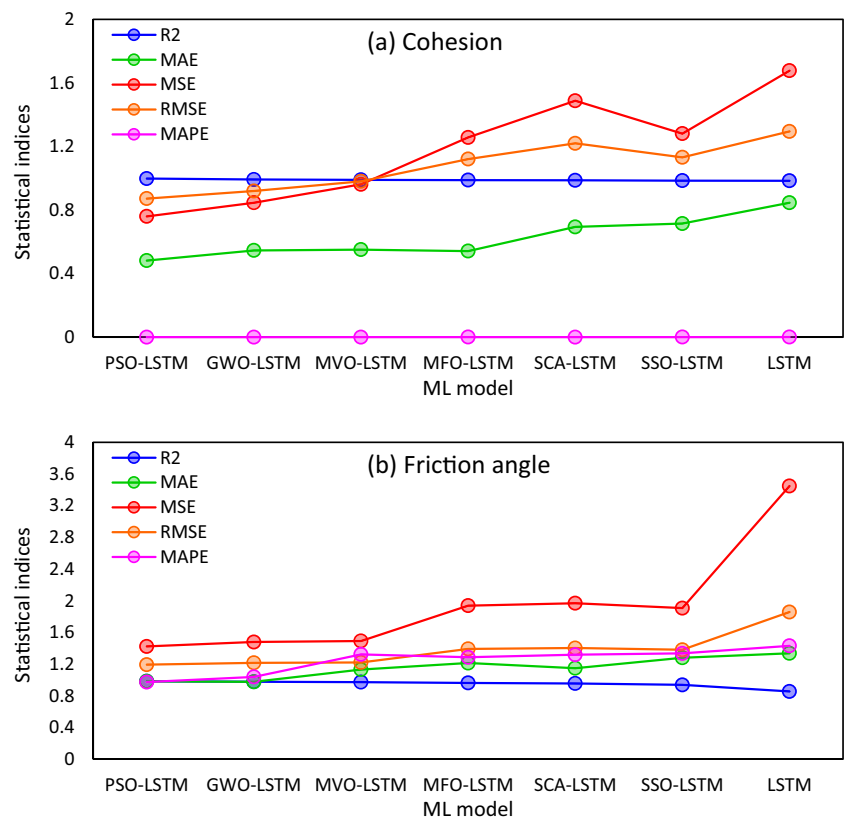


Figure 21 shows five evaluation indices results of the LSTM model and six hybrid models. The results indicate that, all the hybrid models have produced more acceptable and accurate results than the previous non-optimized techniques. According to Fig. 20, the six hybrid models' prediction performance for predicting cohesion parameters from high to low is PSO-LSTM, GWO-LSTM, MVO-LSTM, MFO-LSTM, SCA-LSTM, and SSO-LSTM with ranking scores of 34, 29, 24, 21, 14, and 12, respectively. Also, the six hybrid models' prediction performance for predicting friction angle parameters from high to low is PSO-LSTM, GWO-LSTM, MVO-LSTM, MFO-LSTM, SCA-LSTM, and SSO-LSTM with ranking scores of 34, 31, 23, 18, 15, and 14, respectively. However, the most robust results are produced by the PSO-LSTM model.

Achieving high-precision rock strength parameters in constructing tunnels, dams, buildings, and many other geotechnical engineering problems in the early stages and their construction is fundamental. The proposed PSO-LSTM model in this study is of particular importance in geotechnical engineering. Since laboratory tests require a high time and cost to achieve rock strength parameters, and it is challenging to prepare standard samples for laboratory tests in many cases, the proposed method can be critical. With very little time and cost compared to laboratory tests, this method can produce outstanding results in terms of rock strength parameters.

## 7 Conclusions

This study first proposed four ML models of GPR, SVR, DT, and LSTM to predict shear strength parameters of  $C$  and  $\varphi$ . 244 extensive datasets available in the RockData software, including three input parameters of UCS, UTS,  $\sigma_3$ , and two targets of  $c$  and  $\varphi$  were employed in the models. The fivefold CV method was used to evaluate the prediction performance of the models. The prediction performance of the four models for prediction of cohesion parameter from high to low was LSTM ( $R^2$ : 0.9842; MAE: 0.846; MSE: 1.677; RMSE: 1.295; MAPE: 0.0009), GPR ( $R^2$ : 0.9615; MAE: 1.191; MSE: 4.023; RMSE: 2.005; MAPE: 0.0012), SVR ( $R^2$ : 0.9510; MAE: 1.599; MSE: 5.182; RMSE: 2.276; MAPE: 0.0032), and DT ( $R^2$ : 0.9459; MAE: 1.7001; MSE: 5.6528; RMSE: 2.3775; MAPE: 0.0043).

Also, it was concluded that the prediction performance of the four models for prediction of friction angle parameter from high to low is LSTM ( $R^2$ : 0.8543; MAE: 1.336; MSE: 3.448; RMSE: 1.857; MAPE: 1.4301), GPR ( $R^2$ : 0.7206; MAE: 1.692; MSE: 6.596; RMSE: 2.568; MAPE: 1.827), SVR ( $R^2$ : 0.6981; MAE: 1.316; MSE: 8.281; RMSE: 2.877;

MAPE: 2.9980), and DT ( $R^2$ : 0.5950; MAE: 2.243; MSE: 9.880; RMSE: 3.143; MAPE: 5.138).

Lastly, compared to the other three models, the comprehensive ranking indicated that the LSTM model is the most robust and accurate model to predict the shear strength parameters of  $C$  and  $\varphi$ .

In the next step, to fine-tune the hyper-parameters of the proposed LSTM model, six hybrid models combining LSTM model and six metaheuristic optimization algorithms of PSO, GWO, MVO, MFO, SCA, and SSO were developed. The dropout technique was used to overcome the overfitting problem in the LSTM model and six hybrid models. The six hybrid models' prediction performance for predicting cohesion parameters from high to low was PSO-LSTM, GWO-LSTM, MVO-LSTM, MFO-LSTM, SCA-LSTM, and SSO-LSTM with ranking scores of 34, 29, 24, 21, 14, and 12, respectively. Also, the six hybrid models' prediction performance for predicting friction angle parameters from high to low was PSO-LSTM, GWO-LSTM, MVO-LSTM, MFO-LSTM, SCA-LSTM, and SSO-LSTM with ranking scores of 34, 31, 23, 18, 15, and 14, respectively. However, the most robust results were produced by the PSO-LSTM model.

Finally, the results indicated that using the metaheuristic optimization algorithm to tune the hyper-parameters of the LSTM model can significantly improve the prediction results.

The MIT method was applied in order to sensitivity analysis of the input parameters considered in this study on the prediction of parameters  $C$  and  $\varphi$ . Finally, it was revealed that all the three parameters of  $\sigma_3$  UCS, and UTS with important scores of 1.48, 1.35, and 1.02 in predicting parameter  $C$ , and essential scores of 1.31, 1.17, and 0.85 in predicting parameter  $\varphi$ , respectively, significantly impact the prediction the parameters  $C$  and  $\varphi$ .

It should be noted that the PSO-LSTM hybrid model proposed in this study as the most robust model for predicting the parameters  $C$  and  $\varphi$  just is recommended in similar conditions because it is designed based on the model inputs considered in this study. Furthermore, the procedure to predict the parameters  $C$  and  $\varphi$  introduced by this research can be implemented by other deep learning hybrid models and different optimization algorithms. The focus of this research was on the parameters  $C$  and  $\varphi$ , although LSTM-based hybrid techniques could be used to predict a wide range of geotechnical engineering problems.

Lastly, the hybrid PSO-LSTM model proposed in this study is practicable in estimating the parameters  $C$  and  $\varphi$  under similar conditions in terms of characteristics of rock mass and material conditions. The proposed models in this study can be used as practical techniques to estimate the parameters  $C$  and  $\varphi$  for similar rock mass and material properties in the site investigation step.

It is recommended that other deep learning-based hybrid techniques to predict the parameters  $C$  and  $\varphi$  and other rock strength parameters. In addition, in the next plan, more laboratory tests may be helpful to enrich more datasets to train and construct the deep learning hybrid models for predicting the parameters  $C$  and  $\varphi$ .

## Declarations

**Conflict of interest** There is no conflict of interest.

## References

- Adrien R, Patrice R, Laurent P, Pierre B (2020) Influence of roughness on the apparent cohesion of rock joints at low normal stresses. *J Geotech Geoenviron Eng* 146:04020003. [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0002200](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002200)
- Awad M, Khanna R (2015) Support vector regression. Efficient learning machines. Apress, Berkeley. [https://doi.org/10.1007/978-1-4302-5990-9\\_4](https://doi.org/10.1007/978-1-4302-5990-9_4)
- Beiki M, Majidi A, Givshad AD (2013) Application of genetic programming to predict the uniaxial compressive strength and elastic modulus of carbonate rocks. *Int J Rock Mech Min Sci* 63:159–169. <https://doi.org/10.1016/j.ijrmm.2013.08.004>
- Beyhan S (2008) The determination of G.L.I and E.L.I marl rock material properties depending on triaxial compressive strength. PhD thesis, Osman Gazi University, 224
- Cai M (2010) Practical estimates of tensile strength and hoek-brown strength parameter  $m_i$  of brittle rocks. *Rock Mech Rock Eng* 43:167–184. <https://doi.org/10.1007/s00603-009-0053-1>
- Cevik A, Sezer EA, Cabalar AF, Gokceoglu C (2011) Modeling of the uniaxial compressive strength of some clay-bearing rocks using neural network. *Appl Soft Comput* 11:2587–2594. <https://doi.org/10.1016/j.asoc.2010.10.008>
- Elbaz K, Shen S-L, Zhou A, Yuan D-J, Xu Y-S (2019) Optimization of EPB shield performance with adaptive neuro-fuzzy inference system and genetic algorithm. *Appl Sci* 9(4):780. <https://doi.org/10.3390/app9040780>
- Farah R (2011) Correlations between index properties and unconfined compressive strength of weathered Ocala Limestone. MSc thesis, University of North Florida School of Engineering, 83
- Gokceoglu C (2002) A fuzzy triangular chart to predict the uniaxial compressive strength of Ankara agglomerates from their petrographic composition. *Eng Geol* 66:39–51. [https://doi.org/10.1016/S0013-7952\(02\)00023-6](https://doi.org/10.1016/S0013-7952(02)00023-6)
- Grima MA, Babuška R (1999) Fuzzy model for the prediction of unconfined compressive strength of rock samples. *Int J Rock Mech Min Sci* 36:339–349. [https://doi.org/10.1016/S0148-9062\(99\)00007-8](https://doi.org/10.1016/S0148-9062(99)00007-8)
- Howarth D, Rowlands J (1986) Development of an index to quantify rock texture for qualitative assessment of intact rock properties. *Geotech Test J* 9:169–179. <https://doi.org/10.1520/GTJ10627J>
- Jing H, Nikafshan Rad H, Hasanipanah M, Jahed Armaghani D, Qasem SN (2020) Design and implementation of a new tuned hybrid intelligent model to predict the uniaxial compressive strength of the rock using SFS-ANFIS. *Eng Comput*. <https://doi.org/10.1007/s00366-020-00977-1>
- Kahraman S, Gunaydin O, Alber M, Fener M (2009) Evaluating the strength and deformability properties of Misis fault breccia using artificial neural networks. *Expert Syst Appl* 36:6874–6878. <https://doi.org/10.1016/j.eswa.2008.08.002>
- Karaman K, Cihangir F, Ercikdi B, Kesimal A, Demirel S (2015) Utilization of the brazilian test for estimating the uniaxial compressive strength and shear strength parameters. *J S Afr Inst Min Metall* 115:185–192
- Kurtulus C, Sertçelik F, Sertçelik I (2018) Estimation of unconfined uniaxial compressive strength using schmidt hardness and ultrasonic pulse velocity. *Teh Vjesn* 25:1569–1574. <https://doi.org/10.17559/TV-20170217110722>
- Liu XX, Shen SL, Xu YS, Yin ZY (2018) Analytical approach for time-dependent groundwater inflow into shield tunnel face in confined aquifer. *Int J Numer Anal Methods Geomech* 42:655–673. <https://doi.org/10.1002/nag.2760>
- Mahmoodzadeh A, Zare S (2016) Probabilistic prediction of the expected ground conditions and construction time and costs in road tunnels. *J Rock Mech Geotech Eng* 8(5):734–745. <https://doi.org/10.1016/j.jrmge.2016.07.001>
- Mahmoodzadeh A, Mohammadi M, Daraei A, Rashid TA, Sherwani AFH, Faraj RH, Darwesh AM (2019) Updating ground conditions and time-cost scatter-gram in tunnels during excavation. *Autom Constr* 105:102822. <https://doi.org/10.1016/j.autcon.2019.04.017>
- Mahmoodzadeh A, Mohammadi M, Abdulhamid SN, Ibrahim HH, Hama Ali HF, Salim SG (2021a) Dynamic reduction of time and cost uncertainties in tunneling projects. *Tunn Undergr Space Technol* 109:103774. <https://doi.org/10.1016/j.tust.2020.103774>
- Mahmoodzadeh A, Mohammadi M, Ibrahim HH, Abdulhamid SN, Salim SG, Hama Ali HF, Majeed MK (2021b) Artificial intelligence forecasting models of uniaxial compressive strength. *Transp Geotech* 27:100499. <https://doi.org/10.1016/j.trgeo.2020.100499>
- Mahmoodzadeh A, Mohammadi M, Hama Ali HF, Abdulhamid SN, Ibrahim HH, Noori KMG (2021c) Dynamic prediction models of rock quality designation in tunneling projects. *Transp Geotech* 27:100497. <https://doi.org/10.1016/j.trgeo.2020.100497>
- Mahmoodzadeh A, Mohammadi M, Ibrahim HH, Noori KMG, Abdulhamid SN, Hama Ali HF (2021e) Forecasting sidewall displacement of underground caverns using machine learning techniques. *Autom Constr* 123:103530. <https://doi.org/10.1016/j.autcon.2020.103530>
- Mahmoodzadeh A, Mohammadi M, Daraei A, Faraj RH, Omer RMD, Sherwani AFH (2020a) Decision-making in tunneling using artificial intelligence tools. *Tunn Undergr Space Technol*. <https://doi.org/10.1016/j.tust.2020.103514>
- Mahmoodzadeh A, Mohammadi M, Daraei A, Hama-Ali HF, Abdullah AI, Al-Salihi NK (2020b) Forecasting tunnel geology, construction time and costs using machine learning methods. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-020-05006-2>
- Mahmoodzadeh A, Mohammadi M, Daraei A, Hama-Ali HF, Al-Salihi NK, Omer RMD (2020c) Forecasting maximum surface settlement caused by urban tunneling. *Autom Constr*. <https://doi.org/10.1016/j.autcon.2020.103375>
- Mahmoodzadeh A, Mohammadi M, Ibrahim HH, Rashid TA, Aldalwie AHM, Hama Ali HF, Daraei A (2021d) Tunnel geomechanical parameters prediction using Gaussian process regression. *Mach Learn Appl* 3:100020. <https://doi.org/10.1016/j.mlwa.2021.100020>
- Miah MI, Ahmed S, Zendehboudi S, Butt S (2020) Machine learning approach to model rock strength: prediction and variable selection with aid of log data. *Rock Mech Rock Eng* 53:4691–4715. <https://doi.org/10.1007/s00603-020-02184-2>
- Mohammed DA, Alshkane YM, Hamaamin YA (2020) Reliability of empirical equations to predict uniaxial compressive strength of rocks using Schmidt hammer. *Georisk* 14:308. <https://doi.org/10.1080/17499518.2019.1658881>
- Qiu Y, Zhou J, Khandelwal M, Yang H, Yang P, Li C (2021) Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration. *Eng Comput*. <https://doi.org/10.1007/s00366-021-01393-9>



- Rezaee M, Mojtahedi SFF, Taherabadi E, Soleymani K, Pejman M (2020) Prediction of shear strength parameters of hydrocarbon contaminated sand based on machine learning methods. *Georisk*. <https://doi.org/10.1080/17499518.2020.1861633>
- Rocscience (2012) ‘‘RocData’’. <http://www.rocscience.com/products/4/RocData>. Accessed 10 Sep 2016
- Şahin M, Ulusay R, Karakul H (2020) Point load strength index of half-cut core specimens and correlation with uniaxial compressive strength. *Rock Mech Rock Eng* 53:3745–3760. <https://doi.org/10.1007/s00603-020-02137-9>
- Shen J, Jimenez R (2018) Predicting the shear strength parameters of Sandstone using genetic programming. *Bull Eng Geol Environ* 77:1647–1662. <https://doi.org/10.1007/s10064-017-1023-6>
- Singh A, Ayothiraman R, Rao KS (2020) Failure criteria for isotropic rocks using a smooth approximation of modified Mohr-Coulomb failure function. *Geotech Geol Eng* 38:4385–4404. <https://doi.org/10.1007/s10706-020-01287-5>
- Tariq Z, Elkhatny S, Mahmoud M, Abdulwahab ZA, Abdullaheem A (2017) A new approach to predict failure parameters of carbonate rocks using artificial intelligence Tools. Paper presented at the SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition, Dammam, Saudi Arabia. <https://doi.org/10.2118/187974-MS>
- Teymen A (2019) Estimation of Los Angeles abrasion resistance of igneous rocks from mechanical aggregate properties. *Bull Eng Geol Environ* 78:837–846. <https://doi.org/10.1007/s10064-017-1134-0>
- Ulusay R, Türeli K, Ider MH (1994) Prediction of engineering properties of a selected litharenite sandstone from its petrographic characteristics using correlation and multivariate statistical techniques. *Eng Geol* 37:135–157. [https://doi.org/10.1016/0013-7952\(94\)90029-9](https://doi.org/10.1016/0013-7952(94)90029-9)
- Vapnik VN (1995) *The nature of statistical learning theory*. Springer, New York
- Verron S, Tiplica T, Kobi A (2008) Fault detection and identification with a new feature selection based on mutual information. *J Process Control* 18:479–490. <https://doi.org/10.1016/j.jprocont.2007.08.003>
- Yin ZY, Jin YF, Shen SL, Huang HW (2017) An efficient optimization method for identifying parameters of soft structured clay by an enhanced genetic algorithm and elastic viscoplastic model. *Acta Geotech* 2017(12):849–867. <https://doi.org/10.1007/s11440-016-0486-0>
- Zendehboudi S, Shafiei A, Bahadori A, James LA, Elkamel A, Lohi A (2014) Asphaltene precipitation and deposition in oil reservoirs—Technical aspects, experimental and hybrid neural network predictive tools. *Chem Eng Res Des* 92:857–875. <https://doi.org/10.1016/j.cherd.2013.08.001>
- Zendehboudi S, Rezaei N, Lohi A (2018) Applications of hybrid models in chemical, petroleum, and energy systems: a systematic review. *Appl Energy* 228:2539–2566. <https://doi.org/10.1016/j.apenergy.2018.06.051>
- Zhou J, Qiu Y, Jahed Armaghani D, Zhang W, Li C, Zhu S, Tarinejad R (2021) Predicting TBM penetration rate in hard rock condition: a comparative study among six XGB-based metaheuristic techniques. *Geosci Front* 12:101091. <https://doi.org/10.1016/j.gsf.2020.09.020>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Arsalan Mahmoodzadeh<sup>1</sup> · Mokhtar Mohammadi<sup>2</sup> · Sirwan Ghafoor Salim<sup>3</sup> · Hunar Farid Hama Ali<sup>1</sup> · Hawkar Hashim Ibrahim<sup>4</sup> · Sazan Nariman Abdulhamid<sup>4</sup> · Hamid Reza Nejati<sup>5</sup> · Shima Rashidi<sup>6</sup>

Arsalan Mahmoodzadeh  
arsalan.mahmoodzadeh@uoh.edu.iq

Sirwan Ghafoor Salim  
sirwan.ghafoor@spu.edu.iq

Hunar Farid Hama Ali  
hunar.hamaali@uoh.edu.iq

Hawkar Hashim Ibrahim  
hawkar.ibrahim@su.edu.krd

Sazan Nariman Abdulhamid  
sazan.abdulhamid@su.edu.krd

Hamid Reza Nejati  
h.nejati@modares.ac.ir

Shima Rashidi  
shima.rashid@uhd.edu.iq

<sup>2</sup> Department of Information Technology, College of Engineering and Computer Science, Lebanese French University, Erbil, Kurdistan Region, Iraq

<sup>3</sup> City Planning Department, Technical College of Engineering, Sulaimani Polytechnic University, Sulaymaniyah, Kurdistan Region, Iraq

<sup>4</sup> Civil Engineering Department, College of Engineering, Salahaddin University-Erbil, 44002 Erbil, Kurdistan Region, Iraq

<sup>5</sup> Rock Mechanics Division, School of Engineering, Tarbiat Modares University, Tehran, Iran

<sup>6</sup> Department of Computer Science, College of Science and Technology, University of Human Development, Sulaymaniyah, Kurdistan Region, Iraq

<sup>1</sup> Department of Civil Engineering, University of Halabja, Halabja, Kurdistan Region, Iraq