IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 11, 2018, accepted November 5, 2018, date of publication November 12, 2018,
date of current version December 7, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2880466

Prediction of Pillar Stability for Underground Mines Using the Stochastic Gradient Boosting Technique

HANGXING DING[1], GUANGHUI LI[1], XIN DONG[1], AND YUN LIN[2,3], (Member, IEEE)

[1]School of Resource and Civil Engineering, Northeastern University, Shenyang 110819, China
[2]School of Resources and Safety Engineering, Central South University, Changsha 410083, China
[3]School of Civil, Environmental and Mining Engineering, The University of Adelaide, Adelaide, SA 5005, Australia

Corresponding author: Yun Lin (yunlin617@csu.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 51534003, in part by the National Key Research and Development Program of China under Grant 2016YFC0801600, and in part by the Fundamental Research Funds Project for the Central Universities of the Central South University of China under Grant 2016zzts095. The work of Y. Lin was supported by the Chinese Scholarship Council, for the joint Ph.D. studies at The University of Adelaide.

ABSTRACT The prediction of pillar stability is of great importance because pillar failure can lead to large disasters. In this paper, a stochastic gradient boosting (SGB) model was applied to classify pillar stability. Five potentially relevant factors, including the pillar width, the pillar height, the ratio of the pillar width to the pillar height, the uniaxial compressive strength of the rock, and the pillar stress, were chosen to establish the evaluation index system. The 205 pillar samples were collected, and an SGB model was developed by training 80% of original data (165 samples), and the optimal parameter values of the model were achieved by the method of 10-fold cross-validation. The external testing set (with 40 samples) was used to validate the feasibility of the SGB model. The accuracy and kappa analysis, together with the three within-class classification metrics (recall, precision, and F-measure), and receiver operating characteristic curve were utilized to evaluate the performance of the optimum SGB, random forest (RF), support vector machine (SVM), and MLPNN models. The results revealed that the SGB model has higher credibility than the RF, SVM, and MLPNN models. The sensitivity of the parameters was investigated based on the relative variable importance, in which the pillar stress and the ratio of the pillar width to the pillar height were found to be the major influencing variables for pillar stability.

INDEX TERMS Pillar stability, stochastic gradient boosting (SGB), 10-fold cross-validation, ROC curve, relative variable importance.

I. INTRODUCTION

In long-term, large-scale underground mining, a large number of goafs is one of the most important factors that endangers mine production safety, and pillars are the main structural columns that affect the stability of goafs. The instability of a pillar will lead to the collapse of the roof in the goaf area, resulting in a large number of casualties and serious property losses [1]–[3]. Therefore, it is of great significance to increase the stability of ore pillars to achieve efficient and safe mining of underground mines.

Due to the importance of pillar stability, a large number of publications on a variety of aspects of pillar stability have been published, and many valuable methods have been applied in the past several decades by many scholars for understanding and predicting pillar failures. Overall, these methods can be divided into two categories, namely, empirical methods [4], [5] and numerical simulation [6]–[8]. Empirical methods are usually used for pillar strength estimation based on empirical formulas [9]–[11]. Because it is difficult to determine the actual stress of pillars in underground mines, a factor of safety (FS) of a pillar, which is the ratio of the average strength to the average stress of the pillar, was introduced to evaluate pillar stability [6], [12]. Generally, a pillar with $FS > 1.0$ is stable, while $FS < 1.0$ means the pillar is unstable [6], [12]–[14]. With the emergence of numerical simulation software, numerical simulation technology has been gradually applied to pillar stability analysis. For example, a pillar was designed by Deng *et al.* [6] using the

combined methods of Monte-Carlo and FLAC. Using FLAC3D software, a numerical calculation model was established by Shnorhokian *et al.* [15] to analyze the safety of different pillar sizes under different stopping sequences. Mortazavi *et al.* [7] obtained the deformation and failure process of ore pillars under natural conditions and analyzed the influence of the geometry of the pillars and the mechanical parameters of the rock mass based on UDEC software. All of the studies have greatly improved our understanding of pillar stability, but they are far from solving it completely because of the following disadvantages of the above methods:

a. The stability of a pillar is affected by many parameters, and it is difficult for these two methods to take into account the influence of uncertainties.
b. The boundaries of the safety factor are indefinite and unclear. Theoretically, a pillar with FS>1.0 is stable. However, there have been failed pillars with FS>1.0 in actual engineering scenarios [6], [12].
c. Due to the nonlinear behavior of pillars with high stress in deep mines, the mechanism of pillar failure used in the existing methods is not suitable.

In recent years, data-mining techniques and intelligent evaluation models have been successfully applied to pillar stability analysis and have achieved remarkable results with the increasing availability of pillar parameters. Cauvin *et al.* [16] proposed a probabilistic model for pillar stability prediction based on the uncertainty of the data and the model. Tawadrous and Katsabanis [17] used artificial neural networks (ANN) to predict the stability of crown pillars left over large excavations, and the result showed that the cloud model is a feasible and reliable method for comprehensive stability evaluation of pillars. Logistic regression was introduced by Wattimena [14] and Wattimena *et al.* [18] to calculate the probability of pillar failure, and these authors found that it is a useful tool for analyzing pillar stability. Griffiths *et al.* [19] proposed a model for predicting the stability of pillars combined with random field theory and the Monte-Carlo method. Zhou *et al.* [20] applied supervised learning methods, such as Fisher discriminant analysis, support vector machines (SVM), random forest (RF), linear discriminant analysis (LDA) and multilayer perceptron neural networks, to predict pillar stability and then analyzed the performance of the algorithms by comparing the prediction results. All of the above intelligent models can help us to understand pillar failures, but each of them has its advantages and weaknesses [20]–[22], and none can be applicable for any engineering. Moreover, the focus of each evaluation index is different. For example, Kappa and ROC curves are more useful to unbalanced data. For pillar stability prediction, it is particularly important to predict of failure and unstable cases correctly, which address a need to introduce metrics to compare and analyze the performance of models for cases subject to different classes. Both of these address a need to propose more methods and evaluation metrics to predict pillar stability for underground mines.

Stochastic gradient boosting (SGB) is an improved boosting machine learning algorithm proposed by Friedman [23], [24] for regression and classification studies. The SGB algorithm has the ability to model nonlinear relationships and remains robust in the absence of data and outliers and is particularly suitable for processing high-dimensional data [25]. Additionally, it is not necessary to preselect or transform predictor variables [26], and the prediction accuracy can be increased based on a portion of the training data, which also helps to avoid overfitting the data. The SGB algorithm has been successfully applied in many fields, such as in studies of rock bursts [27], landslides [28], residential structure damage [29], digital imagery [26], time-series data forecasting [30], and tree species [31]. Therefore, it is of interest and motivating to predict pillar stability using the SGB model. The objective of this study is to verify the feasibility and reliability of the SGB algorithm on pillar stability prediction, compare the performance of the SGB and other three models using more evaluation indexes, and investigate the relative importance of influencing variables. It is hoped that the results can provide early warning to mine management for appropriate actions to reduce damage and save lives.

In this paper, pillar stability is predicted using the SGB model. Considering the ambiguity and uncertainty of the evaluation indexes, some typical representative indexes are selected to establish the evaluation index system, and a 10-fold CV method is introduced to optimize the parameters of the model. The performance is evaluated by comparing two global classification metrics (classification accuracy and Cohen's Kappa coefficient), three within-class classification metrics (recall, precision, and F-measure) and ROC curves of SGB, RF, SVM and MLPNN models. The sensitivity of factors is also investigated by calculating the relative variable importance.

II. DATABASE AND PARAMETERS

A. MECHANISMS OF PILLAR FAILURE

The overall response of a pillar to the loading caused by mining depends on the pillar size, the geological structure of the rock mass and the surface constraints imposed by the surrounding rock on the pillar [4], [5]. The failure modes of pillars can be divided into three types [32]: compression shear failure, tensile failure and shear failure along the weak surface. The compression shear failure is caused by a shear fracture of the pillar along the joint crack under the action of shear stress and continues to develop inside the pillar, causing the bearing area of the pillar to gradually shrink and eventually destroy the pillar. Tensile failure occurs when pillars crack with longitudinal joint fissures under the effect of transverse tensile stress. Tensile failure refers to the cracking of pillars with longitudinal joint fractures under the action of transverse tensile stress. The shear failure along the weak surface is the failure of pillars with weak interlayers caused by transverse shear stress, which is generated by the sandwich extrusion under the effect of axial stress. Overall, the stability

of pillar is closely related to geometric structures, geological structures and the redistribution of pillar stress states caused by mining activities.

B. PARAMETER ANALYSIS

To establish classification models, some pillar parameters should be used. Meanwhile, to avoid overtraining the model, three principles of choosing parameters need be relied upon [33]. Firstly, the sensitive and stable parameters reflecting properties of pillar stability should be used as the discriminant indicators. Secondly, the parameters should be physically independent of each other. Finally, the parameter data should be obtained easily or readily available. The stability of a pillar is affected by many factors and the major influencing factors are: 1) the loading of the pillar, 2) the size of the pillar, 3) the geological structure and the strength of the ore body itself. In general, the pillar width, pillar height and the ratio of width to height can reflect the size of a pillar, particularly, the ratio of pillar width to height affects the potential for violent failure of pillars [34] and is the main reference index for the design and strength estimation of pillars in underground mines. The pillar stress can be related to the stress state of the pillar, and the strength of the ore body can be determined by the uniaxial compressive strength of rock. Meanwhile, all or part of these parameters have been utilized to predict the pillar stability and strength. For example, the pillar width and depth were selected by Esterhuizen *et al.* [35] to analyze pillar stability. The study of Song *et al.* [36] showed that pillar width, mining depth and room width have the most significant influence on the stability of the pillar. Wattimena [14] and Wattimena *et al.* [18] predicted pillar stability by developing their model using the ratio of pillar width to height and the ratio of pillar strength to stress. Ghasemi *et al.* [37] developed their model using parameters including depth, pillar width, pillar length and loading condition. Zhou *et al.* [12] used five pillar parameters, such as pillar width, pillar height, the ratio of pillar width to pillar height, and the uniaxial compressive strength of rock and pillar stress. Based on all of the above analysis, five parameters, including pillar width (X_1), pillar height (X_2), the ratio of pillar width to pillar height (X_3), the uniaxial compressive strength of rock (X_4) and pillar stress (X_5), were selected in this study because they are widely used in pillar stability prediction tasks with good results and the data can be accessed easily. Note that the stability of pillar stability is also affected by other factors, such as microseismic monitoring signals and the characteristics of dynamic disturbances. But the data of the two indicators are difficult to obtain, and it can be found that the pillar stability can still be better predicted without the two indicators based on the results of [13], [14], [18], [19], and [37]. Therefore, microseismic monitoring signals and the characteristics of dynamic disturbances were not used as evaluation indicators in this study.

TABLE 1. Descriptive statistics of the five indicators for the SGB model.

Indicators	X_1 / m	X_2 / m	X_3	X_4 / MPa	X_5 / MPa
Maximum	45.00	53.00	4.57	316.00	127.60
Minimum	1.90	1.80	0.31	19.00	1.10
Mean	9.93	10.14	1.35	144.11	46.81
Standard deviation	8.01	11.50	0.77	71.26	29.93

C. CASE DATA AND PRELIMINARY ANALYSIS

To measure and compare the performance of the SGB model, 205 pillar cases (available in Table S1) from underground mines were directly collected from some publications over the period of 1972-2017. The data of the pillar cases were collected from eight different rock mines, which are: Wengfu Phosphate Mine in China [32]; Stone mines in USA [12], [35]; Westmin Resources Ltd.'s H-W mine in Canada [5]; Zinkgruvan mine in Sweden [38]; open stope mines in Canada [39]; Selebi-Phikwe mines in South Africa [40]; Elliot Lake uranium mines in Canada [41]. All data are from these literature and the sources are reliable without any processing.

A box graph of all cases is shown in Fig. 1. Apart from the ratio of pillar width to pillar height (X_3), the medians of most of the parameters are not in the center of the boxes, indicating that the distribution of these parameters is not symmetric. Meanwhile, there are several outliers with very large or small values for all parameters except for the uniaxial compressive strength of rock (X_4). The statistical features (maximum, minimum, mean and standard deviation) of the reduced data set are shown in Table 1.

There are three types of pillar stability for the dataset: stable (S, 73 cases), unstable (U, 37 cases) and failed (F, 95 cases). Fig. 2 shows the scatterplot matrix of the dataset. The distribution of each parameter is demonstrated on the diagonal, and numbers in the upper panels represent the correlation of two parameters. From Fig. 2, it can be seen that most parameters have a relatively poor correlations ($R < 0.5$) with one another [22]. We can also find that pillar width is substantially correlated with pillar height and UCS is moderately correlated with pillar stress. Additionally, it can also be seen that the dataset is quiet widely distributed and the distribution of most influencing variables is not symmetric.

The graph of pillar stability classification with respect to each distinct indicator is shown in Fig. 3. Ideally, to be easily and obviously classified, the values of all indicators should have only one class label value in the figure. It is apparent that the values of some indicators have more than one corresponding value for the same pillar stability classification in some events. This may be because the indicator values do not have apparent limits among the classes of pillar stability. It can also be concluded from Fig. 3 that none of the parameters can distinguish the stability of the pillar well.

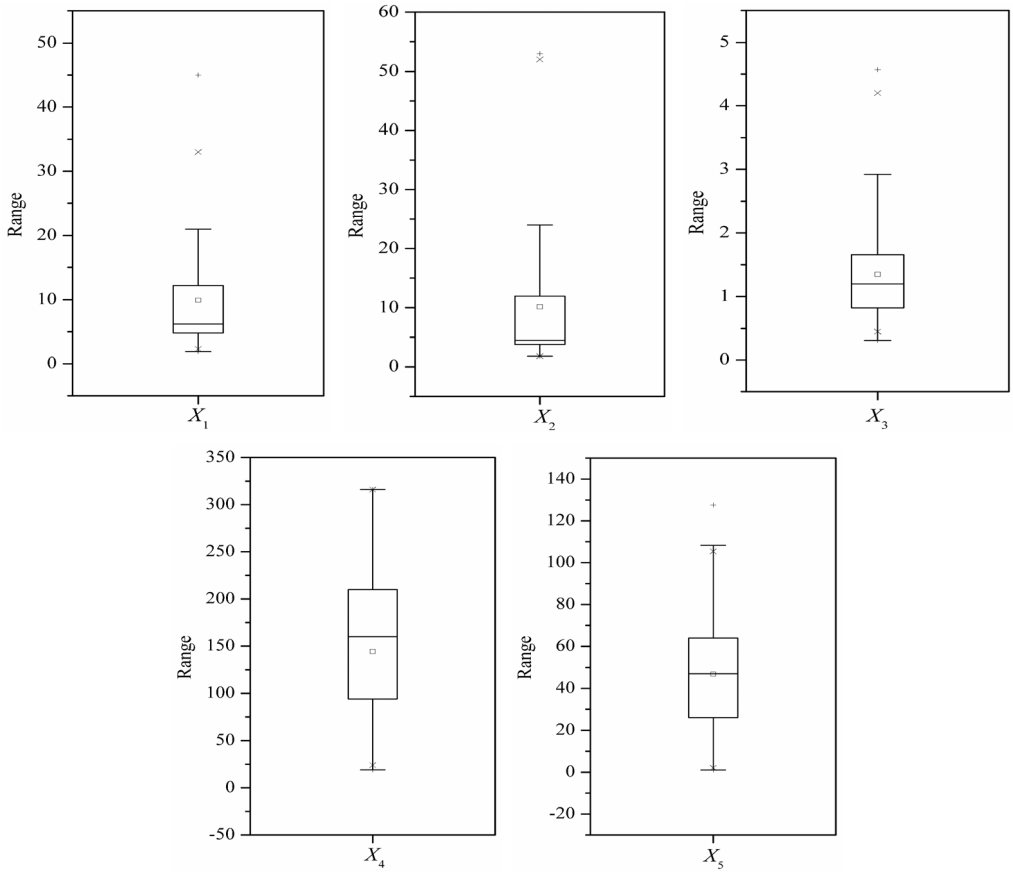

FIGURE 1. Box plot of each variable for rock burst cases.

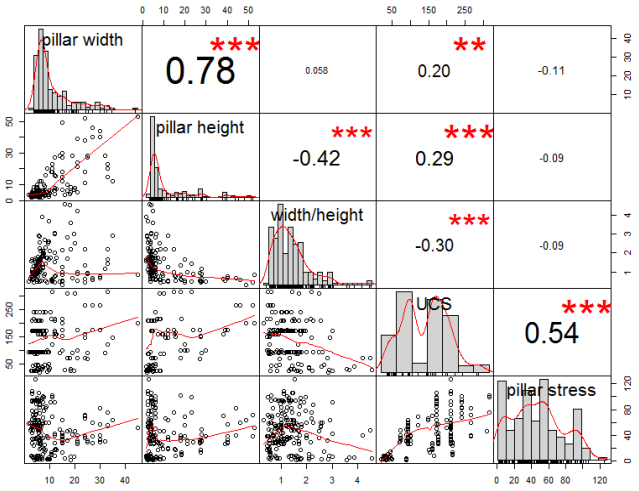

FIGURE 2. Scatter matrix of the five parameters of pillar stability.

III. METHODOLOGY

A. DETAILS OF THE SGB MODEL

The SGB algorithm was built by Friedman [24] by introducing the idea of gradient descent into the boosting algorithm. Gradient boosting is an ensemble learning algorithm combined with boosting and decision trees, and the new model is built along the gradient descent direction of the loss function of the previously established model. The essence of the SGB algorithm is to minimize the loss function between the classification function and real function by training the classification function $F*(X)$.

The distribution of the loss function is the key to the application of the SGB model [20], [24], and SGB algorithm has applicability to all loss functions. For K-class problem, surrogate loss function (multi-class log-loss) is the loss function suggested by Friedman [23], [24] and has been widely applied in many fields [20], [22], [25], [27]–[29]. The loss function can be expressed as

$$\begin{aligned}\psi(y_k, F_k(X)_1^K) &= -\sum_{k=1}^{K} y_k \log p_k(X) \\ &= -\sum_{k=1}^{K} y_k \log[\exp(F_k(X))/\sum_{l=1}^{K}\exp(F_l(X))]\end{aligned} \tag{1}$$

where $X = \{x_1, x_2, ...x_n\}$ is the input variable, k is the number of classes, y is the output variable, $p_k(X)$ is the probability.

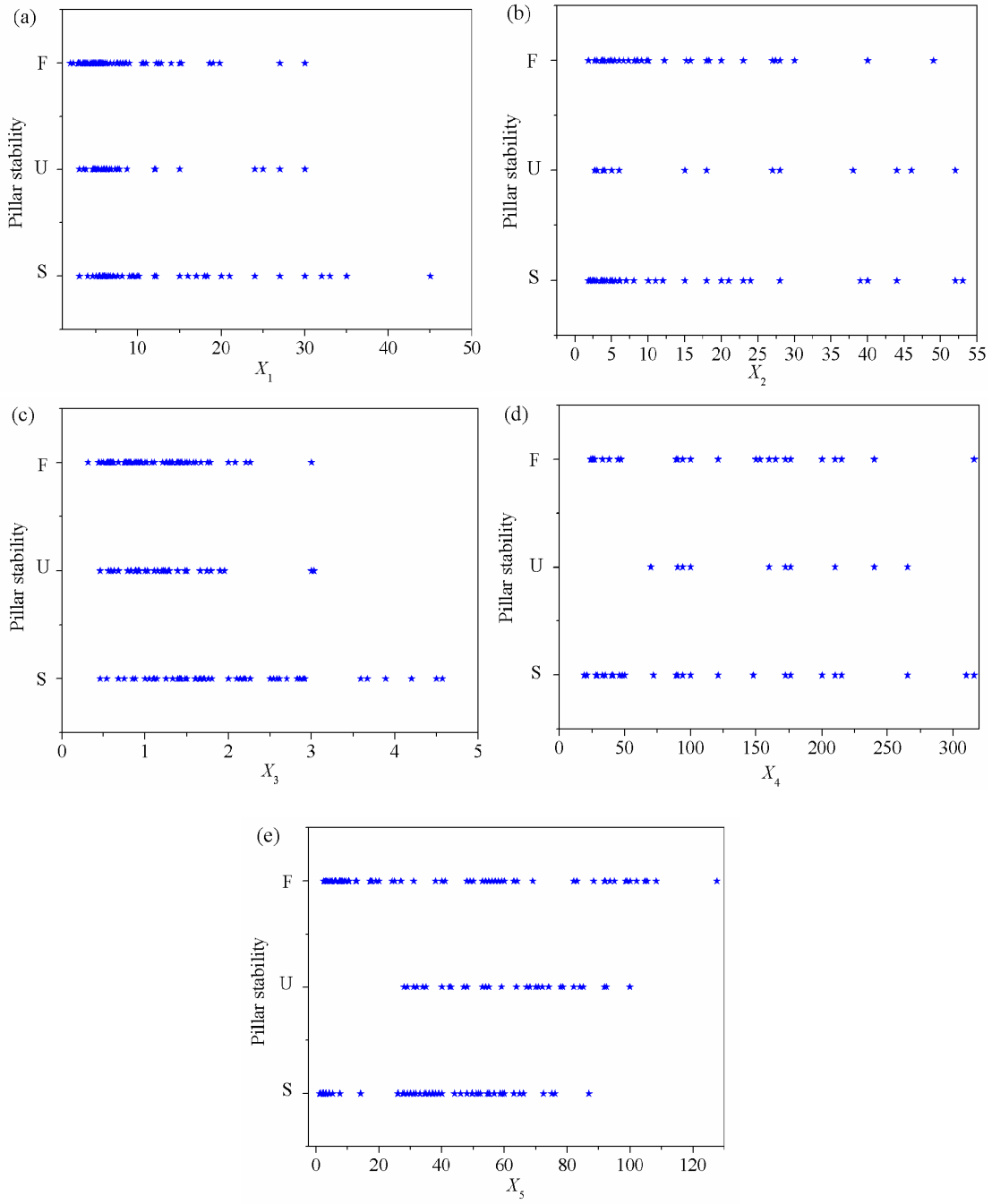

FIGURE 3. Pillar stability with respect to each indicator.

Then the following equation can be obtained

$$\begin{aligned}\tilde{y}_{im} &= -[\frac{\partial \psi(y_i, F_j(x_i))_{j=1}^{K}}{\partial F(x_i)}]_{(F_j(X)=F_{j,m-1}(X))_1^K} \\ &= y_i^k - p_k(x_i) \end{aligned} \tag{2}$$

where $y_i^k - p_k(x_i)$ is the current residuals, and thus, K-trees are induced and. Meanwhile, this will produce K trees each with L-terminal nodes at iteration m, R_{klm}.

Then, each terminal node of each tree can be solved by a separate line search, as shown in Eq. (3).

$$\gamma_{lm} = \arg\min_{\gamma} \sum_{x_i \in R_{lm}} \psi(y_i, F_{m-1}(x_i) + \gamma) \tag{3}$$

Each of the functions is updated and then the SGB is established. The detailed description of SGB algorithm can be found in [27], [29], and [42]. As mentioned in these studies, there are three key parameters for the SGB algorithm: the

number of trees (boosting interactions, M), the interaction depth (the max tree depth, J) and the shrinkage (the learning rate, v). These hyper-parameters have to be tuned to improve the generalization and performance of an SGB model.

B. MODEL PERFORMANCE EVALUATION

One of the basic tools for assessing the confidence of the classifiers is the confusion matrix. The confusion matrix is an $m \times m$ matrix, as shown in Eq. (4). For each model, a confusion matrix is presented [43]. To analyze the predictive performance of the three models, two global classification metrics of the classification accuracy and Cohen's Kappa index, and three within-class classification metrics of recall, precision, and F-measure are utilized to analyze the predictive performance and generalization ability of the algorithms. Note that the symbols in this section are different to that in section A.

$$\text{matrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mm} \end{bmatrix} \tag{4}$$

where x_{ii} on the main diagonal represents the number of samples belonging to class i that are predicted, x_{ij} represents the number of samples belonging to class i that are predicted to class j, *and* m is the number of classes.

The discriminant ability of the models can be evaluated according to the discriminant accuracy rate, and the accuracy can be calculated as Eq. (5). Cohen's Kappa is the index used to assess inter-rater reliability when coding categorical variables. The statistic is also considered to be an improvement over using percentage to evaluate the reliability [44]. The Kappa can be given by Eq. (6).

$$\text{Accuracy} = \left(\frac{1}{n}\sum_{i=1}^{m} x_{ii}\right) \times 100\% \tag{5}$$

$$\text{Kappa} = \frac{n\sum_{i=1}^{m} x_{ii} - \sum^{m}(x_{i+} \cdot x_{+i})}{n^2 - \sum_{i=1}^{m}(x_{i+} \cdot x_{+i})} \tag{6}$$

where n is the number of total samples in the dataset, x_{i+} is the number of samples belonging to class i, and x_{+i} is the number of samples that are predicted to class j.

The range of the Kappa value is from -1 to 1 and can be divided into six groups to represent different levels of consistency (as shown in Table 2). In general, if the value of Kappa is less than 0.4, it represents that the strength of agreement is poor, and if the value of Kappa is greater than or equal to 0.4, then the strength of agreement is good [33], [45].

For pillar stability problems, not only the predictive accuracy of classifiers for all samples, but also the predictive performance on samples subject to a certain class should be considered. Three within-class classification metrics, recall, precision and the F-measure, are widely used to evaluate the discriminant power of models on cases with a certain grade. Precision is defined as the ratio of the total number of samples with a certain grade classified correctly to the total number of cases of that grade assessed by a model. Recall is the ratio of the total number of cases with a certain grade classified correctly to the total number of cases of that grade [46]. The two metrics are often selected to provide a single measure called the F-measure. The three metrics are given as follows:

TABLE 2. The basic scale of agreement with the Kappa value.

Strength of agreement	Kappa value
total disagreement	[-1.0, 0)
Slight	[0, 0.2)
Poor	[0.2, 0.4)
Moderate	[0.4, 0.6)
Substantial	[0.6, 0.8)
Perfect	[0.8, 1.0]

$$\text{Recall}_i = \left(\frac{x_{ii}}{\sum_{j=1}^{m} x_{+j}}\right) \times 100\% \tag{7}$$

$$\text{Precision}_i = \left(\frac{x_{ii}}{\sum_{j=1}^{m} x_{j+}}\right) \times 100\% \tag{8}$$

$$\text{F-measure} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{9}$$

A receiver operating characteristic (ROC) curve is a graphical representation of true positives out of the positives and false positives out of the negatives [33], [47]. The ROC curve can be used to compare the performance of different algorithms. Generally, a classifier has superior discriminant performance when its ROC curve is in the upper-left corner. The performance of different classifiers can also be analyzed by the area under the ROC curve (AUC). The larger the AUC is, the better the performance of the classifier is. In this study, the method is also applied to evaluate and compare the discriminant performance of the four models for predicting pillar stability.

C. MODEL DEVELOPMENT AND PARAMETER OPTIMIZATION

To train and validate the performance of classifiers, the original dataset should be divided into two subsets: a training set and a testing set. The selection of training data is important for the training of the classifiers and the training set must be representative of the whole dataset. For a small training set, the relationship cannot be properly learned. For a too large training set, the generalization capability cannot be verified and over-fitting may occur [22], [48]. Thus, the training and testing set is often determined by an optimization analysis. In this study, 165 cases (80% of the original data) are used as the training set, and the remaining 40 cases are taken as

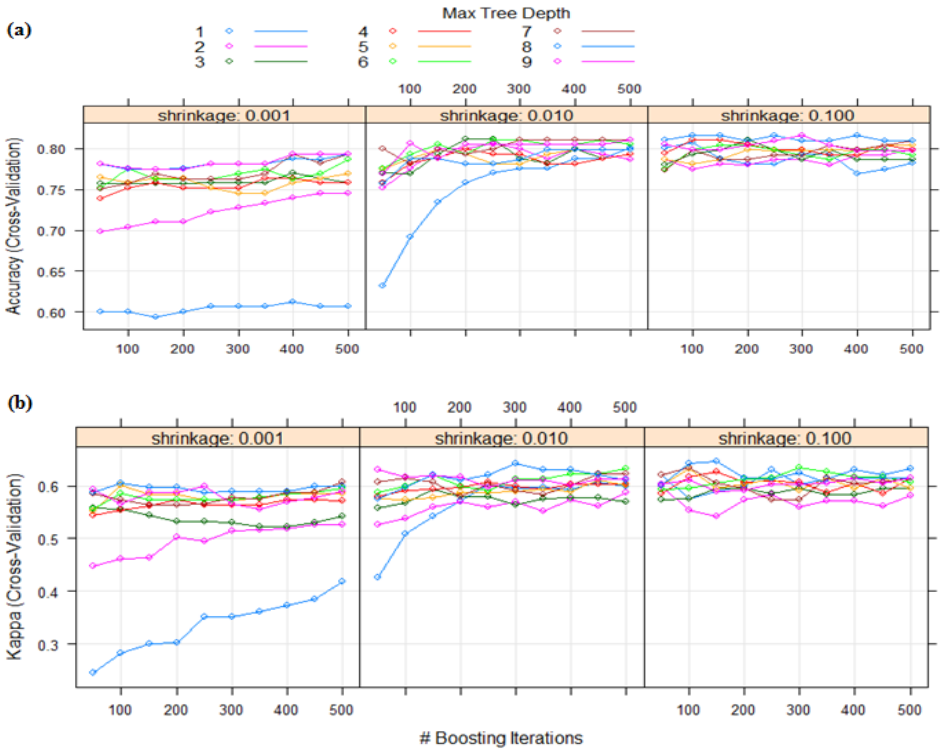

FIGURE 4. Tuning parameters for determining the optimal model of SGB. (a) Accuracy; (b) Kappa.

the testing set after optimization analysis and kennard-stone sampling.

To optimize the key parameters of the SGB model and to obtain better discriminant performance, the 10-fold CV method, which is the number of folds recommended by Rodriguez *et al.* [49], is used as the validation method. During the process of 10-fold CV, the training set is randomly divided into 10 folds: the training sub-set (9 folds) and the validation set (the remaining one fold), as can be seen in Fig. 5. The training sub-set is used for machine learning, and then classifiers are established by matching some hyper-parameters. The validation set is used to adjust the parameters and validate the generalization capability of the classifiers. The process is repeated 10 times on different training sub-set, and at the end, every sample has been used for testing exactly once and for training 9 times. The accuracy of 10-fold CV is calculated by simply averaging the 10 individual accuracy, and the whole 10-fold CV is also repeated 10 times to obtain the reliable results. Then the classifiers with the optimal hyper-parameters are trained using the whole training set, and the testing set (which has never been used to develop the classifiers) is used as an independent testing set to test the predictive ability of the optimal models.

It is generally considered that a model with better generalization capability can be obtained by 10-fold CV method. But the generalization capability of models are obtained based on the training set and considered to be able to generalize to other unknown samples. The testing set is used as unknown samples to test the generalization capability of models.

As described in section 3.1, the SGB algorithm has three key parameters: the number of trees (boosting interaction, M), the interaction depth (the max tree depth, J) and the shrinkage (the learning rate, v). To obtain the best performance of the SGB model, the parameters are examined using 10-fold CV method. According to many literatures, tuning parameters $v = (0.001, 0.01, 0.1)$, $M = (50, 100, 150, \ldots, 500)$, $J = (1, 2, 3, \ldots, 9)$ with a 10-fold CV process in this paper. The accuracy and Cohen's Kappa are used to determine the best combination of the SGB model, and the results are shown in Fig. 4. The final model can be selected using the largest value of accuracy and Kappa. As can be seen, the range of accuracy of SGB models is from 60% to 81%, and the value of Kappa falls into the range [0.35-0.65]. It also can be found that the best choice of the three parameters is $M = 150$, $J = 8$, $v = 0.1$, and the accuracy and Kappa of the SGB model are 81.21% and 0.6420, respectively.

To verify the feasibility and reliability of the SGB model, the performance of the SGB algorithm was compared with the performance of SVM, RF and MLPNN models. These algorithms were chosen because they are increasingly used in pillar stability analysis, and all of them have higher discriminant performance. SVM, RF and MLPNN models are also optimized by the 10-fold CV method. Details of the key parameters choice in this study are as follows:

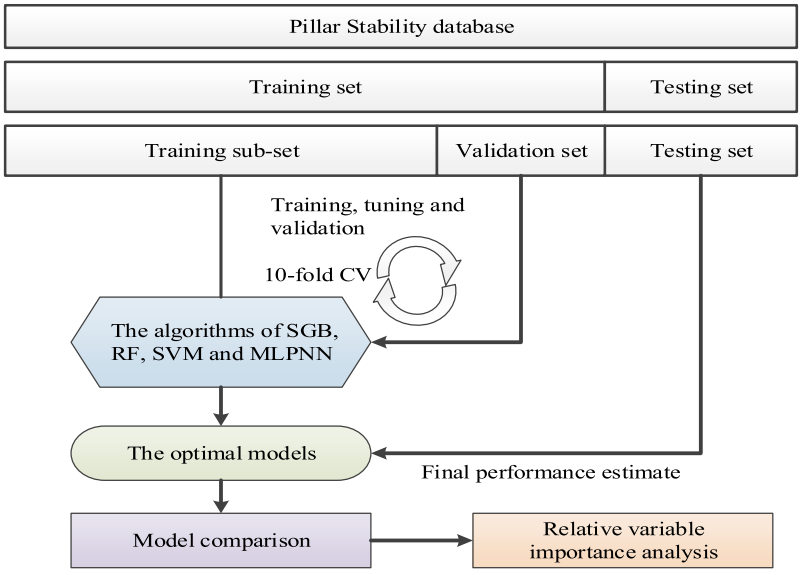

FIGURE 5. Overall procedure for pillar stability prediction using supervised learning methods.

(1) RF model: The algorithm has two parameters (one is the number of classification trees n_{tree}, the other is the number of variables m_{try}) that need to be optimized. m_{try} is more sensitive for discriminant accuracy, and the discriminant accuracy is affected little by n_{tree}. n_{tree} and m_{try} are tested for the number of input factors of the pillar cases. In this paper, the tuning range for the number of trees is 100–1000 with 50 interval. Moreover, there are six indexes in this study, so the values of m_{try} can be (1, 2, 3, 4, 5, 6). The best choice of n_{tree} and m_{try} is 500 and 3, respectively.

(2) SVM model: According to the research of Zhou *et al.* [12], Dong *et al.* [50], and Samui [51], the key parameters of the SVM algorithm are the constraints violation C and sigma. The sigma parameter can be approximately estimated by using the R software ''sigest'' function. Based on the 10-fold CV method, C is tuned for 10 values (2^{-1}, 2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7 and 2^8) to find the optimal parameter in this study. Finally, C $= 2^0$ and sigma $= 0.3525$ are identified as the best values.

(3) MLPNN model: The model tunes the number of hidden neurons H (1, 2, 3, 4, 5, 6, 7, 8, 9 and 10) and tries 10 random-weight initializations. The results show that the optimal parameter values are as follows: hidden units $H = 5$, decay $= 0.1$.

Thus, the over procedure for pillar stability prediction using the SGB model in this study can be expressed as shown in Fig. 5.

IV. RESULTS AND DISCUSSION

A. DISCRIMINANT RESULTS AND PERFORMANCE ANALYSIS

The prediction results of the four optimum models are given in Table 3. The values on the diagonal are the numbers of samples that are correctly predicted. The number of samples that are correctly predicted by the SGB model (with a total number of 36) is the largest. Although the number is lower than those of the SGB model, the RF (34 samples are correctly predicted), SVM (33 samples are correctly predicted) and MLPNN (32 samples are correctly predicted) can also achieve satisfactory results. Therefore, all of the four models have high discriminant power for pillar stability, and the performance of the SGB model is the best.

The value of accuracy and Kappa for each model are shown in Table 4. For the training set, accuracy and Kappa are calculated by 10 times 10-fold CV, and the accuracy of the models falls into the range [71.52–93.33%], as shown in Table 4. Obviously, the RF model exhibited the highest average accuracy rate (93.33%), followed by the SVM model with an average accuracy rate of 83.03%, then the SGB and MLPNN models with average accuracy rates of 81.21% and 71.52%, respectively. Meanwhile, the Kappa values of the four models fall into the range [0.5083–0.8930], and the agreement strength of Kappa in the SGB, RF, SVM, and MLPNN models is from moderate to perfect according to the basic scale shown in Table 2. It is obvious that the Kappa of the RF model is the highest with a value of 0.8930, and then the SVM, SGB and MLPNN models rank successively. The agreement strength of the SGB model is substantial. For the testing set, the accuracy of the models is from 80.00 to 90.00%, and the SGB model has the highest accuracy rate (90.00%), followed by the RF, SVM and MLPNN models with accuracy rates of 85.00%, 82.50% and 80.00%, respectively. Meanwhile, the Kappa values fall into the range [0.6548–0.8351]. The agreement strength of Kappa is from substantial to perfect. It is obvious that the Kappa of the SGB model is the highest with a value of 0.8351, and then the RF, SVM and MLPNN models rank successively. Therefore, the SGB model has superior generalization ability over the

TABLE 3. Prediction results of the testing set across the three models.

Predicted	Observed											
	SGB			RF			SVM			MLPNN		
	F	S	U	F	S	U	F	S	U	F	S	U
F	**19**	2	0	**17**	1	1	**18**	2	3	**18**	1	5
S	0	**12**	2	0	**13**	2	1	**12**	1	1	**13**	1
U	0	0	**5**	2	0	**4**	0	0	**3**	0	0	**1**

The diagonal elements are correct decisions.

TABLE 4. Comparison of accuracy and Kappa for the four models.

Date set	Metrics	SGB	RF	SVM	MLPNN
Training	Accuracy / %	81.21	93.33	83.03	71.52
set	Kappa	0.6420	0.8930	0.7202	0.5083
Testing	Accuracy / %	90.00	85.00	82.50	80.00
set	Kappa	0.8351	0.7586	0.7040	0.6548

testing samples. As known, the performance of classifiers mainly depend on the performance on yet-unseen data. Thus, it can be concluded that the comprehensive prediction performance of the SGB model is the best, and the model is feasible and applicable for pillar stability prediction. Meanwhile, the results show that all four models have a relatively balanced generalization ability for the training set and testing set, which indicates that it is necessary to optimize the models using the 10-fold CV method [22], [47].

The results obtained for three within-class classification metrics based on the testing set are given in Table 5. Notably, the recall, precision and F-measure of the four classifiers for pillar stability prediction based on the testing set exhibit large deviations (recall = 14.29–100%, precision = 75–100%, and F-measure = 25.01–95%). For samples subject to class F and U, the best discriminant performance is exhibited by the SGB model, with recall of 100 and 71.43%, precision of 90.48 and 100%, and an F-measure of 95 and 83.33%, respectively. For samples subject to class S, the performance of the four classifiers is similar. As known, it is important to predict of failure and unstable cases correctly for pillar stability prediction in underground mines. Thus, it can be concluded that the SGB model is more suitable for the prediction of pillar stability in underground mines. Moreover, the RF and SVM models also achieve satisfactory results in all cases. However, the generalization ability of MLPNN model is poor for cases subject to class U.

For the testing set, the ROC curves of the four models are obtained and shown in Fig. 6. From Fig, 6, it can be seen that the shape of the ROC curves of SGB, RF, SVM and MLPNN models are from right to left, successively, denoting that the SGB model achieves higher overall performance. The AUC values of the four classifiers are in the range of 0.791-0.871. According to the classification standard of AUC value in [33],

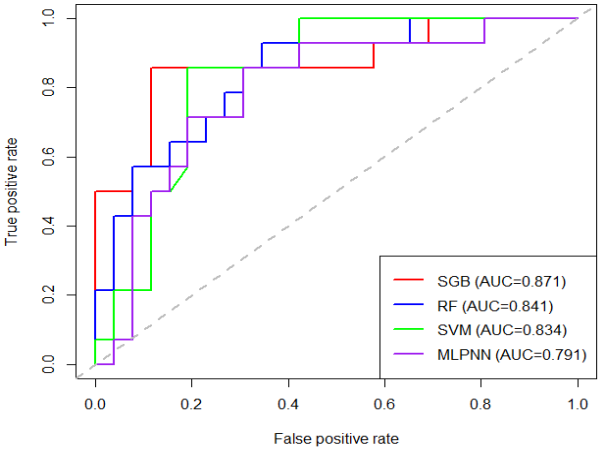

FIGURE 6. ROC curves and AUC values of the four models on the testing set.

the generalization of the four classifiers are from moderate to good. It is obvious that the highest AUC is obtained by the SGB model with an AUC equal to 0.891, followed by RF, SVM and MLPNN models. Both of these indicate that the performance of SGB model is the best and then RF, SVM and MLPNN models, respectively.

Based on the above analysis, it can be concluded that the SGB model with superior generalization is feasible and reliable for the prediction of pillar stability in underground mines.

B. RELATIVE VARIABLE IMPORTANCE

It is meaningful to analyze the sensitivity of parameters when taking measures to prevent pillar failure in underground mines. However, it is difficult to determine the sensitivity because all of the parameters are sensitive to pillar stability and the correlation effect of input parameters. Moreover, there are many methods to obtain the sensitivity of parameters [33], [47], [52]–[56]. In this study, the sensitivity of parameters was investigated using the method of relative importance score based on the optimum SGB and RF models. The selection for these two models was based on their outstanding performance on the testing set.

The generic function *varImp* () in the caret package of *R* software can be used to calculate the importance of parameters. The final variable importance was calculated

TABLE 5. The values of three within-class classification metrics for each model.

Class	MLPNN			SVM			RF			SGB		
	Recall	Precision	F	Recall	Precision	F	Recall	Precision	F	Recall	Precision	F
	%	%	%	%	%	%	%	%	%	%	%	%
F	94.74	75.00	83.72	94.74	78.26	85.72	89.47	89.47	89.47	100.00	90.48	95.00
S	92.86	86.67	89.66	85.71	85.71	85.71	92.86	86.67	89.66	85.71	85.71	85.71
U	14.29	100.00	25.01	42.86	100.00	60.00	57.14	66.67	61.54	71.43	100.00	83.33

Note: F represents the F-Measure.

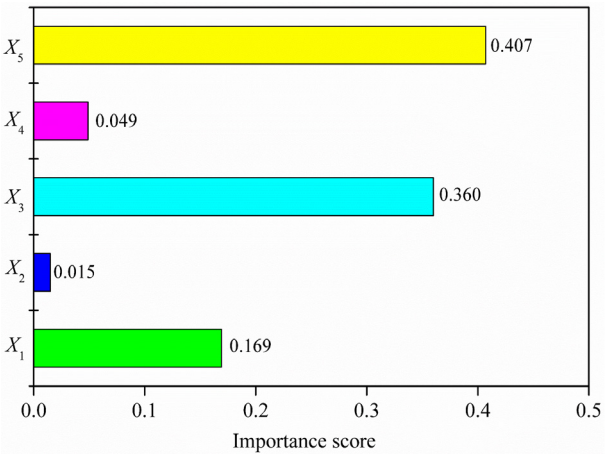

FIGURE 7. Ranking variable importance associated with pillar stability discrimination by SGB and RF.

by averaging the relative variable importance from the optimum RF and SGB models. Variable importance scores are normalized (the sum of all importance scores is one), and the result is shown in Fig. 7.

It is obvious that pillar stress is the most sensitive factor for pillar stability with the importance score of 0.407, indicating that the factor has a significant influence on pillar stability. The reason may be that the stability of a pillar is mainly affected by the mechanical response of the pillar to the load caused by mining, and the mechanical response is directly related to the stress state of the pillar. It has also been found that high pillar stress caused by high extraction ratios can cause spalling of the pillar and spalling can initiate when the average pillar stress exceeds about 10% of the uniaxial compressive strength of the rock [35], [57]. The importance of pillar stress is in accordance with that obtained by Zhou *et al.* [20].

Second in sensitivity is the ratio of pillar width to pillar height with an importance score of 0.360, illustrating that it is also a major influence factor on pillar stability. The ratio is the main reference index for the design and strength estimation of the pillars in underground mines. It was found that the strength of a pillar decreased by nearly 60% when the ratio of width to height changed from 1.0 to 0.5 [35]. Therefore, the ratio of pillar width to pillar height should be considered to predict the stability of pillars in underground mines.

The importance scores of pillar width is 0.169, suggesting that the variable is also an important influence factor of pillar stability. It can also be found that the uniaxial compressive strength of rock (0.049) and pillar height (0.015) are not as sensitive as the former three factors. This conclusion has also been studied in [20].

As describe above, the pillar stress and the ratio of pillar width to pillar height are the major influence variables of pillar stability. Thus, the stability of pillars in underground mines can be improved by improving the stress state of pillars and optimizing the ratio of pillar width to pillar height.

However, it should be noted that the relative importance of variables depends on the dataset and models, and more representative results can be obtained as more valid pillar cases are available.

V. CONCLUSIONS

The SGB algorithm, along with RF, SVM and MLPNN models were introduced in this study for the prediction of pillar stability in underground mines. Five potential relevant indicators, pillar width, pillar height, the ratio of pillar width to pillar height, and the uniaxial compressive strength of rock and pillar stress, were chosen as the prediction indicators. 10-fold CV method is applied to improve the generalization ability of classification models. Based on the analysis above, the following conclusions can be drawn:

(1) The shape of the ROC curve of the SGB model was generally closer to the left and top axes than that of the other models, and the accuracy, Kappa and AUC of the SGB model on the testing sets were 90%, 0.8351 and 0.891, respectively, indicating that the SGB can be considered to have outstanding performance for pillar stability prediction. The order of the performance of the four models is SGB, RF, SVM and MLPNN successively.

(2) The SGB model with higher values of recall, precision and F-measure can be a valuable tool for prediction of all pillar cases in underground mines, particularly for the cases subject to failure and unstable classes, followed by RF, SVM and MLPNN models.

(3) Pillar stress and the ratio of pillar width to pillar height were found to be the two most influential variables for the prediction of pillar stability, which achieved 0.407 and 0.360 importance scores, respectively.

In future research, more samples with other important parameters can be introduced to develop the SGB model to improve its feasibility and reliability, such as microseismic monitoring signals and the characteristics of dynamic disturbances.

REFERENCES

[1] M. Najafi, S. E. Jalali, A. Y. Bafghi, and F. Sereshki, "Prediction of the confidence interval for stability analysis of chain pillars in coal mines," *Saf. Sci.*, vol. 49, no. 5, pp. 651–657, 2011.

[2] M. D. G. Salamon, "Stability, instability and design of pillar workings," *Int. J. Rock Mech. Mining Sci. Geomech. Abstr.*, vol. 7, no. 6, pp. 613–631, 1970.

[3] D. R. Tesarik, J. B. Seymour, and T. R. Yanske, "Long-term stability of a backfilled room-and-pillar test section at the Buick Mine, Missouri, USA," *Int. J. Rock Mech. Mining Sci.*, vol. 46, no. 7, pp. 1182–1196, 2009.

[4] B. H. G. Brady and E. T. Brown, *Rock Mechanics: For Underground Mining*. Springer, 2013. [Online]. Available: https://link.springer.com/book/10.1007/978-1-4020-2116-9

[5] P. J. Lunder, "Hard rock pillar strength estimation an applied empirical approach," Ph.D. dissertation, Appl. Sci., Fac. Mining Eng., Univ. British Columbia, Vancouver, BC, Canada, 1994.

[6] J. Deng, Z. Q. Yue, L. G. Tham, and H. H. Zhu, "Pillar design by combining finite element methods, neural networks and reliability: A case study of the Feng Huangshan copper mine, China," *Int. J. Rock Mech. Mining Sci.*, vol. 40, no. 4, pp. 585–599, 2003.

[7] A. Mortazavi, F. P. Hassani, and M. Shabani, "A numerical investigation of rock pillar failure mechanism in underground openings," *Comput. Geotechn.*, vol. 36, no. 5, pp. 691–697, 2009.

[8] G. York, "Numerical modelling of the yielding of a stabilizing pillar/foundation system and a new design consideration for stabilizing pillar foundations," *J. Southern Afr. Inst. Mining Metall.*, vol. 98, no. 6, pp. 281–297, 1998.

[9] W. A. Hustrulid, "A review of coal pillar strength formulas," *Rock Mech.*, vol. 8, no. 2, pp. 115–145, 1976.

[10] M. Jawed, R. K. Sinha, and S. Sengupta, "Chronological development in coal pillar design for bord and pillar workings: A critical appraisal," *J. Geol. Mining Res.*, vol. 5, no. 1, pp. 1–11, 2013.

[11] C. D. Martin and W. G. Maybee, "The strength of hard-rock pillars," *Int. J. Rock Mech. Mining Sci.*, vol. 37, no. 8, pp. 1239–1246, 2000.

[12] J. Zhou, X.-B. Li, X. Shi, W. Wei, and B. Wu, "Predicting pillar stability for underground mine using Fisher discriminant analysis and SVM methods," *Trans. Nonferrous Metals Soc. China*, vol. 21, no. 12, pp. 2734–2743, 2011.

[13] E. Ghasemi, H. Kalhori, and R. Bagherpour, "Stability assessment of hard rock pillars using two intelligent classification techniques: A comparative study," *Tunnelling Underground Space Technol.*, vol. 68, pp. 32–37, Sep. 2017.

[14] R. K. Wattimena, "Predicting the stability of hard rock pillars using multinomial logistic regression," *Int. J. Rock Mech. Mining Sci.*, vol. 71, pp. 33–40, 2014, doi: 10.1016/j.ijrmms.2014.03.015.

[15] S. Shnorhokian, H. S. Mitri, and L. Moreau-Verlaan, "Stability assessment of stope sequence scenarios in a diminishing ore pillar," *Int. J. Rock Mech. Mining Sci.*, vol. 74, pp. 103–118, Feb. 2015.

[16] M. Cauvin, T. Verdel, and R. Salmon, "Modeling uncertainties in mining pillar stability analysis," *Risk Anal.*, vol. 29, no. 10, pp. 1371–1380, 2009.

[17] A. S. Tawadrous and P. D. Katsabanis, "Prediction of surface crown pillar stability using artificial neural networks," *Int. J. Numer. Anal. Methods Geomech.*, vol. 31, no. 7, pp. 917–931, 2007.

[18] R. K. Wattimena, S. Kramadibrata, I. D. Sidi, and M. A. Azizi, "Developing coal pillar stability chart using logistic regression," *Int. J. Rock Mech. Mining Sci.*, vol. 58, pp. 55–60, Feb. 2013.

[19] D. V. Griffiths, G. A. Fenton, and C. B. Lemons, "Probabilistic analysis of underground pillar stability," *Int. J. Numer. Anal. Methods Geomech.*, vol. 26, no. 8, pp. 775–791, 2002.

[20] J. Zhou, X. Li, and H. S. Mitri, "Comparative performance of six supervised learning methods for the development of models of hard rock pillar stability prediction," *Natural Hazards*, vol. 79, no. 1, pp. 291–316, 2015.

[21] C. Qi, A. Fourie, G. Ma, X. Tang, and X. Du, "Comparative study of hybrid artificial intelligence approaches for predicting hangingwall stability," *J. Comput. Civil Eng.*, vol. 32, no. 2, p. 04017086, 2017.

[22] C. Qi and X. Tang, "Slope stability prediction using integrated metaheuristic and machine learning approaches: A comparative study," *Comput. Ind. Eng.*, vol. 118, pp. 112–122, Apr. 2018.

[23] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, 2001.

[24] J. H. Friedman, "Stochastic gradient boosting," *Comput. Statist. Data Anal.*, vol. 38, no. 4, pp. 367–378, 2002.

[25] R. Blagus and L. Lusa, "Boosting for high-dimensional two-class prediction," *BMC Bioinf.*, vol. 16, no. 1, p. 300, 2015.

[26] R. Lawrence, A. Bunn, S. Powell, and M. Zambon, "Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis," *Remote Sens. Environ.*, vol. 90, no. 3, pp. 331–336, 2004.

[27] J. Zhou, X.-Z. Shi, R.-D. Huang, X.-Y. Qiu, and C. Chen, "Feasibility of stochastic gradient boosting approach for predicting rockburst damage in burst-prone mines," *Trans. Nonferrous Metals Soc. China*, vol. 26, no. 7, pp. 1938–1945, 2016.

[28] L. Lombardo, M. Cama, C. Conoscenti, M. Märker, and E. Rotigliano, "Binary logistic regression versus stochastic gradient boosted decision trees in assessing landslide susceptibility for multiple-occurring landslide events: Application to the 2009 storm event in Messina (Sicily, southern Italy)," *Natural Hazards*, vol. 79, no. 3, pp. 1621–1648, 2015.

[29] J. Zhou, X. Shi, and X. Li, "Utilizing gradient boosted machine for the prediction of damage to residential structures owing to blasting vibrations of open pit mining," *J. Vib. Control*, vol. 22, no. 19, pp. 3986–3997, 2016.

[30] Y. Gao, J. Pan, G. Ji, and F. Gao, "A time-series modeling method based on the boosting gradient-descent theory," *Sci. China Technol. Sci.*, vol. 54, no. 5, pp. 1325–1337, 2011.

[31] G. G. Moisen, E. A. Freeman, J. A. Blackard, T. S. Frescino, N. E. Zimmermann, and T. C. Edwards, "Predicting tree species presence and basal area in Utah: A comparison of stochastic gradient boosting, generalized additive models, and tree-based methods," *Ecol. Model.*, vol. 199, no. 2, pp. 176–187, Nov. 2006.

[32] G. Zhao and J. Liu, "Analysis of the pillar stability based on the Gaussian process for machine learning," *J. Saf. Environ.*, vol. 17, no. 5, pp. 1725–1729, 2017.

[33] Y. Lin, K. Zhou, and J. Li, "Prediction of slope stability using four supervised learning methods," *IEEE Access*, vol. 6, pp. 31169–31179, 2018.

[34] G. Rashed and S. S. Peng, "Change of the mode of failure by interface friction and width-to-height ratio of coal specimens," *J. Rock Mech. Geotech. Eng.*, vol. 7, no. 3, pp. 256–265, 2015.

[35] G. S. Esterhuizen, D. R. Dolinar, and J. L. Ellenberger, "Pillar strength in underground stone mines in the United States," *Int. J. Rock Mech. Mining Sci.*, vol. 48, no. 1, pp. 42–50, 2011.

[36] W.-D. Song, S. Cao, J.-X. Fu, G.-J. Jiang, and F. Wu, "Sensitivity analysis of impact factors of pillar stability and its application," *Rock Soil Mech.*, vol. 35, pp. 271–277, 2014. [Online]. Available: http://ytlx.whrsm.ac.cn/EN/abstract/abstract15198.shtml

[37] E. Ghasemi, M. Ataei, and K. Shahriar, "Prediction of global stability in room and pillar coal mines," *Natural Hazards*, vol. 72, no. 2, pp. 405–422, 2014.

[38] J. Sjoberg, "Failure modes and pillar behaviour in the Zinkgruvan mine," in *Proc. 33rd U.S. Symp. Rock Mech. (USRMS)*. New York, NY, USA: American Rock Mechanics Association, 1992, pp. 491–500.

[39] Y. Potvin, M. Hudyma, and H. Miller, "Rib pillar design in open stope mining," *Bull. Can. Inst. Mining Metall.*, vol. 82, no. 927, pp. 31–36, 1989.

[40] M. R. Von Kimmelmann, B. Hyde, and R. J. Madgwick, "7 The use of computer applications at BCL Limited in planning pillar extraction and the design of mining layouts," in *Proc. Des. Perform. Underground Excavations, ISRM Symp.*, Cambridge, U.K., Sep. 1984, pp. 53–63.

[41] D. G. F. Hedley and F. Grant, "Stope-and-pillar design for the Elliot Lake uranium mines," *Can. Mining Metall. Bull.*, vol. 65, no. 723, p. 37, 1972.

[42] A. Jasra and C. C. Holmes, "Stochastic boosting algorithms," *Statist. Comput.*, vol. 21, no. 3, pp. 335–347, 2010.

[43] S. Visa, B. Ramsay, A. L. Ralescu, and E. Van Der Knaap, "Confusion matrix-based feature selection," in *Proc. MAICS*, 2011, pp. 120–127.

[44] A. Yee, R. T. Corlett, S. Liew, and H. T. Tan, "The vegetation of Singapore—An updated map," *Gardens' Bull. Singapore*, vol. 63, nos. 1–2, pp. 205–212, 2011.

[45] M. L. McHugh, "Interrater reliability: The kappa statistic," *Biochemia Medica*, vol. 22, no. 3, pp. 276–282, 2012.

[46] J. Zhou, X. Li, and H. S. Mitri, "Classification of rockburst in underground projects: Comparison of ten supervised learning methods," *J. Comput. Civil Eng.*, vol. 30, no. 5, p. 04016003, 2016.
[47] C. Qi, A. Fourie, X. Du, and X. Tang, "Prediction of open stope hangingwall stability using random forests," *Natural Hazards*, vol. 92, no. 2, pp. 1179–1197, 2018.
[48] C. Qi, A. Fourie, Q. Chen, and Q. Zhang, "A strength prediction model using artificial intelligence for recycling waste tailings as cemented paste backfill," *J. Cleaner Prod.*, vol. 183, pp. 566–578, May 2018.
[49] J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 569–575, Mar. 2010.
[50] L. Dong, J. Wesseloo, Y. Potvin, and X. Li, "Discrimination of mine seismic events and blasts using the Fisher classifier, Naive Bayesian classifier and logistic regression," *Rock Mech. Rock Eng.*, vol. 49, no. 1, pp. 183–211, 2015.
[51] P. Samui, "Slope stability analysis: A support vector machine approach," *Environ. Geol.*, vol. 56, no. 2, p. 255, 2008.
[52] Y. Qin, Z. Wang, C. Xiang, M. Dong, C. Hu, and R. Wang, "A novel global sensitivity analysis on the observation accuracy of the coupled vehicle model," *Vehicle Syst. Dyn.*, pp. 1–22, Sep. 2018, doi: 10.1080/00423114.2018.1517219.
[53] X. Tang, W. Yang, X. Hu, and D. Zhang, "A novel simplified model for torsional vibration analysis of a series-parallel hybrid electric vehicle," *Mech. Syst. Signal Process.*, vol. 85, pp. 329–338, Feb. 2017, doi: 10.1016/j.ymssp.2016.08.020.
[54] C. Hu, R. Wang, F. Yan, Y. Huang, H. Wang, and C. Wei, "Differential steering based yaw stabilization using ISMC for independently actuated electric vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 2, pp. 627–638, Feb. 2018.
[55] C. Qi, Q. Chen, A. Fourie, and Q. Zhang, "An intelligent modelling framework for mechanical properties of cemented paste backfill," *Minerals Eng.*, vol. 123, pp. 16–27, Jul. 2018.
[56] C. Qi, A. Fourie, G. Ma, and X. Tang, "A hybrid method for improved stability prediction in construction projects: A case study of stope hangingwall stability," *Appl. Soft Comput.*, vol. 71, pp. 649–658, Oct. 2018.
[57] C. J. Pritchard and D. G. F. Hedley, "Progressive pillar failure and rockbursting at denison mine," in *Proc. 3rd Int. Symp. Rockbursts Seismicity Mines, Kingston. Rotterdam, AA Balkema*, 1993, p. 111J6.

HANGXING DING received the Ph.D. degree in mining engineering from Northeastern University, Shenyang, China, in 2013. He has authored over 10 papers published in related journals. He holds three patents. His research interest is mining method for complex orebody. He received the Provincial/Ministerial Science and Technology Award for six projects.

GUANGHUI LI received the M.S. degree in mining engineering from the University of Science and Technology, Anshan, China, in 2017. He is currently pursuing the Ph.D. degree with Northeastern University, majoring in mining engineering. He holds one patent. His research interest is in efficient mining techniques.

XIN DONG received the B.S. degree in mining engineering from Northeastern University, Shenyang, China, in 2018. In 2018, he was recommended to Northeastern University for doctoral studies. He is currently with the Rock Fracture and Instability Research Center, Northeastern University. His research interest is the study of slope engineering stability.

YUN LIN (M'17) received the B.S. degree in geography information system from the China University of Geosciences, Wuhan, China, in 2012, and the M.S. degree in mining engineering from Central South University, Changsha, China, in 2015. From 2015 to 2018, he studied with the State Key Laboratory of Optical Information, Central South University. He has authored over 10 papers published in related international conference proceedings and journals, and holds two patents. His research interests included risk evaluation with machine learning algorithms. He served as a reviewer for some journals.