Research Article

Intelligently Predict the Rock Joint Shear Strength Using the Support Vector Regression and Firefly Algorithm

Jiandong Huang, Jia Zhang, and Yuan Gao

1School of Mines, China University of Mining and Technology, 221116 Xuzhou, China
2School of Chemical Engineering and Technology, China University of Mining and Technology, 221116 Xuzhou, China

Correspondence should be addressed to Yuan Gao; y.gao@cumt.edu.cn

Received 16 June 2021; Accepted 3 August 2021; Published 24 August 2021

Academic Editor: Hao Liu

Copyright © 2021 Jiandong Huang et al. Exclusive Licensee GeoScienceWorld. Distributed under a Creative Commons Attribution License (CC BY 4.0).

To propose an effective and reasonable excavation plan for rock joints to control the overall stability of the surrounding rock mass and predict and prevent engineering disasters, this study is aimed at predicting the rock joint shear strength using the combined algorithm by the support vector regression (SVR) and firefly algorithm (FA). The dataset of rock joint shear strength collected was employed as the output of the prediction, using the joint roughness coefficient (JRC), uniaxial compressive strength ($\sigma_c$), normal stress ($\sigma_n$), and basic friction angle ($\phi_b$) as the input for the machine learning. Based on the database of rock joint shear strength, the training subset and test subset for machine learning processes are developed to realize the prediction and evaluation processes. The results showed that the FA algorithm can adjust the hyperparameters effectively and accurately, obtaining the optimized SVR model to complete the prediction of rock joint shear strength. For the testing results, the developed model was able to obtain values of 0.9825 and 0.2334 for the coefficient of determination and root-mean-square error, showing the good applicability of the SVR-FA model to establish the nonlinear relationship between the input variables and the rock joint shear strength. Results of the importance scores showed that $\sigma_n$ is the most important factor that affects the rock joint shear strength while $\sigma_c$ has the least significant effect. As a factor influencing the shear stiffness from the perspective of physical appearance, the change of the JRC value has a significant impact on the rock joint shear strength.

1. Introduction

It often appeared in mining engineering construction the issues to deal with the rock joint, which is heterogeneous and anisotropic and often contains a large number of faults, bedding, structural cracks, muddy interlayer, and other geological structural planes of different genetic types [1–10]. The strength, deformation, and stability of the rock mass are mainly controlled by the joint surface [11–14]. As a natural defect formed in the long geological process, joints exist in the rock mass with complex geometric shapes, scales, and filling forms, which seriously affect and restrict the mechanical behavior and deformation, and failure characteristics of the rock mass [15–20]. Large-scale rock mass-engineering disasters are mainly caused by the expansion, evolution, and penetration of existing joints [4, 12, 21]. The instability of the engineering rock mass is closely related to the development degree, development location, occurrence, combination characteristics, and engineering properties of the joint fissures [7, 22–28]. Regardless of the hardness of the rock, as long as there is a weak geological interface formed by unfavorable joint cracks in the rock mass, the rock mass may deform and fail along these joint cracks [3, 17]. Therefore, to propose an effective and reasonable excavation plan for rock joints and support design parameters to control the overall stability of the surrounding rock mass and predict and prevent engineering disasters, it is necessary to accurately evaluate the rock joint shear strength ($\tau_p$) [2, 5, 12, 13].

To address this issue, many researchers have been committed to exploring effective methods/models for predicting the shear strength of rock joints with high accuracy [18, 29–36]. Based on the experimental tests, Barton and Choubey proposed the joint roughness coefficient (JRC) as the indicator to describe the surface morphology of the rock joint and developed the JRC-JCS equation for estimating the strength of joints [37]. Since then, many criteria for modeling
the shear strength of rock joints have been introduced in the past studies, which can more truly describe the roughness of the joint surface and its degradation [6, 7, 11, 21, 22, 29, 32, 38]. Grasselli et al. researched the three-dimensional roughness characteristics and shear mechanical properties of rock joints and used three-dimensional roughness indexes to characterize the morphological characteristics of joints [22, 38, 39]. Xie et al. used the fractal theory to describe the roughness of rock joints and analyzed the influence of fractal parameters on shear mechanical behavior [40, 41]. Kulatilake et al. studied the fractal parameters that can describe the anisotropic roughness and proposed the shear strength equation for the prediction [21]. On the other hand, a lot of studies have been carried out in the past on the deformation and strength characteristics of rock joints in the shear process. Patton proposed a bilinear shear strength model by considering the dilatancy effect and introducing the Mohr-Coulomb equation [42]. Ladanyi and Archambault studied the interlocking effect of serrated joints and proposed a nonlinear shear strength model for joints [43]. Zhao introduced the joint coincidence coefficient joint matching coefficient (JMC) to describe the coincidence degree of joint upper and lower plates in the shear process, thus establishing the JRC-JMC shear strength model [44]. It can be summarized that these earlier modeling methods contributed an in-depth understanding of the shear strength of rock joints as well as the parameters that may affect them [22, 29, 45, 46], although these standards follow the same method, which is to build a regression model based on the results obtained from a series of experiments [46]. With the understanding of all aspects of the shear strength of rock joints, a variety of methods can be tried to simulate the nonlinear shear behavior of rock joints, while the nonlinear model with high flexibility conditions cannot provide accurate results for the estimation of the nonlinear behavior of rock joint shear characteristics [47–52].

To address the problems behind traditional methods, recent research has focused more on the use of computational optimization methods to predict rock joint shear strength, such as artificial neural networks (ANN) [53–69]. Support vector regression (SVR) is one of the commonly used machine learning (ML) methods, and it has been widely used in data mining due to its good generalization ability and fast computing power [62, 70, 71]. Many researchers have pointed out that the SVR model is more efficient and accurate in some situations than other AI-based models [72–74]. The SVR model was found to be able to solve classification and linear or nonlinear regression problems. At the same time, the SVR model can draw linear regression functions for the new high-dimensional feature space, while reducing the complexity of the model [75, 76]. Chen et al. have developed a new design of evolutionary hybrid optimization of the SVR model in predicting the blast-induced ground vibration [77]. Xu et al. employed the evolving SVR by the grey wolf optimization to predict the geomechanical properties of rock [78]. Fattahi applied the improved SVR model to predict the deformation modulus of the rock mass [79]. It should be noted that the higher effectiveness of the SVR model for prediction has been proved in the previous studies, but a major drawback is that the performance of SVR is highly dependent on its hyperparameters, which are difficult to adjust with traditional optimization methods [58, 69]. Usually, the hyperparameters of the SVR model are adjusted by the grid search method, which is very complex and can only be used to solve problems with few parameters [53, 72, 73, 80]. To solve this problem, naturally inspired optimization algorithms have been applied in this field, including the particle swarm optimization (PSO) [81, 82], genetic algorithm (GA) [83, 84], and firefly algorithm (FA) [85, 86]. Among these optimization algorithms, FA is widely used in the optimization of ML model hyperparameters due to its advantages of reducing multimode and automatic subdivision [86, 87]. The FA has been used by Wang et al. to simulate the compressive strength of the cemented tailings backfill, and the comparison results were obtained to indicate the good predictive performance [88]. Chahnasir et al. have applied the support vector machine with FA for investigation of the factors affecting the shear strength of angle shear connectors [86]. The implementation of the FA is based on attractiveness that decreases as distance increases, resulting that the whole population can be automatically subdivided into subgroups, each subgroup can achieve local optimal around each mode, and the global optimal solution can be found in these local optimal solutions. Therefore, the efficient and accurate FA is used in this study to search the optimal hyperparameters of SVM.

2. Research Objective

The main research objective of the present study is to predict the rock joint shear strength using the combined algorithm by the SVR and FA. The dataset of rock joint shear strength collected from the earlier studies was employed as the input parameter of the prediction. The physical and mechanical properties of the rock joint, including the JRC, uniaxial compressive strength (σc), normal stress (σn), and basic friction angle (φb), which have been proved to show significant influences on the rock joint shear strength were employed as the input parameters for the machine learning. The data were randomly divided into two parts, in which 70% of the data were used for model training and the remaining 30% for evaluation of predicted performance. Based on the database of rock joint shear strength, the training subset and test subset for machine learning processes are developed, realizing the prediction and evaluation processes. The predicted results can be used for future mining construction and other engineering applications to reduce the time-consuming tests and improve the accuracy of the predicted results.

3. Methodology

3.1. Dataset Collection. The dataset employed to predict the rock joint shear strength was collected from the earlier study by Babanouri and Fattahi [89], including 84 groups of rock joint shear strength data with different JRC, σc, σn, and φb, respectively. The shear strength was performed for the natural rock joint samples with varying physical and mechanical properties. First, the morphology of the rock joints was
part was molded in the cast in plaster or concrete in a silicone mold, and the other captured by silica gel molding. One part of the sample was the stucco and concrete replicas are cylinders with a diameter of 60 mm. The uniaxial compressive strength of the material was measured by laboratory tests. The plaster samples have different plaster/water ratios, and the concrete samples adopt different mixed designs, which lead to the change of mechanical properties of the copied materials. A more detailed testing process and results analysis can be found in the study [89]. Table 1 gives the brief descriptive statistics of the training and testing data-

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Variables</th>
<th>Mean</th>
<th>STD</th>
<th>COV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>JRC</td>
<td>11.17</td>
<td>5.49</td>
<td>0.52</td>
</tr>
<tr>
<td>Training dataset (the number of data points is 60)</td>
<td>σc (MPa)</td>
<td>31.10</td>
<td>16.59</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>σn (MPa)</td>
<td>1.31</td>
<td>0.69</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>φθb (°)</td>
<td>29.43</td>
<td>2.63</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>τp (MPa)</td>
<td>1.05</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>Testing dataset (the number of data points is 24)</td>
<td>JRC</td>
<td>12.03</td>
<td>5.94</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>σc (MPa)</td>
<td>48.07</td>
<td>4.94</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>σn (MPa)</td>
<td>1.39</td>
<td>0.59</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>φθb (°)</td>
<td>34.15</td>
<td>2.12</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>τp (MPa)</td>
<td>1.19</td>
<td>0.63</td>
<td>0.48</td>
</tr>
</tbody>
</table>

![Figure 1: Correlation matrix of the input matrix.](image)

It can be found from Figure 1 that the correlation between the two same variables is 1 on the diagonal from the bottom left to the top right and the correlation coefficient of the part above the diagonal is symmetrical with the correlation coefficient of the part below the diagonal. The correlation coefficients between the different variables are relatively low (most values are lower than 0.5). This indicates that these input variables are independent of each other, so they can be used as input variables for intelligent prediction of the rock joint shear strength without causing multicollinearity issues.

Figure 2 gives the sensitive analysis of the input parameters on the rock joint shear strength. It can be observed from Figure 2 that all input parameters show high sensitivity to the predicted rock joint shear strength, indicating these four input parameters are effective for the subsequent machine learning process. It should be noted that the effects of the φθb and σc on the rock joint shear strength are relatively a little weak, but such effects are enough for the following intelligent prediction.

3.2 Overview of the Machines Learning Techniques

3.2.1. Support Vector Regression (SVR). Support vector regression (SVR) seeks an estimation indicator function, which can be used to classify test samples [80]. By extending the problem from seeking indication function estimation to seeking real-valued function estimation, a support vector machine (SVM) for function estimation (regression) can be obtained [74, 75, 86]. SVM effectively solves the problems of a small number of samples, high dimension, and nonlinearity [71, 72]. However, as a new machine learning algo-

The Pearson correlation coefficient was employed in this study to evaluate the relationship between the input parameters. The Pearson correlation coefficient between two input variables is defined as the quotient of covariance and standard deviation between two variables, as shown in the following equation.

$$
\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y},
$$

(1)

Correspondingly, the correlation of the collected dataset was analyzed by the SPSS software in the present study. Figure 1 gives the correlation matrix between the input variables (including the JRC, σc, σn, and φθb).
The error penalty factor $C$ is to adjust the ratio of confidence range and experience risk of machine learning in the determined data subspace so that the algorithm has the best generalization ability \[70, 76\]. The optimal $C$ is going to be different depending on the data subspace. In the determined data subspace, a small value of $C$ means that the penalty for empirical error is small, and the complexity of machine learning is small but the empirical risk value is large. However, when $C$ exceeds a certain value, the complexity of SVM reaches the maximum allowed in the sample space. At this point, the empirical risk and the ability to spread have barely changed. There is at least one optimal $C$ in each data subspace to maximize the generalization ability of SVM \[71, 74, 80\].

The kernel function parameter $Y$ affects the complexity of the distribution of sample data in the high-dimensional feature space. The change of kernel parameters implicitly changes the mapping function, thus changing the dimension of the sample space. For an indicator function set, if there are $H$ samples that can be separated by all possible forms of two to the $H$ of the function set, then the function set is said to be able to shatter $H$ samples. The Vapnik-Chervonenkis (VC) dimension of a set of functions is the maximum number of samples it can shatter. If there are functions that can shatter any number of samples, then the VC dimension of the set of functions is infinite, and the VC dimension of a bounded real function can be defined by converting it to an indicator function with a certain threshold. In the FA, the dimension of the sample space determines the maximum VC dimension of the linear classification surface that can be constructed in this space and thus determines the minimum empirical error that can be achieved by the linear classification surface \[53, 72\].

3.2.2. Firefly Algorithm (FA). FA is a new natural heuristic optimization algorithm proposed by Yang, a scholar from the University of Cambridge, in 2008 \[90\]. This algorithm simulates the luminous mechanism and behavior of tropical fireflies in nature and is a random search algorithm based
on swarm intelligence. To build the mathematical model of the FA, the following three idealized criteria should be supposed:

(1) All fireflies in the algorithm are not male or female, that is, the attraction between fireflies is only based on brightness information, without considering the influence of gender.

(2) The attraction between fireflies is proportional to the brightness. The greater the brightness, the stronger the attraction. And both will decrease as the distance increases. So for any two fireflies, the lighter one will move toward the brighter one, and the brighter one will move at random.

(3) The brightness of the firefly is related to the value of the objective function to be optimized [85, 86].

FA consists of two elements, namely brightness and attractiveness. Brightness reflects the advantages and disadvantages of the position of the firefly and determines the direction of its movement. Attraction determines the distance the firefly moves. Optimization can be achieved by updating the brightness and attraction continuously. The relative brightness of the firefly is determined as

\[ I = I_0 \cdot e^{-r_{ij}/\gamma}, \]  

where \( I_0 \) is the maximum brightness of fireflies, which is related to the objective function value. \( \gamma \) is the absorption coefficient of light intensity. \( r_{ij} \) is the Euclidean distance between fireflies \( i \) and \( j \). The attractiveness of the firefly is defined as

\[ \beta = \beta_0 \cdot e^{-r_{ij}/\gamma}, \]  

where \( \beta_0 \) is the maximum attractiveness. If the spatial positions of two fireflies \( i \) and \( j \) are, respectively, \( a \) and \( b \), then the distance between them can be expressed as

\[ r_{ij} = \| x_i - x_j \| = \sum_{k=1}^{d} \sqrt{(x_{ik} - x_{jk})^2}. \]

The position that the firefly \( i \) is attracted to move toward the firefly \( j \) is updated with the following equation.

\[ x_i = x_i + \beta(x_j - x_i) + \alpha \left( \text{rand} - \frac{1}{2} \right), \]

where \( \alpha \) is the step size factor and rand is the random value in the range of \((0,1)\).

The basic bionics principle of the FA is as follows: firstly, the points in the search space are used to simulate the individual fireflies in nature; secondly, the search and location updating processes in the optimization process were simulated as the process of mutual attraction and movement among individual fireflies in nature. The objective function value of the problem to be solved is simulated as the brightness information of the position of the individual firefly. Finally, the process of searching for the optimal feasible solution in the iterative process is simulated as the process of individual firefly survival of the fittest [90].

3.3. Hyperparameter Tuning

3.3.1. 10-Fold Crossvalidation (CV). In the present study, the 10-fold crossvalidation (CV) was employed for the hyperparameter tuning in the SVR model, which is including the error penalty factor \( C \) and kernel function parameter \( \gamma \). In the 10-fold CV, the dataset of the rock joint shear strength is divided into 10 subsets, in which the 9 subsets are used for the training process and 1 subset is employed to validate the predicted results. Regarding the 1 subset used to validate the predicted results, the minimum value of the RMSE should be determined after the 50 iterations to represent the optimized structure of the SVR model in this fold. That is to say, such a validation process should be conducted 50 times using the FA algorithm (Figure 4) to obtain the hyperparameters of the SVR model.

Finally, the optimized structure of the SVR model and the corresponding hyperparameters can be determined after the crossvalidation of 10 times. It should be noted that since there is the possibility of overfitting, the predicted performance of the SVR model is required to validate by comparing the testing dataset.

3.3.2. Evaluation of the Predictive Performance. In the present study, two parameters (RMSE (root-mean-square error) and \( R \) (correlation coefficient)) were employed to conduct the model validation process by evaluating the predictive performances of the models developed in the present study. The RMSE can be defined by the following equation:

\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i^* - y_i)^2}, \]

where \( y_i^* \) represents the predicted rock joint shear strength, \( y_i \) represents the measured value of the permeability coefficient, and \( n \) is the number of experimental samples to conduct the tests of rock joint shear strength. \( R \) is determined by the equation as follows:

\[ R = \frac{\sum_{i=1}^{n} (y_i^* - \bar{y})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (y_i^* - \bar{y})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}, \]

where the \( y_i^* \) and \( y_i \) represent the predicted and measured values of the rock joint shear strength.

4. Results and Discussion

4.1. Results of Hyperparameter Tuning. As mentioned in the previous chapter, the FA algorithm and 10-fold CV were used to adjust the hyperparameters of the SVM model to obtain the minimum value of RMSE in the optimized fold. After the 10-fold CV during the hyperparameter tuning
process, Figure 5 gives the relationship between the fold number and the values of RMSE during the 10-fold CV process.

It can be observed from Figure 5 that the minimum value of RMSE was obtained in the 1st fold during the crossvalidation process regarding the rock joint shear strength dataset. Figure 6 gives the relationship between the iteration and RMSE regarding the 1st fold.

It can be seen from Figure 6 that with the increase of the iteration, the values of RMSE gradually reduced, and especially in the first 10 iterations, there is a rapid drop in the values of RMSE. It can also be indicted that for the dataset of the rock joint shear strength, it takes 28 iterations for the RMSE curve convergence. Therefore, it is evident that the FA algorithm can adjust the hyperparameters effectively and accurately, obtaining the optimized SVR model to complete the prediction of rock joint shear strength.

4.2. Predictive Performance of the SVR-FA Algorithm. After the tuning process of the hyperparameters, the predictive ability of the SVR-FA algorithm to evaluate the rock joint shear strength was evaluated. Figures 7 and 8 show the comparison of the measured and predicted $\tau_p$ regarding the training dataset and testing dataset, respectively.
SVR-FA algorithm for the prediction of rock joint shear strength. The predictive ability of the SVR-FA model, by establishing the nonlinear relationship between the input variables (JRC, \( \sigma_n \), \( \sigma_c \), and \( \phi_b \)) and output variable, also these results showed that no issue of overfitting occurred in the developed SVR-FA model as it was used to evaluate and predict the rock joint shear strength.

4.3. Variable Importance. The relative importance of the physical (JRC) and mechanical properties (\( \sigma_n \), \( \sigma_c \), and \( \phi_b \)) related to the rock joint shear strength was also evaluated in this study. Figure 10 gives the importance scores for these variables which were used as the input for the developed SVR-FA model. A higher importance score indicates that the influence of this input variable on shear strength is more significant.

It can be seen from Figure 10 that \( \sigma_n \) is the most important factor that affects the rock joint shear strength, indicating that the normal stress has a significant effect on the rock joint shear strength. However, it is worth noting that the change in \( \sigma_c \) has the least significant effect on the rock joint shear strength. Although \( \sigma_n \) and \( \sigma_c \) are both parameters from the perspectives of the mechanical properties, their ability to influence the rock joint shear strength is indeed very different. As a factor influencing the shear stiffness from the perspective of physical appearance, it can be seen from Figure 10 that the change of the JRC values has a significant impact on the rock joint shear strength. The importance score of \( \phi_b \) is only 0.536, indicating that it is a parameter with a relatively insignificant effect on shear strength.

5. Conclusions and Future Developments

This study is to predict the rock joint shear strength using the combined algorithm by the SVR and FA for future mining construction and other engineering applications. The dataset of rock joint shear strength collected was employed as the output of the prediction, using the JRC, uniaxial compressive strength (\( \sigma_c \)), normal stress (\( \sigma_n \)), and basic friction angle (\( \phi_b \))

In the present study, 64 datasets of \( \tau_p \) were employed for the training process, and 20 datasets were used as the testing dataset for the evaluation. All the measured and predicted values as well as the differences between them (which are shown as rectangles) are shown in the figures. It can be observed from Figures 7 and 8 the good agreements between the measured and predicted \( \tau_p \) for both the training and testing datasets, except for a few noise points that make some difference, indicating a reliable relationship between the predicted and measured values in terms of the calculation of rock joint shear strength. The predictive ability of the SVR-FA algorithm for the \( \tau_p \) dataset was also assessed using relevant statistical indicators (RMSE and \( R \)). Figure 9 gives the measured and predicted \( \tau_p \) as well as the values of RMSE and \( R \).

It can be seen higher values of RMSE (0.9605 and 0.9825 for the training dataset and testing dataset, respectively) and lower values of \( R \) (0.1266 and 0.2334 for the training dataset and testing dataset, respectively) in the comparison between the predicted and measured \( \tau_p \), indicating that the SVR-FA model has a high prediction accuracy and the algorithm is suitable for the dataset of rock joint shear strength. The testing dataset can be used to verify the realistic behavior of the rock joint, and the results obtained (\( R = 0.9825 \) and RMSE = 0.2234) can indicate the accuracy of the proposed model. Also, the closed values of RMSE and \( R \) for both the training and testing dataset demonstrated the good applicability of the SVR-FA model, by establishing the nonlinear relationship between the input variables (JRC, \( \sigma_n \), \( \sigma_c \), and \( \phi_b \)) and output variable. Also, these results showed that no issue of overfitting occurred in the developed SVR-FA model as it was used to evaluate and predict the rock joint shear strength.

4.3. Variable Importance. The relative importance of the physical (JRC) and mechanical properties (\( \sigma_n \), \( \sigma_c \), and \( \phi_b \)) related to the rock joint shear strength was also evaluated in this study. Figure 10 gives the importance scores for these variables which were used as the input for the developed SVR-FA model. A higher importance score indicates that the influence of this input variable on shear strength is more significant.

It can be seen from Figure 10 that \( \sigma_n \) is the most important factor that affects the rock joint shear strength, indicating that the normal stress has a significant effect on the rock joint shear strength. However, it is worth noting that the change in \( \sigma_c \) has the least significant effect on the rock joint shear strength. Although \( \sigma_n \) and \( \sigma_c \) are both parameters from the perspectives of the mechanical properties, their ability to influence the rock joint shear strength is indeed very different. As a factor influencing the shear stiffness from the perspective of physical appearance, it can be seen from Figure 10 that the change of the JRC values has a significant impact on the rock joint shear strength. The importance score of \( \phi_b \) is only 0.536, indicating that it is a parameter with a relatively insignificant effect on shear strength.

5. Conclusions and Future Developments

This study is to predict the rock joint shear strength using the combined algorithm by the SVR and FA for future mining construction and other engineering applications. The dataset of rock joint shear strength collected was employed as the output of the prediction, using the JRC, uniaxial compressive strength (\( \sigma_c \)), normal stress (\( \sigma_n \)), and basic friction angle (\( \phi_b \))

In the present study, 64 datasets of \( \tau_p \) were employed for the training process, and 20 datasets were used as the testing dataset for the evaluation. All the measured and predicted values as well as the differences between them (which are shown as rectangles) are shown in the figures. It can be observed from Figures 7 and 8 the good agreements between the measured and predicted \( \tau_p \) for both the training and testing datasets, except for a few noise points that make some difference, indicating a reliable relationship between the predicted and measured values in terms of the calculation of rock joint shear strength. The predictive ability of the SVR-FA algorithm for the \( \tau_p \) dataset was also assessed using relevant statistical indicators (RMSE and \( R \)). Figure 9 gives the measured and predicted \( \tau_p \) as well as the values of RMSE and \( R \).

It can be seen higher values of RMSE (0.9605 and 0.9825 for the training dataset and testing dataset, respectively) and lower values of \( R \) (0.1266 and 0.2334 for the training dataset and testing dataset, respectively) in the comparison between the predicted and measured \( \tau_p \), indicating that the SVR-FA model has a high prediction accuracy and the algorithm is suitable for the dataset of rock joint shear strength. The testing dataset can be used to verify the realistic behavior of the rock joint, and the results obtained (\( R = 0.9825 \) and RMSE = 0.2234) can indicate the accuracy of the proposed model. Also, the closed values of RMSE and \( R \) for both the training and testing dataset demonstrated the good applicability of the SVR-FA model, by establishing the nonlinear relationship between the input variables (JRC, \( \sigma_n \), \( \sigma_c \), and \( \phi_b \)) and output variable. Also, these results showed that no issue of overfitting occurred in the developed SVR-FA model as it was used to evaluate and predict the rock joint shear strength.

4.3. Variable Importance. The relative importance of the physical (JRC) and mechanical properties (\( \sigma_n \), \( \sigma_c \), and \( \phi_b \)) related to the rock joint shear strength was also evaluated in this study. Figure 10 gives the importance scores for these variables which were used as the input for the developed SVR-FA model. A higher importance score indicates that the influence of this input variable on shear strength is more significant.

It can be seen from Figure 10 that \( \sigma_n \) is the most important factor that affects the rock joint shear strength, indicating that the normal stress has a significant effect on the rock joint shear strength. However, it is worth noting that the change in \( \sigma_c \) has the least significant effect on the rock joint shear strength. Although \( \sigma_n \) and \( \sigma_c \) are both parameters from the perspectives of the mechanical properties, their ability to influence the rock joint shear strength is indeed very different. As a factor influencing the shear stiffness from the perspective of physical appearance, it can be seen from Figure 10 that the change of the JRC values has a significant impact on the rock joint shear strength. The importance score of \( \phi_b \) is only 0.536, indicating that it is a parameter with a relatively insignificant effect on shear strength.

5. Conclusions and Future Developments

This study is to predict the rock joint shear strength using the combined algorithm by the SVR and FA for future mining construction and other engineering applications. The dataset of rock joint shear strength collected was employed as the output of the prediction, using the JRC, uniaxial compressive strength (\( \sigma_c \)), normal stress (\( \sigma_n \)), and basic friction angle (\( \phi_b \))
as the input for the machine learning. Based on the database of rock joint shear strength, the training subset and test subset for machine learning processes are developed, realizing the prediction and evaluation. The conclusions drew from the research process can be highlighted as follows.

1. In the hyperparameter tuning process, the FA algorithm can adjust the hyperparameters effectively and accurately, obtaining the optimized SVR model to complete the prediction of rock joint shear strength. The higher values of RMSE and lower values of $R$ in the comparison between the predicted and measured $\tau$, indicate that the SVR-FA model has a high prediction accuracy and the algorithm is suitable for the dataset of rock joint shear strength. The good applicability of the SVR-FA model is evidenced by the established nonlinear relationship between the input and output variables. Also, no issue of overfitting occurred in the developed SVR-FA model, and it can be used for an effective and reasonable excavation plan for rock joints to control the overall stability of the surrounding rock mass in the practical application.

2. Although $\sigma_n$ and $\sigma_c$ are both parameters from the perspectives of the mechanical properties, their ability to influence the rock joint shear strength is indeed very different: $\sigma_n$ is the most important factor that affects the rock joint shear strength while $\sigma_c$ has the least significant effect. As a factor influencing the shear stiffness from the perspective of physical appearance, the change of the JRC values has a significant impact on the rock joint shear strength.

Considering the limitation that only 84 groups of datasets of rock joint shear strength were employed in this study, more rock joint shear strength datasets using the experimental tests should be collected to improve the accuracy of the developed machine learning model regarding future development. A model comparison study should be carried out as well to evaluate the efficiency and accuracy of different machine learning models in the prediction of the rock joint shear strength. Also, a user-friendly tool should be developed in the future to facilitate the practical application of engineers.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This research was supported by the Fundamental Research Funds for the Central Universities (Grant No. 2021QN1006). The writers are grateful for this support.

References


[38] G. Grasselli, Shear Strength of Rock Joints Based on Quantified Surface Description, EPFL, 2001.


