

## Prediction of ultimate strength of shale using artificial neural network

S. Moshrefi<sup>1</sup>, K. Shahriar<sup>2\*</sup>, A. Ramezanzadeh<sup>3</sup> and K. Goshtasbi<sup>4</sup>

1. Department of Mining Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

2. Mining and Metallurgical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran

3. Faculty of Mining, Petroleum & Geophysics Engineering, Shahrood University of Technology, Shahrood, Iran

4. Department of Mining Engineering, Faculty of Engineering & Technology, Tarbiat Modares University, Tehran, Iran

Received 25 May 2017; received in revised form 23 July 2017; accepted 14 September 2017

\*Corresponding author: k.shahriar@aut.ac.ir (K. Shahriar).

### Abstract

A rock failure criterion is very important for prediction of the ultimate strength in rock mechanics and geotechnics; it is determined for rock mechanics studies in mining, civil, and oil wellborn drilling operations. Also shales are among the most difficult to treat formations. Therefore, in this research work, using the artificial neural network (ANN), a model was built to predict the ultimate strength of shale, and comparison was made with support vector machine (SVM), multiple linear regression models, and the widely used conventional polyaxial failure criteria in the stability analysis of rock structures, Drucker-Prager, and Mogi-Coulomb. For building the model, the corresponding results of triaxial and polyaxial tests have been performed on shales by various researchers. They were collected from reliable published articles. The results obtained showed that a feed forward back propagation multi-layer perceptron (MLP) was used and trained using the Levenberg–Marquardt algorithm, and the 2-4-1 architecture with root-mean-square-error (RMSE) of 24.41 exhibits a better performance in predicting the ultimate strength of shale in comparison with the investigated models. Also for further validation, triaxial tests were performed on the deep shale specimens. They were prepared from the Ramshire oilfield in SW Iran. The results obtained were compared with ANN, SVM, multiple linear regression models, and the conventional failure criterion prediction. They showed that the ANN model predicted ultimate strength with a minimum error and RMSE being equal to 43.81. Then the model was used for prediction of the threshold broken pressure shale layer in the Gachsaran oilfield in Iran. For this, a vertical and horizontal stress was calculated based on a depth of shale layer. The threshold broken pressure was calculated for the beginning and ending of a shale layer to be 154.21 and 167.98 Mpa, respectively.

**Keywords:** *Neural Network, Failure Criterion, Shale, Ultimate Strength, Support Vector Machine.*

### 1. Introduction

Instability of shale-contained formations is the cause of many complicated problems in the mining, civil, and oil and gas wellbore drilling projects. This problem is one of the main challenges in the stability of rock structures, and consumes a large amount of cost and unproductive time. Specially, shales account for nearly 75% of the drilled formations and 90% of the instability of drilling wellbores [1, 2]. Annually 500 million dollars is spent only in the oil industry in relation to the wellbore instability in the shale formations [3].

One of the main aspects of stability analysis of rock structures is the selection of an appropriate failure criterion for predicting the rock ultimate strength [4]. Many theoretical and empirical criteria have been presented by the researchers such as Hoek-Brown [5], Mohr [6], Mogi-Coulomb [7], Drucker-Prager [8], Bienawski [9], Lade [10], and Ramamurthy [11]. However, these conventional criteria are not usually appropriate for all types of rocks and all states of stresses and failure modes. The reason for this limitation is that the failure criteria are not

flexible enough to be used for various conditions [12]. The statistical methods such as simple and multiple regression techniques can be applied to establish predictive models in rock engineering [13]. Principally, statistical methods based on linear regression models suffer from several demerits. Consequently, if the data is expended in a wide range, statistic methods cannot be predicted accurately. Furthermore, this method is tolerant neither to outliers nor to extreme values in the data. Moreover, it should be noted that the statistical method is unsuitable to solve the non-linear and multivariable problems [14].

In the recent years, there has been a growth in the research works in the field of artificial intelligence like the artificial neural network (ANN) and support vector machine (SVM) techniques for developing predictive models in complicated problems. The ANN technique is considered to be one of the most adequate tools for solving intricate systems. This technique has the ability to generalize a solution from the pattern presented to it during the training process. Once the network is trained with a competent number of sample datasets, predictions can be made based on the previous learnings [14]. Due to its multi-disciplinary nature, ANN is becoming popular among the researchers, planners, and designers as an effective tool for the success of their works. In the last two decades, an increase in the ANN applications has been observed in the fields of rock mechanics and geotechnics [15]. These applications demonstrate that ANN is effective in solving problems in geosciences, whose many parameters influence the process.

ANN models are widely used to predict the mechanical properties in rock and soil mechanic engineering problems [16]. For example, Sirat and Talbot (2001) have used ANNs to recognize, classify, and predict patterns of different fracture sets in the top 450 m in crystalline rocks at the Åspö Hard Rock Laboratory (HRL), SE Sweden. Using two hidden layers with tan-sigmoid and linear transfer functions, a series of trials have been carried out using BP neural networks for supervised classification, and the BP networks have recognized different fracture sets accurately [17]. Sonmez et al. (2006) have constructed ANNs to prepare a chart for a generalized prediction of the elastic modulus of intact rock using a large database including UCS, unit weight, and modulus of elastic of intact rock ( $E_i$ ) [18]. Mohammadi and Rahmancejad (2010) have used ANNs to obtain a model for estimating rock mass deformation modulus based on the radial

basis function (RBF). The model displayed high accuracy levels when compared to in-situ tests from the elastic modulus of Karun IV dam [19]. Majdi and Beiki (2010) have used a genetic algorithm to optimize the architecture and heuristics of a BP ANN for predicting the deformation modulus of rock masses. Using a database obtained from four dam sites and powerhouses, the superiority of the ANN technique in comparison with the typical regression methods has been demonstrated [20]. Beiki et al. (2010) have employed an ANN as a tool for conducting a parametric study to determine the sensitivity of the rock mass deformation modulus to the modulus of elasticity of intact rock, UCS, rock mass quality designation, joint frequency, porosity, dry density, and geological strength index (GSI) [21]. Rafiai and Jafari (2011) have trained ANNs to predict the value of major principal stress at failure from uniaxial compressive stress and minor principal stress. They found that, on average, for different rock types, using ANN models led to about 30% decrease in the prediction error relative to the state-of-the-art empirical models [22].

Especially, with ANN of flexible tools in estimating the non-linear functions, in contrast to the conventional methods of regression, it could be implemented in predicting uniaxial strength and the ultimate strength of rock under different confining pressures. The neural network method, being flexible against the conventional failure criteria, exhibits less error in predicting the ultimate strength [12]. Dehghan et al. (2010) [13], Majidi and Rezaee (2013) [14], Cerayan et al. (2012) [23], Rabbani et al. (2012) [24], Tonnizam et al. (2015) [25], Momeni et al. (2015) [26], Barzegar et al. (2016) [27], Asadi (2016) [28], Madhubabu et al. (2016) [29] have used the ANN model for prediction of uniaxial strength using rock properties.

More detailed descriptions of the ANN-based failure criterion method have been reported by Meulenkamp and Grima (1999) [30], Singh et al. (2001) [15], Anakci and Pala (2007) [31], Tiryaki (2008) [32], Zorlu et al. (2008) [33], Rafiai and Jafari (2011) [22, 34], Kaunda (2014) [35], and Asteris and plevris (2017) [36]. Also a LSSVM-based rock failure criterion was described by Zhu et al. (2015) [37].

In this research work, considering the importance of the failure criterion in stability analysis, also problems associated with shaly formations in rock projects, especially drilling the oil and gas wellbores based on the database that includes the

triaxial and polyaxial compressive strength tests performed on the shale specimen, an ANN model was developed for prediction of the ultimate strength of shales. Then the ANN model was compared with SVM, multiple regression models, and two conventional failure criterion basis of the same database. Also for further examination, prediction of ANN model, SVM, multiple regression models, and two convention failure criteria were compared with the triaxial test results. Tests were performed on shale samples that were prepared from Ramshire oilfield in SW Iran. Then the ANN model was used to determine the threshold broken pressure of shale layer for drilling operation in the Gachsaran oilfield in Iran.

**2. Artificial neural network (ANN)**

ANN is inspired by the biological neural system (human brain) that processes the information but, in comparison to the human brain, it is relatively simple with a more straight forward functionality. The early studies corresponding to the neural network go back to 1943 [38-40]. The advantages of using ANNs are their parsimonious data requirements, rapid implementation time, and capability to yield models, where the relationship between inputs and outputs is not fully understood [41]. ANNs have the potential to model complex and non-linear relations between the input and output variables of a system and so they are commonly used in non-linear engineering problems. In the design of engineering projects, the neural network systems can be used to confirm and refine design solutions [42]. A particular network is defined using three fundamental components: transfer function, network architecture, and learning law. One has to define these constituents depending on the problem to be solved [14].

**2.1. Multi-layer perceptron (MLP)**

MLP is one of the commonly used ANN approaches for prediction studies. An MLP network consists of five parts: input, bias, weights, performance function, and output. The inputs are the input information to the network bias. A neural network is comprised of weight, indicating the effect of input to output, and bias, indicating the effect of a constant input to the neuron. Weight and bias could be adjusted, and the performance function is selected by the designer. Based upon the selected performance function and the training algorithm type, the parameters weight and bias are adjusted. Training means that by changing the weight and bias, a

logical relation is found between the inputs and outputs [34]. The number of input and output neurons is determined by the actual number of input and output variables. Figure 1 shows the structure of an MLP neural network model. In this figure, i, j, and k denote the input layer, hidden layer, and output layer neurons, respectively, and w is the applied weight by the neuron.

$$y_k = f_0 \left[ \sum_{i=1}^{M_N} W_{kj} f_h \left( \sum_{i=1}^{N_N} W_{ji} X_i W_{j0} \right) + W_{k0} \right] \quad (1)$$

where  $W_{ji}$  is a weight in the hidden layer connecting the  $i_{th}$  neuron in the input layer and the  $j_{th}$  neuron in the hidden layer,  $W_{j0}$  is the bias for the  $j_{th}$  hidden neuron,  $f_h$  is the activation (transfer) function of the hidden neuron,  $W_{kj}$  is a weight in the output layer connecting the  $j_{th}$  neuron in the hidden layer and the  $k_{th}$  neuron in the output layer,  $W_{k0}$  is the bias for the  $k_{th}$  output neuron,  $f_0$  is the activation function for the output neuron,  $X_i$  is the  $i_{th}$  input variable for the input layer, and  $y_j$  is the computed output variable.  $N_N$  and  $M_N$  are the numbers of neurons in the input and hidden layers, respectively [26].

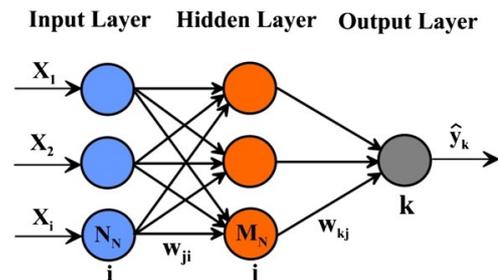


Figure 1. ANN structure [26].

The number of hidden layers, number of neurons in these layers, and number of training data are assessed based on the trial and error.

The reason for investigating these parameters is that an increase in the number of training data results in over-training, which enhances the precision of training data. However, concerning the assessment data, it results in a large error but the low number of neurons and the training data result in a lower precision per the inputs used for producing the output. Also a large number of neurons cause over-training and error in the new data.

The extra layers are dependent on the precision, and the results of the output data and the number of neurons in the input layer are dependent on the input number. Also the number of intermediate neurons depends on the assessment and skillfulness of the network designer [35].

### 3. Support vector machine (SVM)

SVM is an estimation technique based on the principals of the statistical learning theory [43]. The algorithm estimates unknown values using an optimal linear regression model in a new feature space, which is defined by mapping the input data from the original space into a higher m-dimensional space. Consider a given training data in a p-dimensional input vector and a 1D target vector. The objective is to formulate between the input and output data in the following form [43]:

$$y = f(x) = W^T \varphi(x) + b \quad (2)$$

where  $\varphi$  is a non-linear mapping function and  $W$  and  $b$  are the weighting vector and bias term of the regression equation, respectively. The optimal  $W$  and  $b$  are determined by minimizing the following risk function using the slack variables as  $\xi_i, \xi_i^*$  subjected to:

$$R(f) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^L (\xi_i, \xi_i^*) \quad (3)$$

$$\begin{cases} y_i - W^T \varphi(x_j) - b \leq \varepsilon + \xi_i \\ W^T \varphi(x_j) + b - d_i \leq \varepsilon + \xi_i^*, i = 1, \dots, L \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (4)$$

Here,  $C$  is a constant parameter that defines the trade-off between the flatness and estimation error. Quality of approximation is measured by tube in the loss function. Eq. (4) is solved based on a foundation of dual problem formulation and defining the Lagrange multipliers,  $\alpha_i, \alpha_i^* \in [0, C]$ , and ultimately, the following solution is obtained [36]:

$$f(x) = \sum_{i=1}^L (\alpha_i - \alpha_i^*) k(x_i - x_i^*) + b \quad (5)$$

where  $k(x_i - x_i^*)$  is the kernel function. The kernel function plays an important role in SVM. With a

suitable choice of kernel, the data can become separable in feature space, while the original input space is still non-linear. Thus whereas the data for n-parity or the two spirals problem is non-separable by a hyper plane in the input space, it can be separated in the feature space by the proper kernels. Table 1 gives some of the most common kernels.

Similar to other multivariate statistical models, the performances of SVM for regression depend on a combination of several parameters. They are the capacity parameter  $C$ , of insensitive loss function, and the kernel type  $K$  and its corresponding parameters.  $C$  is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the training error. In order to make the learning process stable, a large value should be set up for  $C$ . The optimal value for  $C$  depends on the type of noise present in the data, which is usually unknown. Even if enough knowledge of the noise is available to select an optimal value for, there is the practical consideration of the number of the resulting support vectors. Insensitivity prevents the entire training set meeting boundary conditions, and so allows for the possibility of sparsity in the dual formulations solution. Therefore, choosing an appropriate value is critical in the theory [36]. The schematic diagram of a SVM is shown in Figure 2.

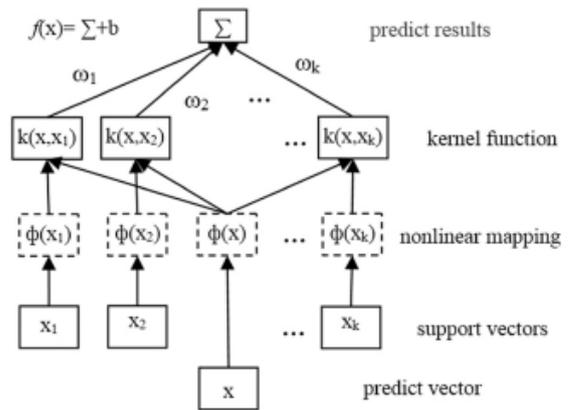


Figure 2. Schematic representation of SVM for regression [43].

Table 1. Types of kernel functions.

Name	Definition	Parameter
Linear kernel function	$k(x_i - x_i^*) = (x_i - x_i^*)$	-
Polynomial kernel function	$k(x_i - x_i^*) = [(x_i - x_i^*) + 1]^q$	q
Radial basis function	$k(x_i - x_i^*) = \exp\left(-\gamma \ x_i - x_i^*\ ^2\right)$	$\gamma$

**4. Multiple regression analysis**

**4.1. Multiple linear regression model (MLR)**

As a way to provide a visual illustration of the concept of multiple regression analysis, a quasi Venn diagram is used to explain the shared variance in correlation or regression. Simple regression analysis can show how a single dependent variable is affected by the values of one independent variable. This method only concerns the  $X_i$  variable as a predictor (i.e. independent variable) and the  $Y$  variable as an outcome (i.e. dependent variable). Thus if two or more predictors are used for the simple regression analysis, each predictor can separately show an individual relationship with the outcome variable. Another anomaly of the simple regression analysis is that it cannot predict the most significant  $X$  variable among the independent variables [44]. A multiple linear regression model is generally expressed by the relationship between a single outcome variable ( $Y$ ) and some explanatory variables ( $X_i$ ), given as:

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \tag{6}$$

where the term  $Y$  is the predicted value of  $Y$  (estimated from  $X_i$ ),  $a$  is the intercept, and  $b_i$  is the partial regression coefficient. The multiple regressions present two different overlaps, the overlap for the combined effect and the overlap for the individual effect [45].

**4.2. Multiple non-linear regression models**

Many different empirical failure criteria ( $f(\sigma_1, \sigma_2, \sigma_3)=0$ ) have been used for prediction of rock strength based on regression models in polyaxial states. In this work, the Drucker-Prager and Mogi-Coulomb failure criteria were selected because they are used in the stability analysis and they are multiple non-linear in the principal stress space ( $\sigma_1, \sigma_2, \sigma_3$ ).

**4.2.1. Drucker-Prager criterion**

This criterion was first developed for soil mechanics. This criterion was expressed in terms of the principal stresses, as follows:

$$\tau_{oct} = k + m\sigma_{oct} \tag{7}$$

$$\sigma_{oct} = \frac{\sigma_1 + \sigma_2 + \sigma_3}{3} \tag{8}$$

$$\tau_{oct} = \frac{1}{3}\sqrt{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_1 - \sigma_3)^2} \tag{9}$$

where  $m$  and  $k$  are the material constants. These values could be obtained from the drawn failure push in the  $\tau_{oct} - \sigma_{oct}$  space [7]. It can be observed that the criterion is linear in the  $\tau_{oct} - \sigma_{oct}$  space but is non-linear in the principal stress space.

$$f(\sigma_1, \sigma_2, \sigma_3) = \frac{1}{3}\sqrt{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_1 - \sigma_3)^2} - k - m\left(\frac{\sigma_1 + \sigma_2 + \sigma_3}{3}\right) = 0 \tag{10}$$

**4.2.2. Mogi-Coulomb criterion**

The Mogi-Coulomb criterion presented by Al-Ajmi, given below, has widely been used in the oil wellbore stability analysis. In fact, the Mogi-Coulomb criterion is the extended form of the Mohr-Coulomb criterion in three dimensions [46, 7].

$$\tau_{oct} = a + b\sigma_{m.2} \tag{11}$$

where  $\tau_{oct}$  is the octahedral shear stress, which is calculated by Eq. (9) and  $\sigma_{m.2}$  is also the octahedral normal stress, given by the following expression:

$$\sigma_{m.2} = \frac{\sigma_1 + \sigma_3}{2} \tag{12}$$

It was found that the criterion is non-linear in the principal stress space.

$$f(\sigma_1, \sigma_2, \sigma_3) = \frac{1}{3}\sqrt{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_1 - \sigma_3)^2} - a - b\left(\frac{\sigma_1 + \sigma_3}{2}\right) = 0 \tag{13}$$

**5. Criterion for model performance**

In this research work, to compare the efficiency of the models, coefficient of correlation ( $R$ ), coefficient of determination ( $R^2$ ), mean square error (MSE), and root mean square error (RMSE) were used. Their corresponding relationships are as follow:

$$R = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 \sum(Y - \bar{Y})^2}} \tag{14}$$

$$R^2 = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X}) \sum(Y - \bar{Y})}} \tag{15}$$

$$MSE = \frac{1}{n} \sum (Y - Y_i)^2 \tag{16}$$

$$RMSE = \sqrt{\frac{\sum (Y - Y_i)^2}{n}} \tag{17}$$

In the above expressions,  $Y_i$  is the measured value,  $Y$  is the estimated value,  $\bar{Y}$  is the mean estimated value,  $X$  is the observed value,  $\bar{X}$  is the mean observed value, and  $n$  is the total number of data [28, 47, 48].

The mean square error (MSE), root mean square error (RMSE), and coefficient of correlation ( $R$ ) are calculated for all the models, and accordingly, the model with minimum RMSE and MSE and maximum  $R$  or  $R^2$  is chosen as the optimum one.

### 6. Data collection and preparation

Data collection is one of the most important stages in ANN, SVM, and regression modeling. In this work, the results of triaxial and polyaxial tests performed on shale samples by various researchers were collected from reliable published articles [49-53]. 83 datasets were used for training, validation, and testing. The data included maximum, intermediate, and minimum stresses at the failure stage. The ranges of maximum, intermediate, and minimum stresses are shown in Table 2. Also for the more assessment of the neural network method and failure criteria, the multi-stage triaxial tests under different confining pressure values were performed on three shale rock specimens prepared from the Ramshire oilfield in SW Iran. Figure 3 and Table 3 show the characteristics of deep shale samples. Also Table 4 shows the results of the multi-stage triaxial tests.

**Table 2. Description of input and output parameter in modeling.**

Type of data	Parameter	Symbol	Unit	Min	Max
Input	Minimum stress	$\sigma_3$	Mpa	0	400
	Intermediate stress	$\sigma_2$	Mpa	0	400
Output	Maximum stress	$\sigma_1$	Mpa	32.40	745.90



**Figure 3. Samples of shale from southern and southwestern oil fields in Iran.**

**Table 3. Characterization of deep shale samples.**

Sample	Filed	Well No.	Depth (m)	Formation
O1	Ramshire	9	3260.10	Pabdeh
Y1	Ramshire	9	3266.20	Pabdeh
P1	Ramshire	9	3246.30	Pabdeh

**Table 4. Results of multi-stage triaxial test on deep shale samples.**

Number	$\sigma_2=\sigma_3$ (Mpa)	$\sigma_1$ (Mpa)
1	17	162.20
2	25	207.70
3	30	244.00
4	33	224.00
5	38	236.00
6	40	256.00

### 7. Analysis of strength data using ANN

MATLAB 2016 software was used for ANN modeling. In the MATLAB procedure, the training and testing data are chosen randomly. In this work, 70% of the data was used for training, 15% for validation, and 15% for testing.

The error back propagation algorithm is one of the most known training algorithms for multi-layer perceptron that utilizes a special training algorithm for the transfer of error from the end layer to the preceding layer, and it adjusts the weights and bias with a minimum time [28, 47]. Also the LM algorithm is known to be the fastest

method for training the moderate-sized feed-forward neural networks [33].

In order to develop a three-layered ANN model, two inputs including  $\sigma_2$  and  $\sigma_3$  are used in the first layer, and  $\sigma_1$ , as the output, is utilized in the last layer. The feed-forward neural network is trained with the Levenberg–Marquardt algorithm (Train LM). The number of hidden neurons for MLP models is selected via the trial and error method. For this, root mean square error (RMSE) is calculated for different numbers of hidden neurons, and accordingly, the model with minimum RMSE is chosen as the optimum one [14]. There is no direct method available to identify the number of hidden layers and the number of neurons in each hidden layer. The optimal number of neurons in the hidden layers is also obtained by the trial and error method based on the minimum RMSE [16].

Table 5 shows the network performance for different numbers of neurons in the hidden layers. As it can be seen, the best model has RMSE equal to 24.41 for the test datasets. This model is an optimum model with a 2-4-1 architecture, (Figure 4); it has two input neurons, four neurons in the hidden layer, and one output neuron, respectively. The results of the network are presented in this section to demonstrate its performance. The coefficient of correlation between the predicted and measured values of ultimate strength is taken as the network performance. The prediction was based upon the input datasets (discussed in the previous section). Figure 5 illustrates the coefficient of correlation ( $R$ ) for the proposed ANN model including the training, test, validation, and overall data.

### 8. Analysis of strength data using SVM

The data (input:  $\sigma_2$ ,  $\sigma_3$ ; and output:  $\sigma_1$ ) was randomly divided into two subsets: 70% of the total data was allotted to the training data for the

SVM model construction and 30% was allocated to the test data used to assess the reliability of the development model. Now it would be necessary to select a suitable kernel function. A lot of kernel functions have been proposed in the literature. Among them, the ones based on Radial Basis Functions (RBF) are widely employed. In this work, we decided to employ the RBF kernel since it has been previously used in defect prediction and context [54-56], and usually yields a better performance than the other kernels [57].

When SVM is used by the RBF kernel, the three parameters  $C$ ,  $\epsilon$ , and  $\gamma$  have to be set by the user. The selection of appropriate values for these parameters is crucial to obtain a good regression performance. As described earlier,  $C$  is the penalty factor for misclassified points. If it is too large, a higher penalty for non-separable points are added, leading to store too many support vectors and thus over-fit. On the other hand, if  $C$  is too small, an under-fitting can occur. The  $\gamma$  parameter specifies the radius of RBF, also having a strong impact on the accuracy. To obtain the optimum parameters  $C$ ,  $\epsilon$ , and  $\gamma$  of the SVM model, different values of these parameters must be examined based on the trial and error method. In this research work, Weka 3.6.9 software was used for SVM modeling. Then root mean square error (RMSE) was calculated for all the models, and accordingly, the model with minimum RMSE was chosen as the optimum model [56]. Figure 6 shows the output of Weka software for the SVM model.

The training result is presented in Table 6. As shown in this table, the coefficient of correlation of training data and RMSE are 0.96 and 32.54, implying the proper performance of SVM. Also Figure 7 shows the relation between the measured values and the failure criteria predicted values ultimate strength.

Table 5. Results of ANN model with different architectures.

ANN Model	Transfer Function	Training Law	Net Architecture	Training R	Validation R	Test R	Model R	MSE	RMSE
MLP	TANSIG	LM	2-1-1	0.97	0.96	0.98	0.96	839.50	28.97
MLP	TANSIG	LM	2-2-1	0.96	0.98	0.95	0.95	913.45	30.22
MLP	TANSIG	LM	2-3-1	0.97	0.98	0.97	0.97	759.28	27.55
MLP	TANSIG	LM	2-4-1	0.96	0.98	0.98	0.97	596.10	24.41
MLP	TANSIG	LM	2-5-1	0.95	0.98	0.96	0.96	1151.15	33.92

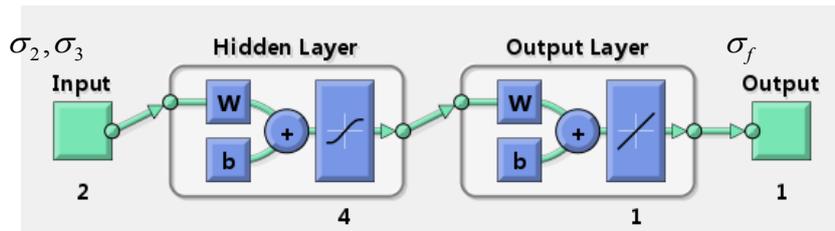


Figure 4. Topology of ANN used for prediction of polyaxial strength.

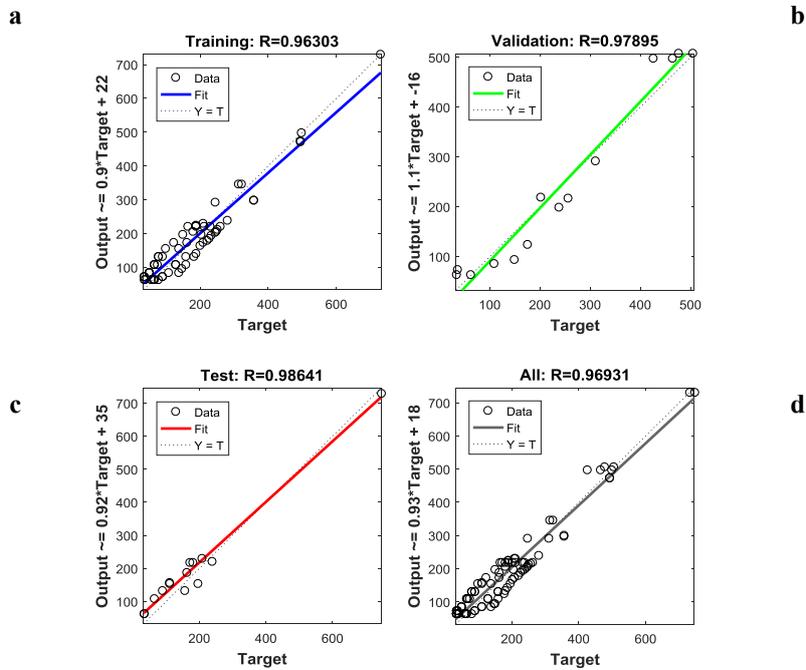


Figure 5. Relation between measured (Target) values and ANN model predicted values of ultimate strength (output) for training (a), validation (b), testing (c), and overall datasets (d) in Mpa.

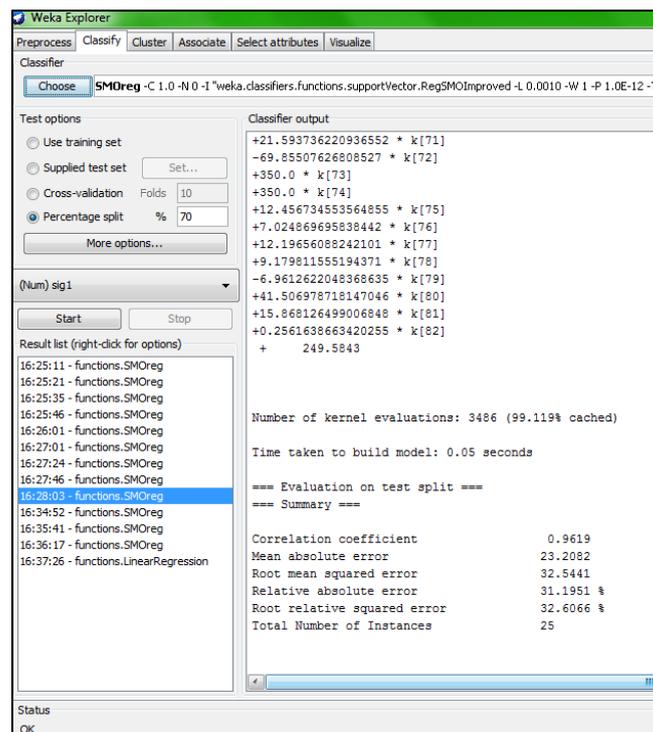
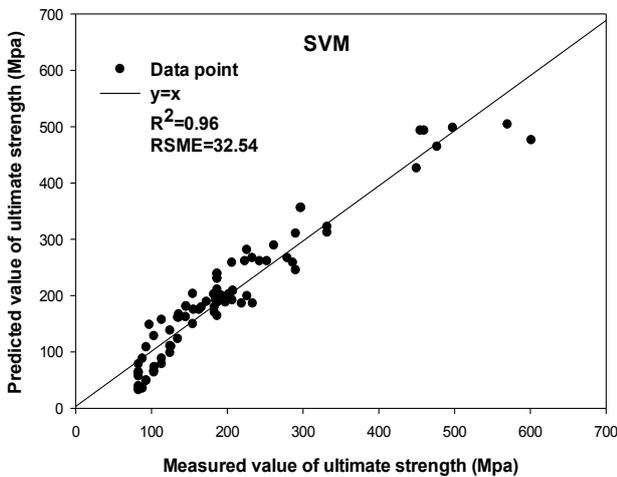


Figure 6. Weka software output for SVM model.

**Table 6. SVM model factors.**

c	$\gamma$	$\epsilon$	R	R <sup>2</sup>	RMSE
350	0.01	0.001	0.96	0.92	32.54



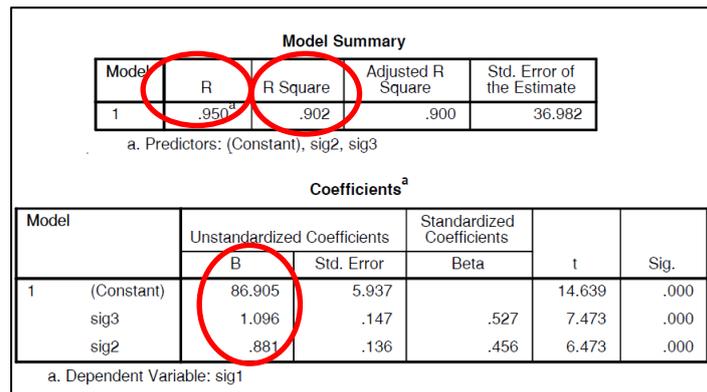
**Figure 7. Relation between measured values and SVM model predicted values ultimate strength.**

**9. Analysis of strength data using multiple linear regression and convention failure criterion**

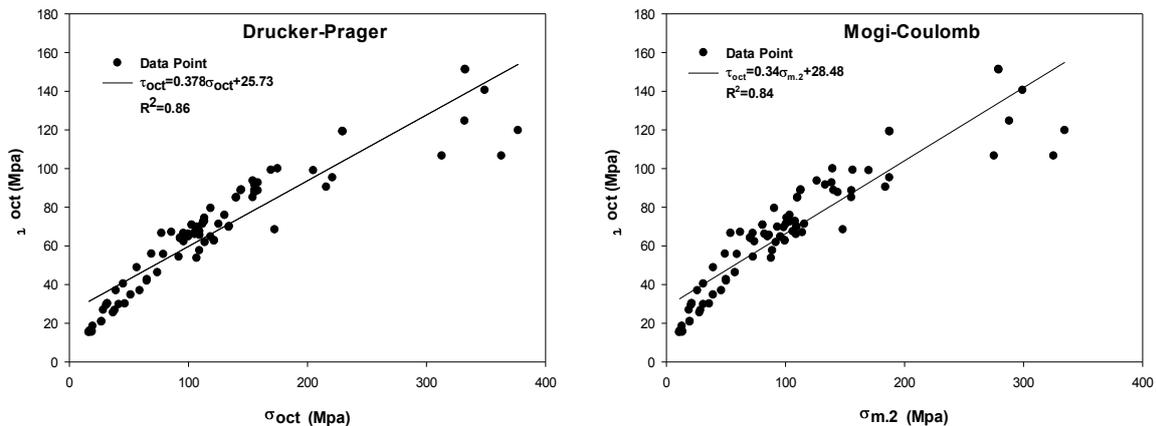
The same datasets were used for training the ANN and SVM models. They are analyzed to obtain the parameters of multiple linear and conventional criteria. The parameters are introduced in Section 4.

Figures 8 and 9 show the SPSS software output for the multiple linear regression model and fitting curves on test data of shale to determine the parameters of Drucker-Prager and Mogi-Coulomb failure criteria. The results of the models are summarized in Table 7. In this research work, SPSS, Sigmaplot, and Excel software were used for determination and plot of the regression models.

Figure 10 and Table 8 show the relation between the measured values and predicted values of ultimate strength based on the multiple linear, Drucker-Prager, and Mogi-Coulomb failure criteria.



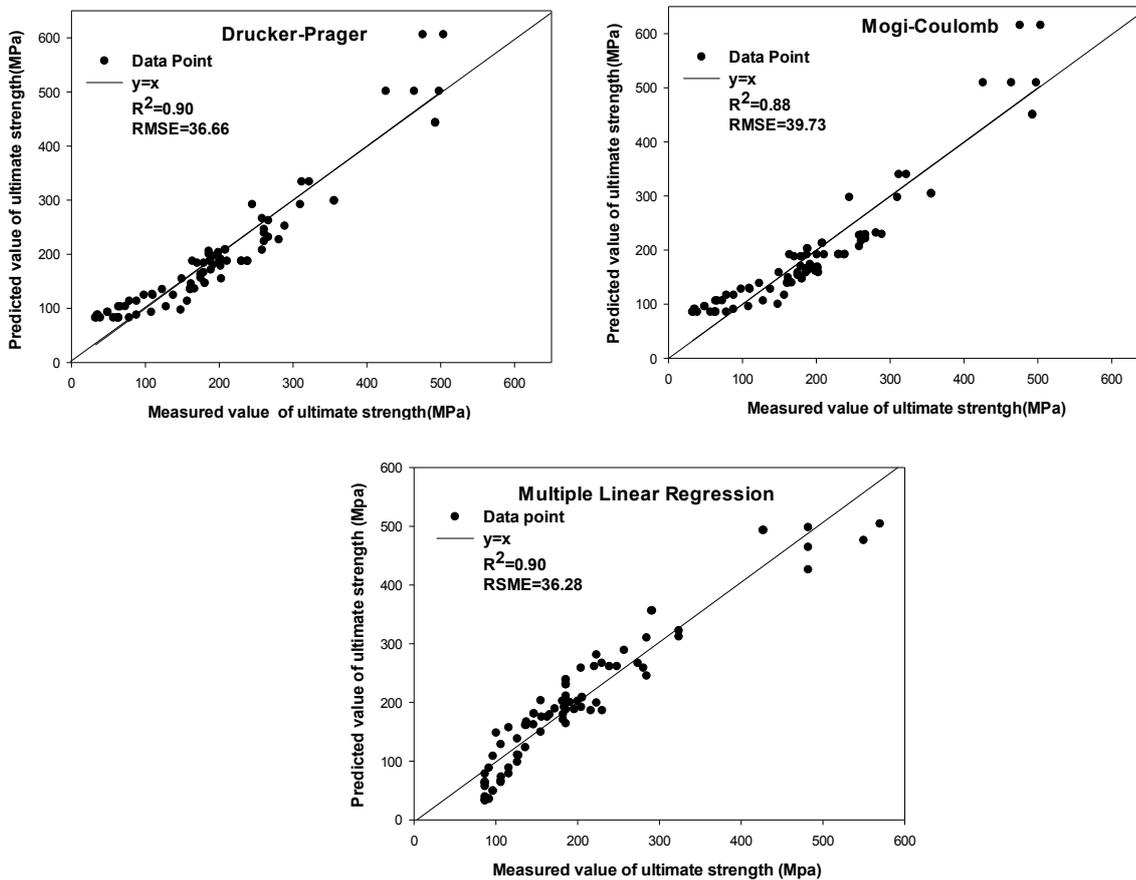
**Figure 8. SPSS software output for multiple linear regression model.**



**Figure 9. Fitting curves on test data of shale to determine parameters of Drucker-Prager and Mogi-Coulomb failure criteria.**

**Table 7. Parameters and coefficient of determination of multiple linear, Drucker-Prager, and Mogi-Coulomb failure criteria.**

Model	Equation	R	R <sup>2</sup>	Parameter	
MLR	$\sigma_1 = a + b_1\sigma_2 + b_2\sigma_3$	0.95	0.90	a=86.90	b <sub>1</sub> =1.09, b <sub>2</sub> =0.88
Drucker-Prager	$\tau_{oct} = k + m\sigma_{oct}$	0.93	0.86	k=25.73	m=0.378
Mogi-Coulomb	$\tau_{oct} = a + b\sigma_{m,2}$	0.92	0.84	a=28.48	b=0.34



**Figure 10. Relation between measured values and failure criteria predicted values ultimate strength of multiple linear, Drucker-Prager, and Mogi-Coulomb failure criteria in Mpa.**

**Table 8. Performance indices of multiple linear, Drucker-Prager, and Mogi-Coulomb failure criteria.**

Models	R	R <sup>2</sup>	RMSE	MSE
Drucker-Prager	0.95	0.90	36.66	1343.95
Mogi-Coulomb	0.94	0.88	39.73	1578.47
MLR	0.95	0.90	36.28	1207.56

**10. Comparison of models**

In order to control the prediction performances of the ANN, SVM, multiple linear, Drucker-Prager, and Mogi-Coulomb failure criteria, their predicted ultimate strengths were compared with the measured ones. For this purpose, three key performance indices including the coefficient of correlation (R), root mean square error (RMSE), and mean square error (MSE) were used. These indices are described in Section 5. The model performance indices are summarized in Table 9.

It can be observed that the neural network model with minimum RMSE (24.41) predicts the ultimate strength of shale with respect to SVM, multiple linear regression, and two conventional failure criteria. For more comparison, the accuracy of different failure criteria can be made by considering their predicted values for uniaxial compressive strength. The measured values of uniaxial compressive strength with the corresponding predicted value using SVM, multiple linear regression two conventional criteria, and ANN-based criteria are shown in

Table 10. However, from the results obtained, it was found that the neural network had the highest precision in predicting the shale rock strength with respect to the failure criteria.

**11. Further examination of ANN model**

In the preceding sections, the accuracy of ANN model (the same model in Section 7) in prediction of ultimate strength of shale under triaxial and polyaxial state of stress is considered. For this, multi-triaxial tests were performed in the confining pressures of 17, 25, 30, 33, 38, 40 Mpa

on shale samples in the Ramshire oilfield in Iran. The results were compared with the ultimate strength predicted value using the ANN, SVM, multiple linear regression models, and the conventional failure criteria based on the same data used in the previous sections. Table 11 shows the predicted values of ultimate strength of every model. It was seen that the ANN model predicted the ultimate strength of shale with minimum RMSE (equal to 43.81).

**Table 9. Performance indices of models.**

Models	R	R <sup>2</sup>	RMSE	MSE
Drucker-Prager	0.95	0.90	36.66	1343.95
Mogi-Coulomb	0.94	0.88	39.73	1578.47
ANN	0.97	0.94	24.41	596.10
SVM	0.96	0.92	32.54	10568.85
MLR	0.95	0.90	36.28	1207.56

**Table 10. Measured and predicted values of uniaxial compressive strength in Mpa and associated error values in percent.**

Type	Experimental	Drucker-Prager		Mogi-Coulomb		ANN		SVM		MLR	
		Predict	Error	Predict	Predict	Predict	Error	Predict	Error	Predict	Error
UCS	50.80	74.49	46.63	94.61	86.24	45.70	10.03	62.10	22.24	86.90	71.06

**Table 11. Comparison of ultimate strength prediction in different confining pressures using ANN model, SVM, multiple linear regression, and conventional failure criteria for Ramshire oilfield samples.**

Confining pressure (Mpa)	Experimental (Mpa)	Ultimate strength prediction (Mpa)				
		Drucker-Prager	Mogi-Coulomb	ANN	SVM	MLR
15	161.20	118.05	121.40	137.38	134.20	116.54
25	207.70	134.80	138.36	174.38	167.61	136.30
30	244.00	145.27	148.96	188.76	181.34	146.18
33	224.00	151.55	155.32	194.66	194.77	152.10
38	236.00	162.02	165.92	201.21	198.38	162.00
40	256.00	166.21	170.16	203.00	200.12	165.94
<b>RMSE</b>	-	<b>79.56</b>	<b>75.89</b>	<b>43.81</b>	<b>44.08</b>	<b>79.30</b>

**12. Application of ANN model in determining of threshold broken pressure in Gachsaran oilfield**

A number of effective parameters of the drilling rate are as follow: load on bit, rotation, depth, rock strength, bit condition, formation pressure, mud weight, mud type, bit diameter, and mud flow rate [40]. One of the main aspects of the drilling operation is ultimate strength of rock or break point of rock in drilling depth [58, 50]. In this section, the break point of the shale layer is determined using the ANN-based failure criteria (the same model in Section 7) in the Gachsaran oilfield in SW Iran.

Shale layer has 141 m thickness in the Gachsaran oilfield (depth 2248-2589 m). Vertical stress was calculated at the beginning and ending layer using Eq. (18).

$$\sigma_v = \frac{\rho g H}{1000} \tag{18}$$

where H is the layer of depth, g is the acceleration of gravity,  $\rho$  is the layer density (average density was assumed  $2.7 \frac{Kg}{Cm^3}$ ), and  $\sigma_v$  is the vertical stress in Mpa.

Vertical stress is 60.61, 69.90, corresponding to the beginning and ending layer. Minimum horizontal stress is minimum stress in the normal stress region. It was calculated using Eq. (19) [37, 52].

$$\begin{cases} \sigma_h = K_0 \sigma_v \\ k_0 = \frac{\gamma}{1-\gamma} \end{cases} \quad (19)$$

where  $\sigma_h$  is the minimum horizontal stress,  $k_0$  is the effective of stress ratio, and  $\gamma$  is the Poisson's ratio.

The effective of stress ratio changes about 0-0.50. If the Poisson's ratio is assumed 0.25 for the shale layer,  $k_0$  is obtained to be 0.33. Therefore, the

confining pressure is 20.20, 23.30 Mpa, corresponding to the beginning and ending of layer. Now for drilling operation, threshold broken pressure ( $\sigma_s$ ) is estimated in the 20.20 and 23.30 Mpa confining pressures using the ANN model. The results show that the threshold broken pressure is 154.21, 167.98, corresponding to the beginning (depth 2248 m) and ending (depth 2589 m) of layer. The results obtained are summarized in Table 12.

**Table 12. Vertical stress, minimum horizontal stress, and threshold broken pressure at beginning and ending of shale layer.**

Depth (m)	$\sigma_v$ (Mpa)	$\sigma_h$ (Mpa)	$\sigma_s$ (Mpa)
2248	60.61	20.20	154.21
2589	69.90	23.30	167.98

### 13. Conclusions

In this research work, a neural network was utilized for predicting the shale ultimate strength in the condition of conventional triaxial and polyaxial stresses. The database is comprised of 83 collected data from various sources.

The best model for predicting the shale ultimate strength is the multi-layer perceptron network with sigmoid activating function for the hidden layer, linear function for the output layer, and error back-propagation training rule (Levenberg-Marquardt) with the 2-4-1 architecture. This model calculates the shale ultimate strength with the MSE and RMSE values of 596.10 and 24.41, respectively. Also the neural network has the highest precision in predicting the uniaxial strength (UCS) shale.

The coefficient of correlation values obtained from the neural network model are greater than those of SVM, multiple linear regression, and conventional failure criteria. Also the corresponding MSE and RMSE values of the neural network model were smaller than the index values of the other models, and this reveals the high efficiency of the neural network model with respect to the investigated models.

For further examination, the results of the multi-stage triaxial tests performed on the specimens produced from the Ramshire oil fields in Iran were compared with the ANN model, SVM, multiple linear regression, and conventional failure criteria ultimate strength prediction. It was found that the ANN model predicted the ultimate strength with a higher accuracy, RMSE 43.81.

The threshold broken pressure was estimated at the beginning and ending layer using the ANN model. One of the main aspects of the drilling operation is the ultimate strength of rock or break point of rock in drilling depth. The threshold broken pressure or break point of rock was determined using the ANN-based Failure criteria for the Gahsaran oilfield in Iran. The results obtained showed that the threshold broken pressure was 154.21, 167.98, corresponding to the beginning (depth 2248 m) and ending (depth 2589 m) of layer.

### Acknowledgments

We wish to acknowledge the favor and cooperation of the National Iranian South Oil Company (NISOC) authorities, especially Mr. Baghdashtaki, who helped us in the preparation of the deep shale specimens.

### References

- [1]. Van Oort, E. (2003). On the physical and chemical stability of shales. *Journal of Petroleum Science and Engineering*. 38: 213-235.
- [2]. Zhang, J. (2013). Borehole stability analysis accounting for anisotropies in drilling to weak bedding planes. *International Journal of Rock Mechanics & Mining Sciences*. 60: 160-170.
- [3]. Chen, X. and Tan, C. (2000). Numerical evaluation of the deformation behavior of thick-walled hollow cylinders of shale. *International Journal of Rock Mechanics and Mining Sciences*. 37: 947-961.
- [4]. Mirzabozorgnali, A., Fathianpour, N., Ghasemzadeh, H. and Salarian, M. (2011). A New Approach in Casing Collapse Design Using the

Geomechanical Model and Heaviest Drilling Fluid. *Petroleum Science and Technology*. 29: 1948-1962.

[5]. Hoek, E. and Brown, E.T. (1980). *Underground excavations in rock*. London. Inst Min Metall.

[6]. Mohr, O. (1900). Welche umstände bedingen die elastizitätsgrenze und den bruch eines materials?. *VDI-eitschrift*. 44: 1524.

[7]. Al-Ajmi, A.M. and Zimmerman, R.W. (2005). Relation between the Mogi and the Coulomb failure criteria. *Int J Rock Mech Min Sci*. 42: 431-439.

[8]. Drucker, D. and Prager, W. (1952). Soil mechanics and plastic analysis or limit design. *Q Appl Math*. 10: 56-157.

[9]. Bieniawski, Z.T. (1974). Estimating the strength of rock materials. *J. S. Afr. Inst. Min. Metall*. 74: 312-320.

[10]. Lade, P.V. (1993). Rock strength criteria: the theories and the evidence. In Hudson J., Brown ET (eds). *Comprehensive rock engineering*. Elsevier Inc. London. pp. 225-284

[11]. Ramamurthy, T., Rao, G.V. and Rao, K.S. (1985). A strength criterion for rocks. In: *Proceedings of the Indian Geotechnical Conference*. Roorkee. 1: 59-64.

[12]. Rafiai, H., Jafari, A. and Mahmoudi, A. (2013). Application of ANN-based failure criteria to rocks under polyaxial stress conditions. *International Journal of Rock Mechanics & Mining Sciences*. 59: 42-49.

[13]. Dehghan, S., Sattari, G.H., Chehreh Chelgani, S. and Aliabadi, M.A. (2010). Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using regression and artificial neural networks. *Mini Sci Technol*. 20 (1): 0041-0046

[14]. Majdi, A. and Rezaei, M. (2013). Prediction of unconfined compressive strength of rock surrounding a roadway using artificial neural network. *Neural Comput & Applic*. 23: 381-389.

[15]. Singh, V.K., Singh, D. and Singh, T.N. (2001). Prediction of strength properties of some schistose rocks from petrographic properties using artificial neural networks. *Int J Rock Mech Min Sci*. 38 (2): 269-284.

[16]. Gholaminejad, J., Bahaaddini, H. and Rastegar, M. (2013). Prediction of the deformation modulus of rock masses using artificial neural networks and regression methods. *Journal of mining & Environment*. 4 (1): 35-43.

[17]. Sirat, M. and Talbot, C.J. (2001). Application of artificial neural networks to fracture analysis at the Äspö HRL, Sweden: fracture sets classification. *International Journal of Rock Mechanics and Mining Sciences*. 38 (5): 621-639.

[18]. Sonmez, H., Gokceoglu, C., Nefeslioglu, H.A. and Kayabasic, A. (2006). Estimation of rock modulus:

for intact rocks with an artificial neural network and for rock masses with a new empirical equation. *International Journal of Rock Mechanics and Mining Sciences*. 43 (2): 224-235.

[19]. Mohammadi, H. and Rahmancejad, R. (2010). The estimation of rock mass deformation modulus using regression and artificial neural networks analysis. *The Arabic Journal of Science and Engineering*. 35 (1A): 205-217.

[20]. Majdi, A. and Beiki, M. (2010). Evolving neural network using a genetic algorithm for predicting the deformation modulus of rock masses. *International Journal of Rock Mechanics and Mining Sciences*. 47 (2): 246-253.

[21]. Beiki, M., Bashari, A. and Majdi, A. (2010). Genetic programming approach for estimating the deformation modulus of rock mass using sensitivity analysis by neural network. *International Journal of Rock Mechanics and Mining Sciences*. 47 (7): 1091-1103.

[22]. Rafiai, H. and Jafari, A. (2011). Artificial neural networks as a basis for new generation of rock failure criteria. *International Journal of Rock Mechanics and Mining Sciences*. 48 (7): 1153-1159.

[23]. Ceryan, N., Okkan, U. and Kesimal, A. (2012). Prediction of unconfined compressive strength of carbonate rocks using artificial neural networks. *Environ Earth Sci*. 68 (3): 807-819.

[24]. Rabbani, E., Sharif, F., Koolivand, M. and Moradzadeh, A. (2012). Application of neural network technique for prediction of uniaxial compressive strength using reservoir formation properties. *Int J Rock Mech Min Sci*. 56: 100-111.

[25]. Tonnizam, E., Jahed, D. and Momeni, E. (2015). Prediction of the unconfined compressive strength of soft rocks: a PSO-based ANN approach. *Bulletin of Engineering Geology and the Environment*. 74 (3): 745-757.

[26]. Barzegar, R., Sattarpour, M., Nikudel, M. and Moghaddam, A. (2016). Comparative evaluation of artificial intelligence models for prediction of uniaxial compressive strength of travertine rocks, Case study: Azarshahr area, NW Iran. *Modeling Earth Systems and Environment*. 2 (2): 76.

[27]. Momeni, E., Jahed, D., Hajihassani, M. and Mohad, M. (2015). Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks. *Measurement*. 60: 50-63.

[28]. Asadi, M. (2016). Optimized Mamdani fuzzy models for predicting the strength of intact rocks and anisotropic rock masses. *Journal of Rock Mechanics and Geotechnical Engineering*. 8 (2): 218-224.

- [29]. Madhubabu, N., Singh, P., Kainthola, A., Mahanta, B. and Tripathy, A. (2016). Measurement. Vol. 88. pp. 202-213.
- [30]. Meulenkamp, F. and Grima, M. (1999). Application of neural networks for the prediction of the unconfined compressive strength from Equotip hardness. *International Journal of Rock Mechanics and Mining Sciences*. 36 (1): 29-39.
- [31]. Anakci, H. and Pala, M. (2007). Tensile strength of basalt from a neural network. *Engineering Geology*. 94 (1): 10-18.
- [32]. Tiryaki, B. (2008). Predicting intact rock strength for mechanical excavation using multi variate statistics, artificial neural networks, and regression trees. *Engineering Geology*. 99 (1): 51-60.
- [33]. Zorlu, K., Gokceoglu, C., Ocakoglu, F., Nefeslioglu, H. and Acikalin, C. (2008). Prediction of uniaxial compressive strength of sandstones using petrography-based models. *Engineering Geology*. 96 (3): 141-158.
- [34]. Rafiai, H. and Jafari, A. (2011). Implementation of ANN-based rock failure criteria in numerical simulations. in *Proceedings of the 12<sup>th</sup> International Congress on Rock Mechanics of the International Society for Rock Mechanics (ISRM '11)*. Beijing, China. pp. 501-506.
- [35]. Kaunda, R. (2014). New artificial neural networks for true triaxial stress state analysis and demonstration of intermediate principal stress effects on intact rock strength. *Journal of Rock Mechanics and Geotechnical Engineering*. 6 (4): 338-347.
- [36]. Asteris, P.G. and Plevris, V. (2017). Anisotropic masonry failure criterion using artificial neural networks. *Neural Computing and Applications*. 28 (8): 2207-2229.
- [37]. Zhu, C.H., Zhao, H. and Ru, Z.H. (2015). LSSVM-Based rock failure criterion and its application in numerical simulation. *Mathematical Problem in Engineering*. Article ID 246068, 13 P.
- [38]. Cannady, J. (1998). Artificial neural networks for misuse detection. In *National information systems security conference*. pp. 368-81.
- [39]. Topcu, I. and Saridemir, M. (2008). Prediction of Compressive Strength of Concrete Containing Fly Ash Using Artificial Neural Networks and Fuzzy Logic. *Computational Materials Science*. 41 (3): 305-311.
- [40]. Ince, R. (2010). Artificial Neural Network Based Analysis of Effective Crack Model in Concrete Fracture. *Fatigue & Fracture of Engineering Materials & Structures*. 33 (9): 595-606.
- [41]. Bahrami, S. and Dulati Ardrjani, F. (2015). Application of artificial neural network and genetic algorithm to modelling the groundwater inflow to an advancing open pit mine. *Journal of mining & Environment*. 6 (1): 21-30.
- [42]. Topcu, I., Karakurt, C. and Saridemir, M. (2008). Predicting the Strength Development of Cements Produced with Different Pozzolans by Neural Network and Fuzzy Logic. *Materials & Design*. 29 (10): 1986-1991.
- [43]. Bodaghi, A., Ansari, H. and Gholami, M. (2015). Optimized support vector regression for drilling rate of penetration estimation. *Open Geosci*. 1: 870-879.
- [44]. Enayatollahi, I., Agajani, A. and Asadi, A. (2014). Comparison Between Neural Networks and Multiple Regression Analysis to Predict Rock Fragmentation in Open-Pit Mines. *Rock Mech Rock Eng*. 47: 799-807.
- [45]. Smith, P.C. (1969). The measurement of satisfaction in work and retirement: A strategy for the study of attitudes.
- [46]. Gholami, R., Moradzadeh, A., Rasouli, V. and Hanachi, J. (2014). Practical application of failure criteria in determining safe mud weight windows in drilling operations. *Journal of Rock Mechanics & Geotechnical Engineering*. 6: 13-25.
- [47]. Mert, E., Yilmaz, S. and Inal, M. (2011). An assessment of total RMR classification system using unified simulation model based on artificial neural networks. *Neural Comput Appl*. 20: 603-610.
- [48]. Afshin, H., Chupani, N. and Fathipur, H. (2011). Estimating Fracture Energy of Concrete Using Artificial Neural Networks. *Journal of Civil & Environmental Engineering*. 42 (1): 2-9.
- [49]. Carter, B., Duncan, E. and Lljatj, E. (1991). Fitting strength criteria to intact rock. *Geotech and geological engineering*. 9: 73-81.
- [50]. Bieniawski, Z. and Denkhaus, H. (1967). Personal communication from M.A. Kwasniewski. Gliwice.
- [51]. Handin, J., Hager, R., Friedman, M. and Feacher, J. (1963). Experimental deformation of sedimentary rocks under confining pressure: Pore pressure tests. *Bull. Amer. Assoc. Pet. Geologists*. 47: 717-755.
- [52]. Hanna, K., Conner, D. and Hharamy, K. (1991). Coal mine entry intersection behavior study. *USBM Rep. Inv*, 9337.
- [53]. Sheorey, P.R. (1997). Empirical rock failure criteria. Rotterdam. Balkema.
- [54]. Alimoradi, A., Moradzadeh, A. and Bakhtiari, M. (2013). Application of Artificial Neural Networks and Support Vector Machines for carbonate pores size estimation from 3D seismic data. *Journal of Mining & Environment*. 4 (1): 1-14.
- [55]. Jain, S. and Bhatia, M.P.S. (2013). Performance investigation of support vector regression using

meteorological data. International journal of database theory and application. 6 (4): 109-118.

[56]. Sonebi, M., Cevik, A., Grunewald, S. and Walraven, J. (2016). Modelling the fresh properties of self-compacting concrete using support vector machine approach. Constr Build Mater. 106: 55-64.

[57]. Al-Anazi, A.F. and Gates, I.D. (2010). Support vector regression for porosity prediction in a

heterogeneous reservoir: a comparative study. Comput Geosci. 36: 1494-1503.

[58]. Cheraghi, M. and Ehteshami, P. (2013). Estimating the drilling rate in Ahvaz oilfield. Journal of petroleum exploration and production technology. 3: 167-173.

## پیش‌بینی مقاومت نهایی شیل با استفاده از شبکه عصبی مصنوعی

سهیل مشرفی<sup>۱</sup>، کوروش شهریار<sup>۲\*</sup>، احمد رمضان‌زاده<sup>۳</sup> و کامران گشتاسبی<sup>۴</sup>

۱- گروه مهندسی معدن، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

۲- دانشکده مهندسی معدن و متالورژی، دانشگاه صنعتی امیرکبیر، ایران

۳- دانشکده مهندسی معدن، نفت و ژئوفیزیک، دانشگاه صنعتی شاهرود، ایران

۴- بخش مهندسی معدن، دانشکده فنی مهندسی، دانشگاه تربیت مدرس، ایران

ارسال ۲۵/۵/۲۰۱۷، پذیرش ۱۴/۹/۲۰۱۷

\* نویسنده مسئول مکاتبات: k.shahriar@aut.ac.ir

### چکیده:

معیار شکست سنگ برای پیش‌بینی مقاومت نهایی در علم مکانیک سنگ و ژئوتکنیک بسیار مهم است. مقاومت نهایی سنگ برای مطالعات مکانیک سنگی در پروژه‌های معدنی، عمرانی، عملیات حفاری چاه نفت و غیره مورد نیاز است. همچنین شیل‌ها مشکل‌سازترین سازندهایی هستند که در سازه‌های سنگی با آن‌ها سروکار داریم؛ بنابراین در این تحقیق یک مدل برای پیش‌بینی مقاومت نهایی شیل با استفاده از شبکه عصبی مصنوعی ساخته شد. این مدل، با مدل‌های ماشین بردار پشتیبان، رگرسیون چند متغیره خطی و معیارهای چند محوره پرکاربرد در تحلیل ناپایداری ساختارهای سنگی مانند دراگر- پراگر و موگی- کلمب مقایسه شد. برای ساخت مدل از نتایج آزمایش‌های سه و چند محوره انجام شده بر روی نمونه‌های شیل توسط محققین مختلف که در مقاله‌های معتبر چاپ شده است، استفاده شد. نتیجه‌ها نشان داد که مدل شبکه عصبی پرسپترون چند لایه پیش‌خور با الگوریتم آموزش لونبرگ- مارکوت و معماری ۱-۴-۲ با مجذور میانگین خطای ۲۴/۴۱ کمترین خطا را در پیش‌بینی مقاومت نهایی در مقایسه با سایر مدل‌های شیل نشان می‌دهد. همچنین برای اعتبار سنجی بیشتر، آزمایش‌های سه محوره بر روی نمونه‌های عمقی شیل انجام شد. نمونه‌ها از میدان رامشیر واقع در جنوب غربی ایران تهیه شد. نتایج به دست آمده با پیش‌بینی مقاومت نهایی شیل توسط مدل‌های شبکه عصبی، ماشین بردار پشتیبان، رگرسیون چند متغیره خطی و معیارهای متداول شکست مقایسه شد. نتایج نشان داد مدل شبکه عصبی مقاومت نهایی شیل را با کمترین خطا ( $RMSE=43/81$ ) پیش‌بینی می‌کند. سپس مدل برای پیش‌بینی فشار آستانه شکست لایه شیل در میدان نفتی گچساران در ایران مورد استفاده قرار گرفت. برای این کار تنش‌های قائم و افقی بر اساس عمق لایه شیل محاسبه شد. فشار آستانه شکست برای ابتدا و انتهای لایه شیل برابر ۱۵۴/۲۱ و ۱۶۷/۹۸ مگا پاسکال محاسبه شد.

**کلمات کلیدی:** شبکه عصبی، معیار شکست، شیل، مقاومت نهایی، ماشین بردار پشتیبان.