ORIGINAL PAPER

Rockburst Intensity Grade Prediction Based on Data Preprocessing Techniques and Multi-model Ensemble Learning Algorithms

Zhi-Chao Jia[1] · Yi Wang[1] · Jun-Hui Wang[1] · Qiu-Yan Pei[1] · Yan-Qi Zhang[1]

Abstract

Rockburst is a mine dynamic disaster caused by the rapid release of elastic strain energy of surrounding rock. As the depth of engineering project operations increases, accurate classification of rockburst intensity cannot be achieved based on conventional criteria due to high uncertainty and unpredictability of rockburst. In this regard, an AOA-Voting-Soft ensemble machine learning was proposed in this study by combining seven individual classifiers, i.e., eXtreme gradient boosting, support vector machines, multilayer perceptron, k-nearest neighbor, random forest, naive Bayesian, and gradient boosting decision Tree. In addition, outliers were eliminated by means of density-based spatial clustering of applications with noise, and CURE-MeanradiusSMOTE was adopted to obtain a balanced data structure. Furthermore, the optimal combination of classifiers in Voting was determined by the game theory and the exhaustive search method. Weights of individual learners in Voting were determined through the arithmetic optimization algorithm and fivefold cross-validation. The results show that the prediction accuracy of the ensemble algorithm proposed in this study is 4.4% higher than that of the individual classifier with optimal performance. The importance analysis indicates that the elastic energy index is the most important variable that affects rockburst intensity grades. Moreover, this rockburst ensemble method can be applied further to solve other classification problems in underground engineering projects.

Highlights

1. This study improves the data preprocessing method, outliers were eliminated by means of density-based spatial clustering of applications with noise, and CURE-MeanradiusSMOTE was proposed to obtain a balanced data structure.
2. This study presents a hybrid ensemble model for Rockburst intensity grade prediction, combining a new metaheuristic method with the Voting-Soft model.
3. This study combines game theory and method of exhaustion to determine the best classifier combination in voting.
4. The weights of individual learners in Voting were determined through arithmetic optimization algorithm and fivefold cross-validation.
5. Sensitivity study was conducted on input variables with RBD-FAST, and the results suggest that W_{et} is the most important input variable.

✉ Yi Wang
wangyi@tyut.edu.cn

Zhi-Chao Jia
jiazhichao1560@link.tyut.edu.cn

Jun-Hui Wang
wangjunhui@tyut.edu.cn

Qiu-Yan Pei
peiqiuyan@tyut.edu.cn

Yan-Qi Zhang
zhangyanqi1612@link.tyut.edu.cn

1 College of Safety and Emergency Management Engineering, Taiyuan University of Technology, Taiyuan 030024, China

1 Introduction

Rockburst is a sudden geological disaster induced by the rapid release of accumulated elastic strain energy, and it leads to brittle failure such as rock throwing, ejection, and spalling in mines, tunnels and other geotechnical engineering (Askaripour et al. 2022; Zhao and Chen 2020). The U.S. witnessed a total of 172 rockburst cases in the period 1936–1993 (Mark 2016), and countries such as Canada, India, Sweden, and China have also undergone rockburst for many times (Leveille et al. 2017; Sepehri et al. 2020; Simser 2019; Sun et al. 2021), yet this is not an exhaustive list. As the depth of mining and the locations of excavation activities pose greater challenges (Xu et al. 2022), more cases of rockburst have occurred. For the purpose of predicting intensity grades of rockburst accurately and finding targeted preventive measures, scholars worldwide proposed a large number of methods, including empirical criteria(Gong et al. 2020, 2019; Xue et al. 2020a), numerical simulation (Lu et al. 2021; Yang et al. 2022; Yu et al. 2023), and on-site monitoring (Hu et al. 2023a; Xue et al. 2021; Zhang et al. 2018). The empirical criteria method is an empirical method in which the parameters can be obtained through experiments, but it fails to be transferred effectively between different engineering scenarios. The results generated by the numerical simulation method always deviate from actual

on-site situations, being far from satisfactory. Although relevant parameters can be obtained in real time by means of on-site monitoring, it is still difficult to determine the rockburst threshold. Machine learning (ML), as a branch of artificial intelligence, can better explore the nonlinear relationships between various indices and samples (Kadkhodaei and Ghasemi 2022a; Shukla et al. 2021), and its prediction will become increasingly accurate over time. Therefore, studying the prediction of rockburst intensity grades based on ML is of great significance (Sun et al. 2021).

Research on ML-based rockburst intensity grade prediction mainly focuses on the rockburst data structure and the algorithms for rockburst intensity grade prediction. In terms of data structure, the dataset may encounter data anomalies or data imbalance (Fig. 1). Considering this problem, scholars proposed various methods to eliminate outliers (Hu et al. 2023b; Li et al. 2018; Zeng et al. 2022). Tan et al. (2021) and Asniar et al. (2022) detected and eliminated data outliers with the aid of the local outlier factor (LOF). Nnamoko and Korkontzelos (2020) and Xue et al. (2020b) identified outliers through interquartile range (IQR). It is generally believed that oversampling is one of the most effective methods to resolve data imbalance (Yi et al. 2022; Zhang et al. 2022a, 2022b). Wang et al. (2023) adopted the random oversampling method, but this method tends to cause overfitting. Li et al. (2023) and Feng et al. (2021) addressed data imbalance by the synthetic minority
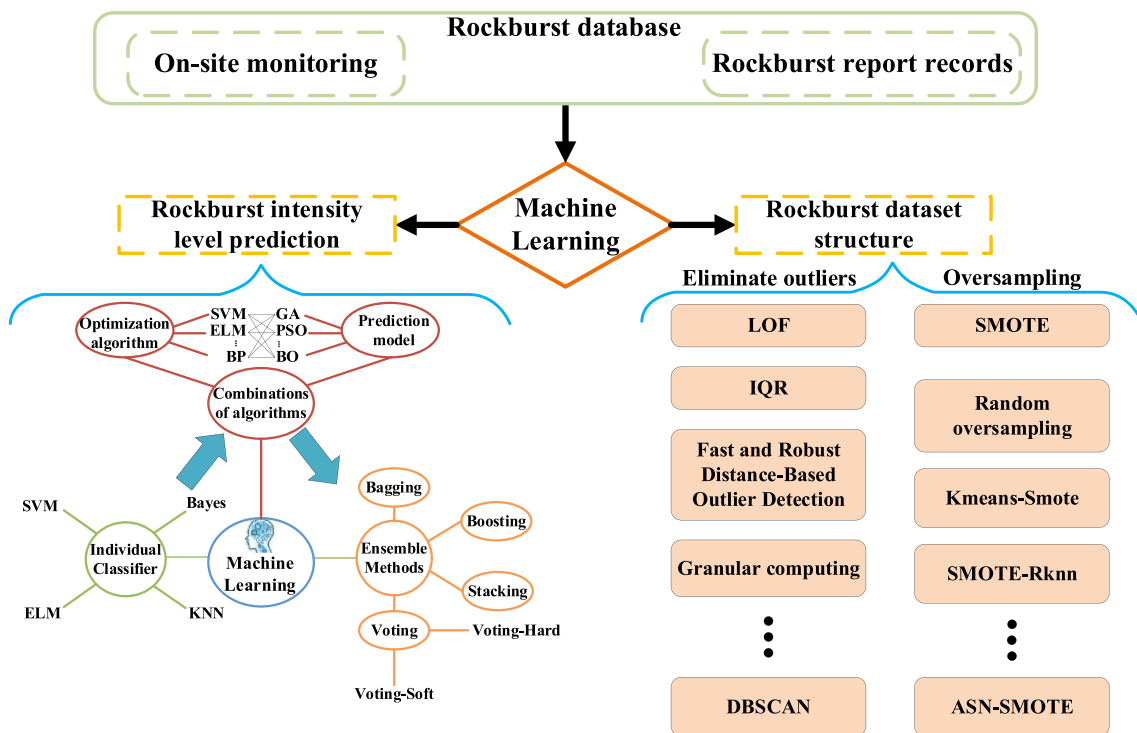

Fig. 1 ML methods for classification and prediction of rockburst intensity grades

oversampling technique (SMOTE). On the other hand, in terms of algorithms, single models were applied predict rockburst intensity grade prediction first (Fig. 1). Kadkhodaei et al. (2022b) using stochastic modeling based on Monte Carlo (MC) simulation predict rockburst potential, the results reveal that stochastic modeling can effectively predict rockburst potential. Ullah et al. (2022) predicted rockburst intensity grades with extreme gradient boosting (XGBoost). Yang et al. (2021) established a prediction model based on the self-organizing feature map (SOFM) neural network. However, since many network parameters in the model need to be set from experience, its performance remains to be improved. Lin et al. (2018) established a cloud model evaluation system based on common influencing factors of rockburst. Ghasemi et al. (2020) for the first time applied C5.0 decision tree algorithm to rockburst prediction; this model can show the relationship between the input and output. Subsequently, it is proved that combination algorithms perform better in prediction than single algorithms, and the optimal parameters of the models can be determined by optimization algorithms. Li et al. (2017) and Xue et al. (2020b) used GA algorithm and PSO algorithm to determine the hyper-parameters in extreme learning machine (ELM) respectively. Ji et al. (2020) determined hyper-parameters in support vector machine (SVM) through genetic algorithm (GA) and built a GA-SVM rockburst prediction model. Liu and Hu (2019) used PSO algorithm to optimize back propagation neural network (BP), probabilistic neural network (PNN), and support vector machine (SVM); the results show that the prediction performance of PSO–PNN model is the best. In recent years, ensemble algorithms have found wide application in predicting rockburst intensity grades. Ensemble algorithms mainly comprise Bagging, Boosting, and Stacking (Fig. 1). Taking six classifiers as individual learners, Liang et al. (2021) created ensemble classifiers with different weightings based on different weighting rules. Voting, an ensemble learning model that follows the majority rule, can integrate performances of individual classifiers better than other ensemble algorithms. Li et al. (2022a) and Tan et al. (2022) built various models of Voting, Bagging, and AdaBoost, and the comparison reveals that Voting-Soft performs the best. However, all the Voting models calculate the voting weights of classifiers according to accuracy, recall, and F1-score, and thus have limitations to some extent. Some scholars determined the weights of classifiers in Voting-Soft by means of optimization algorithms. Zhang et al. (2020) optimized the Voting-Soft model with the beetle antennae search (BAS) algorithm which has a low convergence rate and takes a long time to train the model.

Though ensemble algorithms, especially Voting, are usually superior to single algorithms with respect to prediction, the determination of the type and number of individual learners in Voting is unfounded, and the determination of voting weights of individual learners has a significant impact on model performance. Moreover, the datasets used to train models are likely to be beset by data imbalance or data anomalies. SMOTE can make data structures balanced, but it fails to eliminate class differences between data of different types. Therefore, the following work was done in this study to address problems in existing researches: (1) The method for rockburst data preprocessing was improved. Outliers were eliminated by means of density-based spatial clustering of applications with noise (DBSCAN), and CURE-MeanradiusSMOTE was proposed to obtain a balanced data structure. (2) The AOA-Voting-Soft model was established to predict the rockburst intensity grades. (3) The best classifier combination in Voting was determined by the game theory and method of exhaustion. (4) The weights of individual learners in Voting were determined through arithmetic optimization algorithm (AOA) and fivefold cross-validation (CV).

In this study, a data preprocessing method and an ensemble algorithm prediction model were proposed and applied to rockburst intensity grade prediction. The manuscript is organized as follows. First, rockburst data were preprocessed. Next, the AOA-Voting-Soft ensemble learning model was established. Furthermore, the effect of data preprocessing was verified, and the predictive performance of the proposed model was compared with those of single models and other ensemble models. Finally, the model was applied to rockburst intensity grade prediction of projects like the Dochu La Tunnel, and its generalization ability was verified.

2 Dataset Preparation

2.1 Data Collection

All data in this paper are from worldwide rockburst cases that have been extensively cited and employed. The dataset consists of data of 319 rockburst cases, of which 213 rockburst cases are from Zhou et al. (2016), 21 from Wang et al. (2013), 19 from Zhou et al. (2013), 46 from Dong et al. (2013), and 20 from Xue et al. (2019). These data involve different underground projects such as mines, traffic tunnels, diversion tunnels, and underground chambers. Their diversity ensures that the established ML model has excellent generalization ability.

The occurrence of rockburst is influenced by many factors. When evaluating the intensity grade of rockburst, different scholars choose different parameters as evaluation indicators of criterion for rockburst. Currently, the mainstream input

parameters of criteria for rockburst are the maximum tangential stress of the surrounding rock (σ_θ), the uniaxial compressive strength of the rock (σ_c), the uniaxial tensile strength of the rock (σ_t), the rock brittleness coefficient (σ_c/σ_t), the rock stress coefficient (σ_θ/σ_c), and the elastic strain energy index (W_{et}). Specifically, w_{et} σ_θ can reflect the strata stress characteristics of rockburst. σ_t and σ_c can effectively characterize the lithologic conditions of the surrounding rock in practical engineering. σ_c/σ_t considers the influences of joints and the block size of the rock mass. σ_θ/σ_c, one of the commonly used rockburst evaluation indicators, represents the stress concentration of the surrounding rock after excavation. Rockburst is also related to the energy stored in the rock mass. W_{et} reflects the ability of rock to store elastic energy. The above six indicators reflect the characteristics of rockburst from different aspects (Xu et al. 2022) and are generally acknowledged by scholars worldwide to be closely related to rockburst (Li et al. 2023; Xu et al. 2022; Zhou et al. 2016).

The output indices of the dataset are the rockburst intensity grade. In this study, the database was classified into four grades (Zhou et al. 2012) according to the general standard for rockburst intensity classification, namely, Grades 0, 1, 2, and 3, which represent cases of none rockburst, weak rockburst, moderate rockburst, and strong rockburst, respectively (Table 1).

2.2 Data Analysis

The data in the database built in this study are fairly unbalanced, with data of Grade 0 accounting for 18.2% (58 cases), Grade 1 for 27.6% (88 cases), Grade 2 for 36.1% (115 cases), and Grade 3 for 18.2% (58 cases). The data sizes of none rockburst and strong rockburst are the smallest, and that of moderate rockburst is the largest. The ratio of data sizes of Grades 0–3 is 1:1.5:1.9:1.

The data distribution characteristics and ranges of the four grades are exhibited by the maximum, minimum, mean, and variable coefficient, and descriptive statistics are presented in Table 2. σ_θ characterizes the in situ stress in the rockburst sector, and it is affected by geometry of the opening. The larger the value of σ_θ, the higher the risk and intensity grade of the rockburst. The data of σ_θ in Table 2 increase as the intensity grade rises. For instance, the mean value of σ_θ increases from 25.8 MPa to 115.8 MPa, by 348.4%. σ_c is a base rock mechanics index that shows the hardness of rock. The higher σ_c is, the greater the elastic strain energy is, and the higher the risk and intensity grade of rockburst are. Likewise, the data of σ_c in Table 2 increase as the intensity grade rises. For example, the mean value of σ_c increases from 107.9 to 135.7 MPa, an increase of 25.8%. The initiation and propagation of cracks during rockburst are closely related to the uniaxial tensile strength of rock mass σ_t. The data in Table 2 indicate that σ_t increases with the increase of rockburst intensity grade. It is worth noting that since brittle crack failure occurs in rock mass during rockburst, the brittle coefficient σ_c/σ_t can be deemed a characteristic index of rockburst intensity. In addition, Tang and Wang (2002) proposed a new empirical criterion based on σ_c/σ_t. The values of σ_c/σ_t in Table 2 are barely correlated with the rockburst intensity grade, and the data are highly discrete, which may result from the large number of outliers in the database. The gob-side rock mass with a larger stress coefficient σ_θ/σ_c has poorer stability of equilibrium and thereby is more prone to rockburst. Russenes (1974) proposed a rockburst criterion based on σ_θ/σ_c. This can be verified by the positive correlation between σ_θ/σ_c and the rockburst intensity grade (Table 2). W_{et} represents the ratio of the accumulated energy from elastic deformation of the surrounding rock to the released energy from its plastic deformation. Kidybinski (1981) proposed a classic rockburst classification criterion by testing the energy storage characteristics of rocks. Clearly, if more energy is accumulated while less energy is released, the risk and intensity grade of rockburst would be higher. The mean value of W_{et} in Table 2 rises from 2.9 of Grade 0 to 8.8 of Grade 3, by 206.2%. Moreover, some of the data in Table 2 have a large coefficient of variation. For example, the coefficient of variation of σ_θ/σ_c in data of Grade 0 is 0.8, and that of σ_θ/σ_c in data of Grade 3 is 1. This demonstrates outliers in the database affect the predictive performance of the ML model.

Table 1 Standard for rockburst intensity classification

Rockburst grate	Meaning	Failure characteristics
0	None rockburst	There is no sound of rock burst or rockburst activities
1	Weak rockburst	The surrounding rock is spalled, cracked or stripped, and there is a weak sound but no ejection phenomenon
2	Moderate rockburst	The surrounding rock is more severely deformed and fractured with a considerable amount of rock chip ejecting, loose and sudden destruction, accompanied by a crunchy squeak which often appears in local cavern in the surrounding rock
3	Strong rockburst	The surrounding rock is severely bursted and suddenly thrown or shot into the tunnel, accompanied by strong bursts and roaring sounds, air jets, the continuity of storm phenomena, and the rapid expansion into deep surrounding rock

Pairs plots of the database built in this study are depicted in Fig. 2, where different colors of scattered points denote data of different rockburst grades. Kernel density maps of variables are on the diagonal lines of pairs plots, and the scatter diagrams of correlation between the two variables are in the non-diagonal areas. Besides, the Pearson correlation coefficient of the two variables is presented. Discrete points can be observed in all these correlation scatter diagrams. Meanwhile, data points of different rockburst grades overlap each other significantly in correlation scatter diagrams, which influences the predictive performance of ML models. In Kernel density maps in Fig. 2, data of σ_θ are concentrated in the range of 30–70 MPa (accounting for 66.7% of the total data), σ_c in 90–130 MPa (48.6%), σ_t in 3–7 MPa (55.2%), and W_{et} in 3–7 MPa (80.3%). The maximum Pearson correlation coefficient is 0.49, which means variables are mutually independent .

2.3 Data Preprocessing

2.3.1 Eliminating Outliers by DBSCAN

According to the analysis in Sect. 2.2, outlier points in the original data have a considerable impact on the predictive performance of the ML model. Therefore, abnormal data in the dataset were detected and eliminated by means of DBSCAN. DBSCAN, one of the density-based clustering algorithms, is used to determine the clusters of any shape

that may exist in a given dataset (Ester et al. 1996). As this algorithm does not require human efforts to determine the number of clusters in advance, it can tackle with errors resulting from unreasonable setting of the cluster number.

DBSCAN only requires two parameters to create a new cluster, namely the radius of the cluster (eps) and the minimum number of points (MinPts) within a circle of radius eps (Hao et al. 2015). And it classifies all points into three types: core points, border points, and noise points. In Fig. 3a, A is the core point because the number of points within its cluster radius eps is more than or equal to MinPts; B is the border point because it is not a core point but it is within the cluster of a core point; a point that belongs to neither the type of Point A nor the type of Point B is a noise point, e.g., Point C in Fig. 3a.

DBSCAN is quite sensitive to the eps value. Specifically, an excessively small eps value may lead to an expanded range of noise points, whereas an excessively large eps value may result in a poor detection effect of outliers. The *K*-distance graph technique is used to determine the eps (Starczewski et al. 2020). As illustrated in Fig. 3b, with the number of nearest neighbors of each point in the rockburst database *K* regarded as the abscissa, *K*-distance was calculated and taken as the ordinate. Subsequently, k_{dist} graph was plotted in ascending order, and the ordinate of its maximum curvature point is the best eps. The value of MinPts was determined based on Eq. (1) (Arafa et al. 2022).
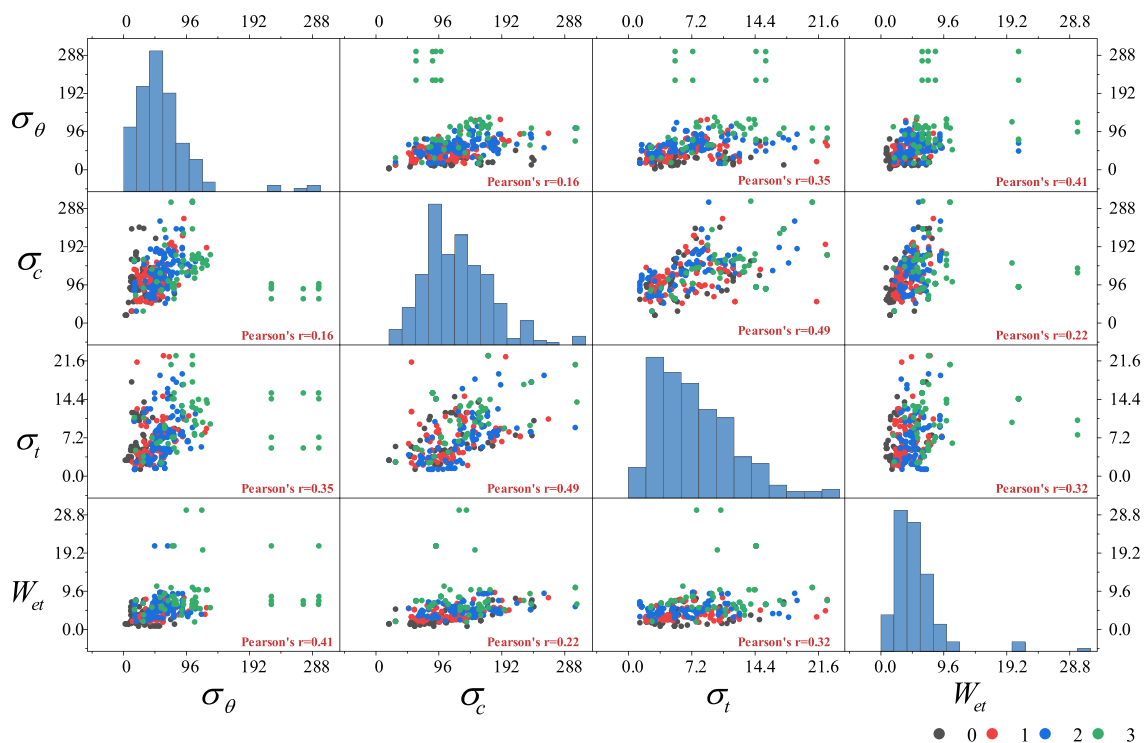

Fig. 2 Matrix scatter plot

Table 2 Statistical parameters of different rockburst grades

Rockburst grade	Rockburst feature						
	Statistical parameter	σ_θ MPa	σ_c MPa	σ_t MPa	σ_c/σ_t	σ_θ/σ_c	W_{et}
0	Maximum	77.69	241.00	17.66	47.90	1.05	0.81
	Minimum	2.60	20.00	1.30	5.38	0.05	7.80
	Mean	25.83	107.86	6.20	21.13	0.29	2.89
	Coefficient	0.61	0.48	0.59	0.56	0.83	0.65
1	Maximum	126.70	263.00	22.60	69.60	0.90	9.30
	Minimum	13.50	30.00	1.90	2.52	0.12	1.80
	Mean	50.19	117.91	7.19	20.36	0.46	3.98
	Coefficient	0.44	0.39	0.58	0.51	0.41	0.42
2	Maximum	118.77	304.00	19.17	80.00	1.27	21.00
	Minimum	16.30	30.00	1.30	5.53	0.11	2.03
	Mean	58.32	130.74	7.12	25.90	0.48	5.55
	Coefficient	0.38	0.37	0.61	0.71	0.40	0.48
3	Maximum	297.80	306.58	22.60	32.20	4.87	30.00
	Minimum	16.43	30.00	2.50	5.53	0.11	1.90
	Mean	115.81	135.69	10.93	13.91	1.10	8.84
	Coefficient	0.65	0.44	0.45	0.42	1.01	0.67

Fig. 3 DBSCAN. **a** DBSCAN clustering; **b** K-distance graph
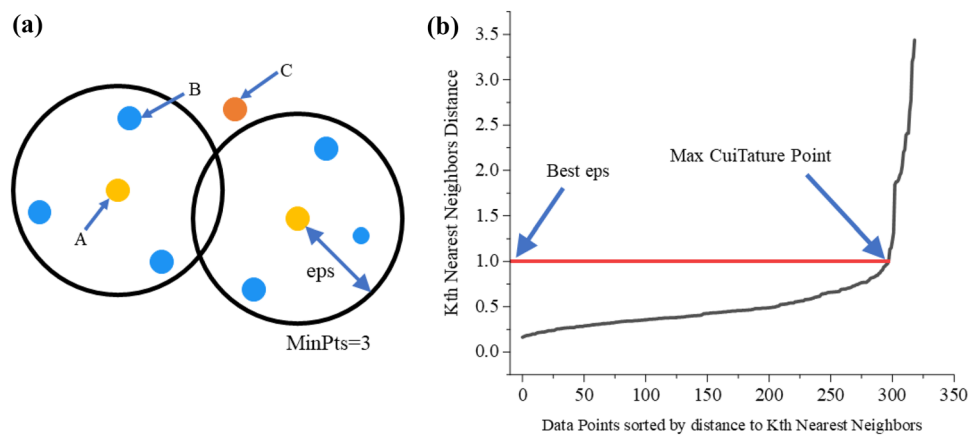
$$\text{MinPts} = \ln N, \tag{1}$$

where N is the number of samples in the rockburst database.

The results indicate that the values of MinPts and eps are 6 and 1.014, respectively. Moreover, two samples of moderate rockburst and fifteen samples of strong rockburst are removed from the original database by means of DBSCAN.

2.3.2 Eliminating Dataset Imbalance Through Cure-MeanradiusSMOTE

According to the analysis in Sect. 2.2, the initial rockburst dataset is imbalanced. Consequently, ML models may mistake minority-class samples as majority-class ones, thereby weakening the predictive performance of ML models. Hence, the

rockburst database needs to be oversampled. It is noteworthy that oversampling should be performed on rockburst datasets where the outliers have been omitted so as to prevent the generation of new outliers during the oversampling process. The CURE-MeanradiusSMOTE method proposed in this study is based on Kmeans-SMOTE, a common oversampling method.

CURE-MeanradiusSMOTE includes three stages, i.e., clustering, filtering, and oversampling. In the clustering stage, clustering is conducted on rockburst data by the clustering using representatives (CURE) algorithm. In the oversampling stage, the radius and geometric center are considered when new data are generated. Compared with the Kmeans-SMOTE, the proposed algorithm is more efficient for datasets of any shape, and can generate new data that are more likely to be distributed around the average radius of minority-class samples. In this way, it enhances ML models' ability to recognize decision boundaries.
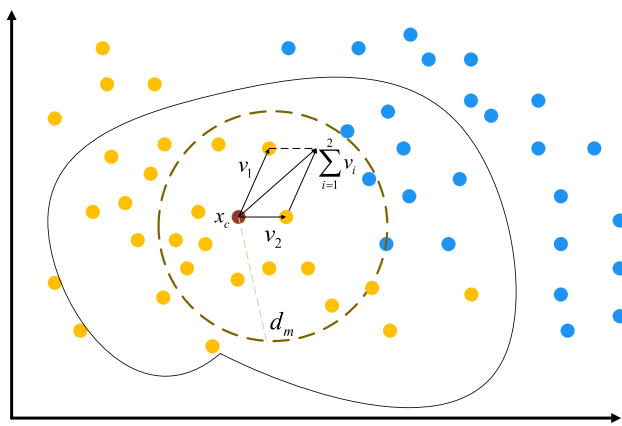
Fig. 4 Schematic diagram of CURE-MeanradiusSMOTE

The CURE-MeanradiusSMOTE schematic is shown in Fig. 4, CURE-MeanradiusSMOTE schematic is shown, the specific procedure of CURE-MeanradiusSMOTE is introduced as follows:

1. Clustering is conducted on rockburst data by the CURE algorithm, and clusters with a high proportion of minority-class samples are retained. The center of these clusters is calculated and denoted by x_c.
2. The weights of rockburst characteristic indices are calculated by principal component analysis.
3. The weighted Euclidean distances between points in retained clusters in Step 1 and the cluster center are calculated, and then their average is calculated and represented by d_m.
4. k minority-class samples are randomly selected from clusters retained in Step 1, and then k vectors v_i from the sample center to samples are calculated. Resultant vector $\sum_{i=0}^{k} v_i$ of k vectors is calculated.
5. The distance between the new sample and the cluster center x_c is determined based on the average distance d_m and the parameter θ. The new sample is created according to Eq. (2):

$$x_{new} = x_c + r * \sum_{i=0}^{k} v_i \quad r \sim (\frac{d_m}{\theta}, d_m). \tag{2}$$

6. Steps 3–5 are repeated until the number of majority-class and minority-class samples becomes balanced.

After data were made balanced, 55 data of none rockburst, 26 data of weak rockburst, and 71 data of strong rockburst were generated. The new rockburst database has a total of 452 rockburst data, the ratio of data of Grades 0–3 being 1:1:1:1 .

3 ML Modeling

3.1 Fundamental Theory of AOA and Voting

3.1.1 Voting-Soft Algorithm

Ensemble learning, also known as a multi-classifier system or committee-based learning, is an algorithm that creates and combines multiple learners to complete learning tasks (Wang et al. 2020). Voting is a kind of ensemble learning that consists of multiple heterologous individual classifiers (Rojarath and Songpan 2021). The schematic diagram of Voting is depicted in Fig. 5. Voting is carried out in two ways: Voting-Hard and Voting-Soft. For Voting-Hard, the final result is determined in line with the majority rule. As presented in Fig. 5, three out of five classifiers choose Type B, so the prediction result is B. For Voting-Soft, the average of probability that all model prediction samples are of a certain type is regarded as the standard, and then the type with the highest probability is the final prediction result. As can be seen from Fig. 5, since the average of probability that five classifiers choose Type A is 0.6 and that for Type B is 0.3, the prediction result is A. It should be noted that the voting weights of C1–C5 in Fig. 5 are all set to 1. Unlike Voting-Hard, Voting-Soft requires each individual learner to calculate the probability of each type. It can give higher weights to individual learners with excellent predictive performance, so it typically outperforms Voting-Hard in terms of prediction.

3.1.2 AOA

AOA is a population-based metaheuristic algorithm proposed by Abualigah et al. in (2021). The inspiration for AOA comes from the application of arithmetic operators (addition, subtraction, multiplication, and division) in solving arithmetic problems. The algorithm can deal with optimization without calculating derivatives. It is simple, and boasts few control parameters and excellent output performance. AOA consists of three stages, i.e., initialization, exploration, and exploitation, and its schematic diagram is shown in Fig. 6.

Step 1 Initialization. Optimization strategies are selected with the aid of math optimizer accelerated (MOA).

AOA performs global search when r_1 is greater than MOA, while it conducts local search when r_1 is smaller than MOA. MOA is calculated based on Eq. (3):

$$MOA(C_{Iter}) = Min + C_{Iter} \times \left(\frac{Max - Min}{M_{Iter}} \right), \tag{3}$$

where $MOA(C_{Iter})$ is the function value at the tth iteration; C_{Iter} is the present iteration whose value lies between 1 and
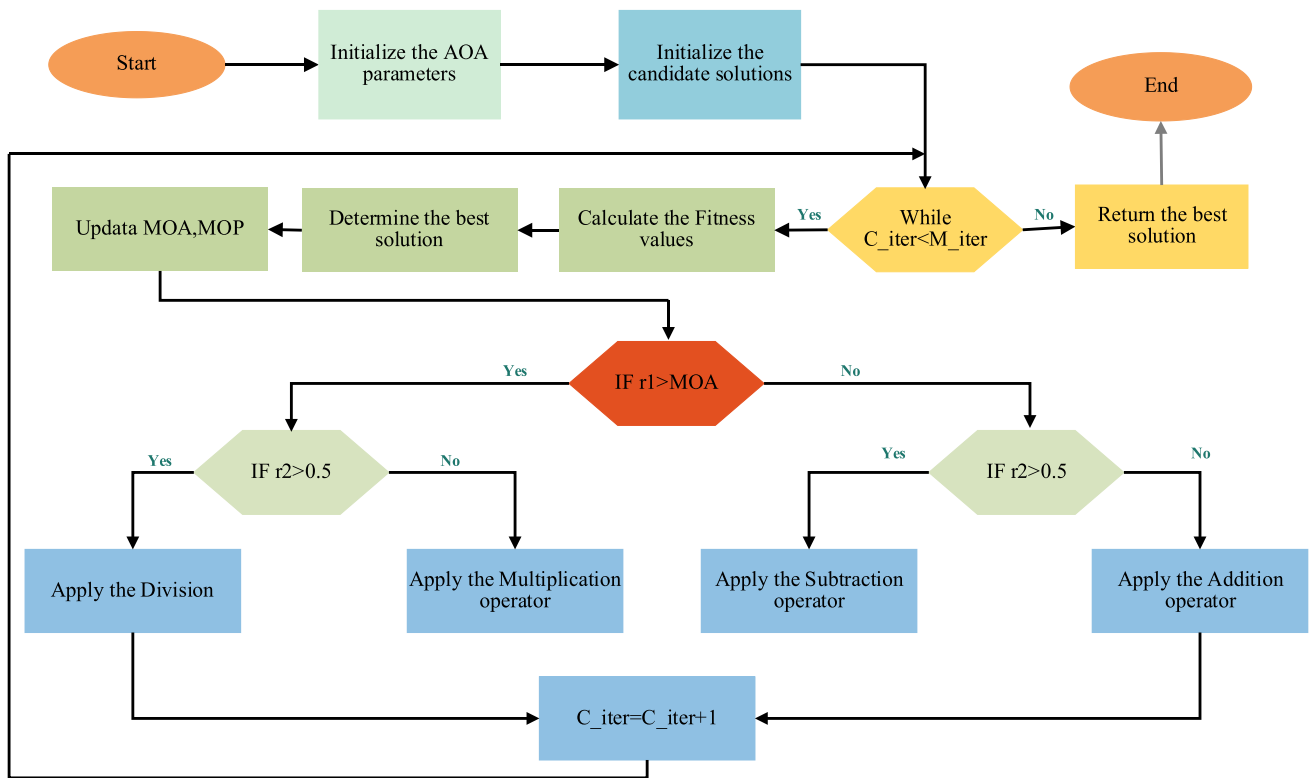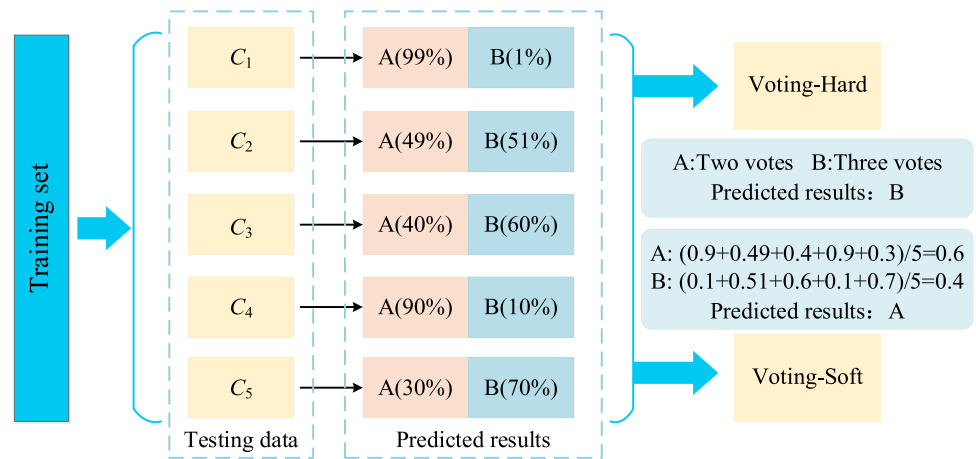
Fig. 5 Schematic diagram of Voting

Fig. 6 Flowchart of AOA

M_{Iter}; Max and Min are the maximum and minimum values of the acceleration function, respectively.

Step 2 Exploration.

AOA carries out diversified global search by means of multiplication or division in the exploration stage. It performs division when r_2 is smaller than 0.5, while it conducts multiplication when r_2 is greater than or equal to 0.5. r_2 is a random number, and $r_2 \in [0,1]$. The location update strategy is:

$$x_{i,j}(C_{Iter}+1) = \begin{cases} \text{best}(x_j) \div (E+\varepsilon) \times \left[\left(\text{UB}_j - \text{LB}_j\right) \times \mu + \text{LB}_j\right] & r_2 < 0.5 \\ \text{best}(x_j) \times \text{MOP} \times \left[\left(\text{UB}_j - \text{LB}_j\right) \times \mu + \text{LB}_j\right] & \text{otherwise,} \end{cases} \tag{4}$$

where $x_{i,j}(C_{\text{Iter}} + 1)$ is the ith solution in the next iteration; $x_{i,j}(C_{\text{Iter}})$ is the jth position of the ith solution in the present iteration; best(x_j) is the jth position of the best solution obtained so far; ε is a decimal number; UB_j and LB_j are the upper and lower limits of the jth position, respectively; μ is the control parameter that adjusts the searching process. Multi-objective optimization problem (MOP), which is the coefficient of mathematical optimization rate, is calculated according to Eq. (5):

$$\text{MOP}(C_{\text{Iter}}) = 1 - \frac{C_{\text{Iter}}^{1/a}}{M_{\text{Iter}}^{1/a}}, \tag{5}$$

where $\text{MOP}(C_{\text{Iter}})$ is the function value at the tth iteration; M_{Iter} is the maximum number of iterations; a is the sensitive parameter, which defines the development accuracy of iteration.

Step 3 Exploitation.

AOA searches for the optimal solution in many dense regions in the exploration stage through the strategy of search by subtraction or addition. r_3 is a random number, and $r_3 \in [0,1]$. When r_3 is smaller than 0.5, the strategy of search by subtraction is taken; when r_3 is greater than 0.5, the strategy of search by addition is adopted. The location update strategy is:

$$x_{i,j}(C_{\text{Iter}} + 1) = \begin{cases} \text{best}(x_j) - \text{MOP} \times \left[\left(UB_j - LB_j \right) \times \mu + LB_j \right] & r_3 < 0.5 \\ \text{best}(x_j) + \text{MOP} \times \left[\left(UB_j - LB_j \right) \times \mu + LB_j \right] & \text{otherwise.} \end{cases} \tag{6}$$

3.2 CV

CV is a common method for evaluating the generalization ability of ML models. In CV, the dataset is divided many times, and several models need to be trained. K-fold CV is the most common CV method (Wong and Yeh 2020). In this study, k was set to 5 according to previous studies

(Motsinger and Ritchie 2006). In addition, the training set was randomly and evenly divided into five subsets, of which four were used for training models and determining their hyper-parameters, and one was for verifying these models' generalization ability. The above process was repeated five times to obtain five different hyper-parameters, whose average was considered the final hyper-parameter (Fig. 7).

3.3 Modeling and Hyper-parameter Tuning

3.3.1 Game Theory-Based Model Combination

According to Sect. 3.1.1, though Voting-Soft has some advantages over Voting-Hard, it requires each individual learner to have remarkable predictive performance and obtain diversified predictive results, which is extremely difficult to achieve. Hence, individual learners in Voting-Soft are required to take into consideration both accuracy and diversity. The accuracy of models shows their predictive performance, and their diversity can be seen from the correlation between their predictive results. In this study, the conflict between diversity and accuracy of models was mitigated by using the combination weighting method of game theory (Feng et al. 2019). Besides, the best classifier combination in Voting-Soft was determined through the exhaustive search method. The flow chart is shown in Fig. 8, the steps of calculation are as follows:

1. Prediction is performed by N ML models to obtain their prediction results and accuracy.
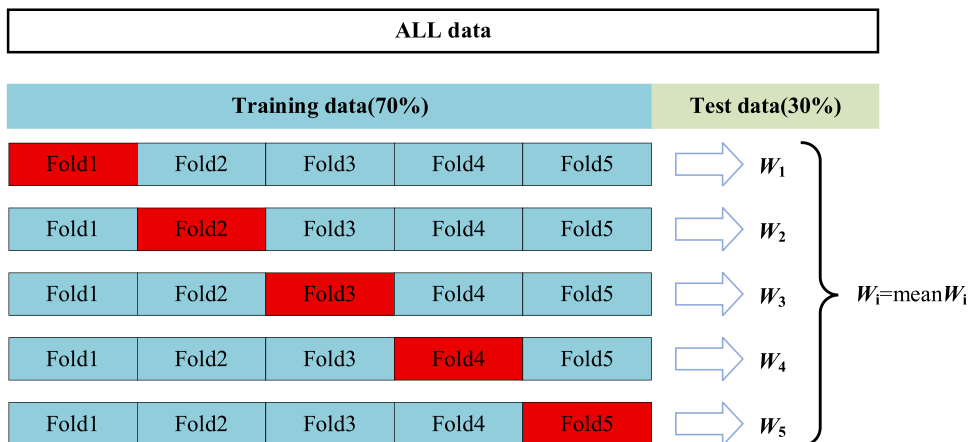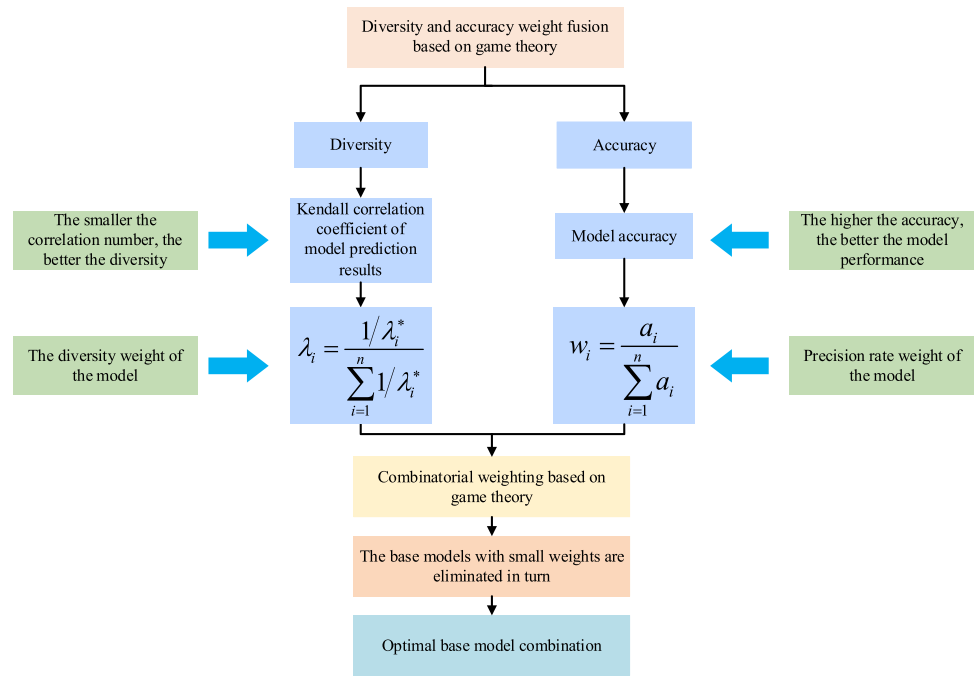
Fig. 7 Fivefold CV

Fig. 8 Schematic diagram of weight calculation by the game theory

2. The diversity and accuracy weights of each individual learner are calculated.

First, the Kendall correlation coefficients (Kendall) of prediction results of N models are calculated. The sum of correlation coefficients between an individual learner and others is averaged to obtain the correlation between this individual learner and others. The smaller the correlation is, the greater the difference between this individual learner and others. The diversity weight of the model is calculated through Eqs. (7) and (8):

$$\lambda_i^* = \frac{\sum_{i=1}^{n-1} b_i \Big/ n-1}{\sum_{i=1}^{n} (\sum_{i=1}^{n-1} b_i \Big/ n-1)}, \tag{7}$$

$$\lambda_i = \frac{1/\lambda_i^*}{\sum_{i=1}^{n} 1/\lambda_i^*}, \tag{8}$$

where b is the correlation between a model and others; n is the number of models; λ_i^* is the diversity coefficient of the model; λ_i is the diversity weight of the model (Fig. 8).

The greater the accuracy of a model is, the better its performance is. The accuracy weight of a model is calculated by Eq. (9):

$$w_i = a_i \Big/ \sum_{i=1}^{n} a_i, \tag{9}$$

where a is the accuracy of the model; and w_i is the accuracy weight of the model.

3. The comprehensive weight of each individual learner is calculated by the combination weighting method of game theory according to Eq. (10):

$$\begin{pmatrix} \lambda_1 \lambda_1^T & \lambda_1 a^T \\ a\lambda_1^T & aa^T \end{pmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \lambda_1^T \\ aa^T \end{bmatrix}, \tag{10}$$

$$W = b_1^* \lambda_1^T + b_2^* \lambda_2^T, \tag{11}$$

where λ is the diversity weight matrix; a is the accuracy weight matrix; b_1 and b_2 are the linear combination coefficients to be solved; and W is the combination weight matrix. b_1^* and b_2^* are calculated through Eqs. (12) and (13):

$$b_1^* = \frac{b_1}{b_1 + b_2}, \tag{12}$$

$$b_2^* = \frac{b_2}{b_1 + b_2}. \tag{13}$$

4. N models are placed into Voting for training and prediction, and models with small weights are eliminated in sequence to select the optimal model combination for rockburst intensity grade prediction in Voting.

3.3.2 Modeling

The dataset, consisting of 419 data, is split into the training and test sets. ML models are trained on the training set, and their generalization ability is tested on the test set. Because the type, number, voting weight, and predictive capability of individual learners in Voting-Soft models influence their predictive performance, in this study, the optimal combination of individual learners was determined based on the game theory, and the hyper-parameter and voting weight of each individual learner were determined by means of AOA and fivefold CV. Furthermore, the maximum number of iterations was set to 100, with ten individuals in each iteration. The sensitive parameter was 5, and the control parameter was 0.499. All parameters of AOA were determined through experimental testing. The modeling process is displayed in Fig. 9, and its steps are as follows:

1. Data are collected and analyzed.
2. Outliers are detected and eliminated by DBSCAN, and the data structure is made balanced by MeanRadius-SMOTE.
3. The preprocessed rockburst database is split into the training and test sets at a ratio of 7:3.
4. Hyper-parameters of base classifiers are determined by means of AOA and fivefold CV.

5. The combination weight of each individual learner is calculated through the combination weighting method of game theory.
6. The Voting-Soft model is built, and voting weights of base classifiers in Voting-Soft are determined through AOA and fivefold CV.
7. Individual learners with low weights are removed in sequence.
8. Whether the termination condition is met is determined. If it is, the Voting-Soft-AOA model is established based on the optimal combination of individual learners and voting weight; otherwise, Step 6 is performed.
9. The generalization ability is tested, and importance analysis is conducted on characteristic variables.

3.4 Model Evaluation

Accuracy and recall, which are common indicators to evaluate predictive ability of classification models, are calculated with the confusion matrix (Fig. 10). The confusion matrix is widely adopted for evaluating the predictive accuracy of classification models in binary classification. In the confusion matrix for multi-class classification, each class is deemed positive in turn, and others negative. In this way, multi-class classification is converted into binary
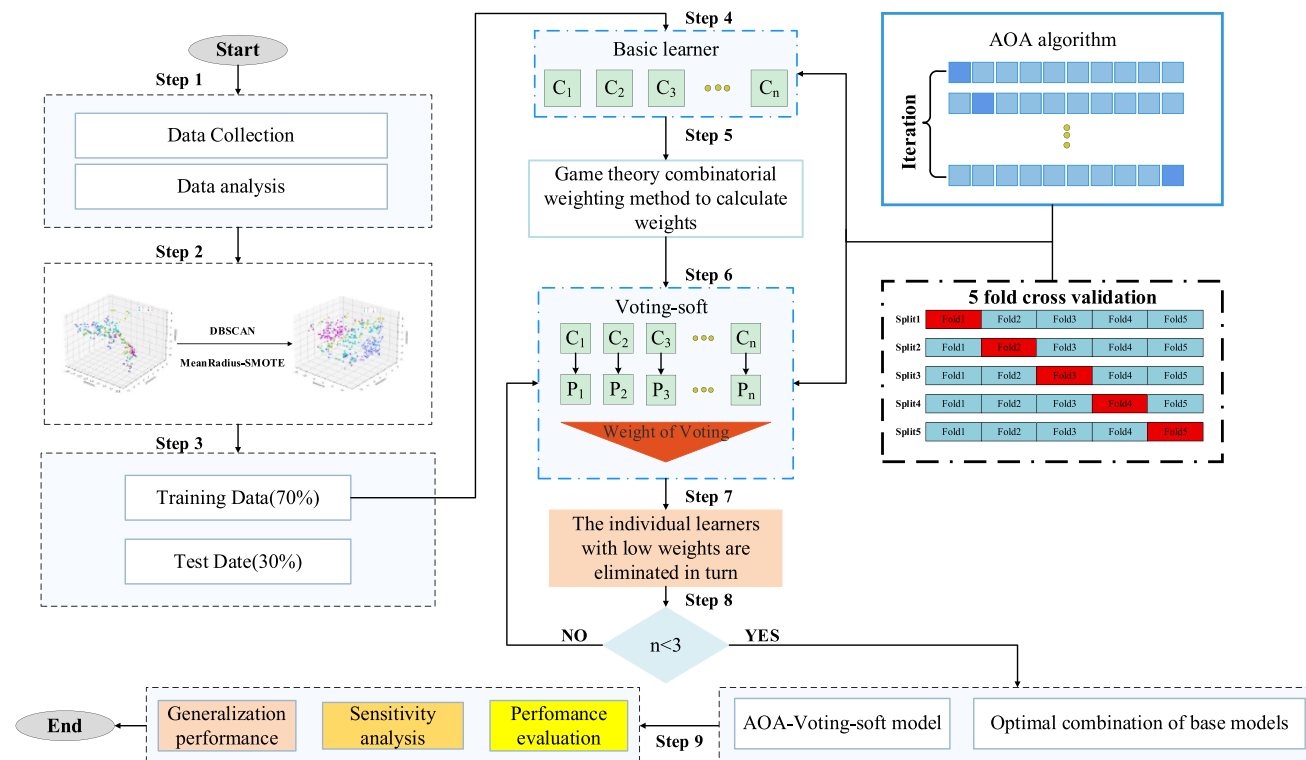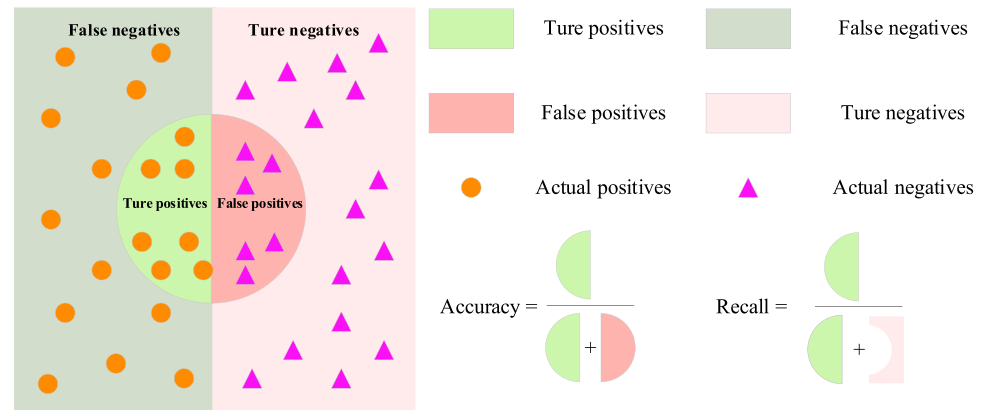

Fig. 9 Flowchart of modeling

Fig. 10 Schematic diagram of confusion matrix

classification (Trajdos and Kurzynski 2018). The schematic diagram is exhibited in Fig. 10.

In addition to the above metrics, the receiver operating characteristic (ROC) and a larger area under the curve (AUC) evaluation metrics were applied in this study. False-positive rate and true-positive rate at different thresholds need to be calculated to draw a ROC curve whose abscissa and ordinate are the false-positive rate and the true-positive rate, respectively. AUC usually means higher classification accuracy. ROC and AUC can demonstrate the false-positive rate and the true-positive rate comprehensively. It is noteworthy that AUC and ROC can only be used for binary classification. Curves of each class are plotted for rockburst intensity grade prediction by binary decomposition. Four ROC curves were drawn, and four AUC values were calculated. Subsequently, these four ROC curves were averaged to obtain the curve of multi-class classification, and the four AUC values were also averaged to obtain the value of multi-class classification. Generally, a higher AUC value is indicative of better predictive performance of a model (Chen et al. 2022).

4 Results and Discussion

4.1 Verification of Data Preprocessing Effect

Voting in this study contains seven heterogenous individual classifiers, including three ensemble learning algorithms (XGBoost, GBDT, and RF), one neural network algorithm (MLP), and three single classical ML algorithms (KNN, SVM, and Bayesian). For the purpose of verifying the effect of the data preprocessing method (DBSCAN and Cure-MeanradiusSMOTE) used in this study, the prediction effects of these seven individual learners in the original rockburst database and the preprocessed rockburst database were compared by regarding model accuracy as the evaluation indicator. As presented in Table 3, the prediction accuracy of GBDT in the original rockburst data is 0.677. After being processed by SMOTE (Chawla et al. 2002; Fernandez et al. 2018), Kmeans-SMOTE (Douzas et al. 2018)and the data preprocessing method presented in this study, the model's prediction accuracy is raised by 5.8%, 7.5%, and 11.7%, respectively. Obviously, the algorithm accuracy of seven individual learners is improved to varying degrees after data are preprocessed by the method presented in this study.

For the purpose of better demonstrating the preprocessing effect of rockburst data, dimensionality reduction and

Table 3 Prediction accuracy of individual learners with different data preprocessing methods

Model	Original database	Database processed by Smote	Database processed by Kmeans-SMOTE	Database processed by the method presented in this study
GBDT	0.677	0.735	0.752	0.794
XGBoost	0.635	0.757	0.752	0.765
SVM	0.531	0.735	0.659	0.794
MLP	0.542	0.706	0.667	0.772
KNN	0.531	0.721	0.674	0.721
RF	0.688	0.743	0.723	0.750
Bayesian	0.646	0.676	0.617	0.728

(a)

(b)

Original

SMOTE
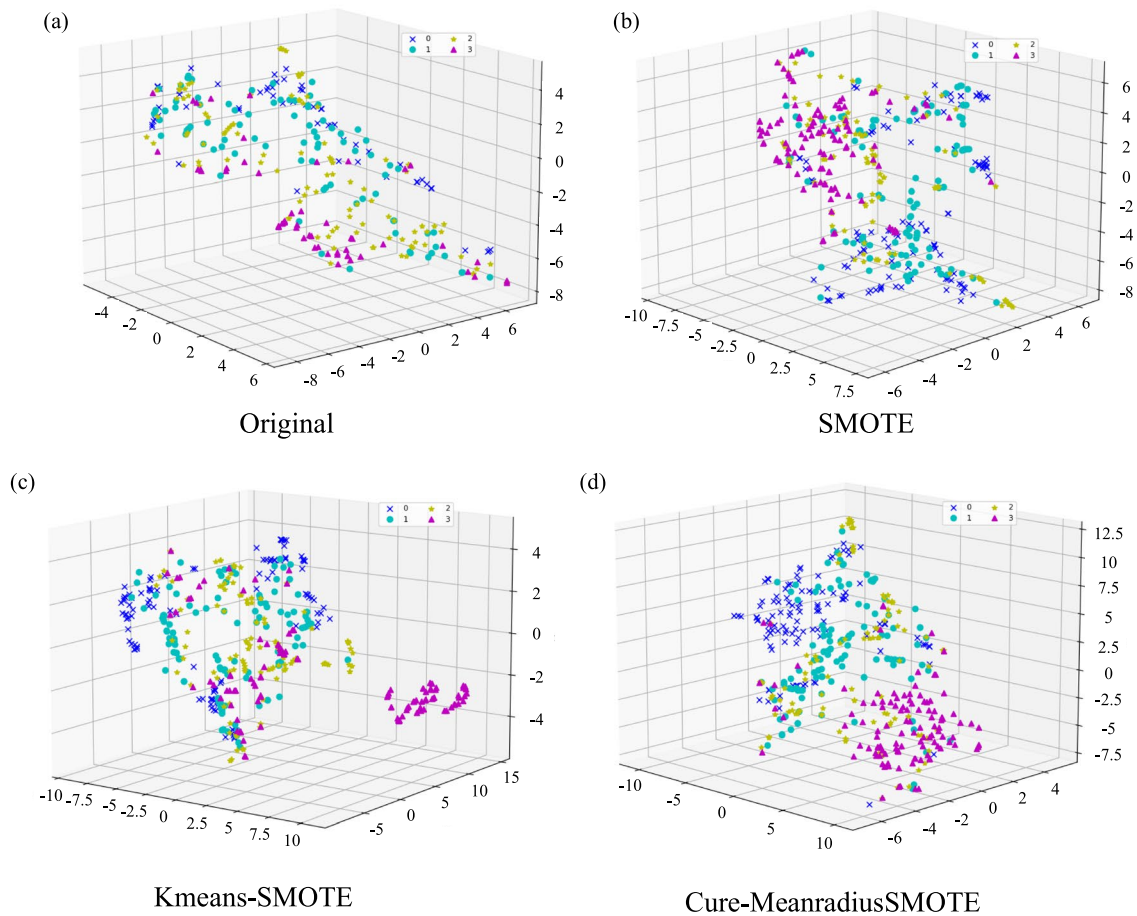
(c)

(d)

Kmeans-SMOTE

Cure-MeanradiusSMOTE

Fig. 11 3D spatial distribution map of rockburst data. **a** Original data; **b** Data preprocessed by SMOTE; **c** Data preprocessed by KMeans-SMOTE; **d** Data processed by the method presented in this study

visualization were conducted on rockburst data by t-distributed stochastic neighbor embedding (TSNE). TSNE is a visualization tool that can maintain data separability of low-dimensional spaces in high-dimensional ones (Zhu et al. 2019). The distribution of the rockburst dataset before and after preprocessing is depicted in Fig. 11. It can be found from Fig. 11 that many outliers exist in the original rockburst data, and data of all intensity grades are mixed together. SMOTE creates new classes in the area of majority-class samples, and the new samples generated may be outliers. Samples created by KMeans-SMOTE seriously overlap, which may lead to overfitting. As presented in Fig. 11d, new samples generated by the method in this study are distributed uniformly in the space of minority class, and rockburst data of an intensity grade cluster together without outliers.

Table 4 Hyper-parameters of individual learners

Model	Hyper-parameter	Value
RF	Number of DTS	96
	DTS maximum depth	100
MLP	Alpha	0.618
	Learning rate	0.004
XGBoost	Number of boosting rounds	17
	Learning rate	0.01
GBDT	Maximum depth	3
	Number of boosting iterations	88
	Learning rate	0.133
KNN	Nearest neighbor number	11
SVM	Penalty coefficient	1.013
	Kernel function	RBF

Fig. 12 Hyper-parameter tuning
for individual learners
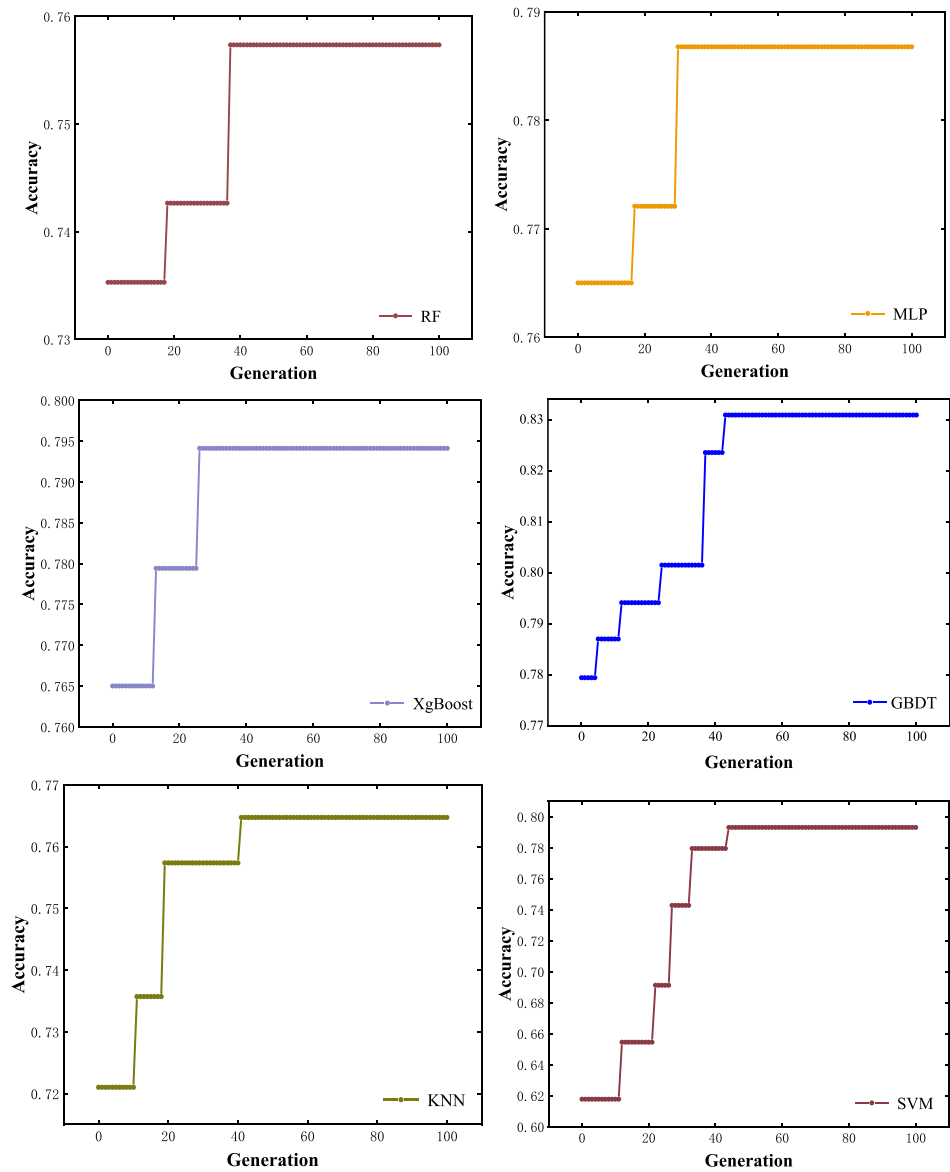
4.2 Hyper-parameter Tuning for Base Classifiers

Hyper-parameters of seven individual learners in Voting were optimized by means of AOA and fivefold CV. Bayesian does not need optimization owing to its particularity. Accuracy was set as the objective function of AOA to find optimal hyper-parameters for other individual classifiers. Hyper-parameters and optimal values of classifiers are displayed in Table 4.

Figure 12 shows the iteration process in which AOA finds the maximum accuracy. Due to randomness of initial points in AOA, objective functions of different models

Table 5 Comprehensive weights of individual learners

Model	Diversity weight	Accuracy weight	Combination weight
GBDT	0.144	0.152	0.150
XGB	0.146	0.146	0.146
SVM	0.138	0.145	0.143
MLP	0.139	0.144	0.143
KNN	0.144	0.140	0.141
RF	0.144	0.139	0.140
Bayesian	0.146	0.133	0.136

have different values in the initial state. For instance, in SVM, the accuracy of objective functions increases gradually as the iteration is performed, which means AOA is effective in tuning SVM architecture. The highest accuracy is 0.62 at the 1st iteration, while it increases to 0.79 at the 50th iteration. At this time, the penalty coefficient is 1.012567, and the Kernel function is the radial basis function (RBF).

4.3 Optimal Combination of Base Classifiers

To calculate the combination weight of each individual classifier, the diversity weight, accuracy weight, and combination weight of each individual classifier were calculated with the method introduced in Sect. 3.3.1 based on the prediction results of optimized individual classifiers. The calculation results in Table 5 indicate that Bayesian has the lowest combination weight, 0.136, while GBDT has the highest combination weight, 0.150.

Voting-Soft-AOA models were built based on seven classifiers, and then those with low weights were eliminated sequentially by the exhaustive search method. According to the results in Fig. 13, the Voting-Soft-AOA model built with seven base classifiers has the best performance, with an overall accuracy of 0.875. As the number of base classifiers reduces, the performance of Voting-Soft-AOA models becomes progressively worse. The Voting-Soft-AOA model built with three base classifiers has the lowest overall accuracy, 0.80147. Hence, the Voting-Soft-AOA model build with seven base classifiers was ultimately chosen in this study.
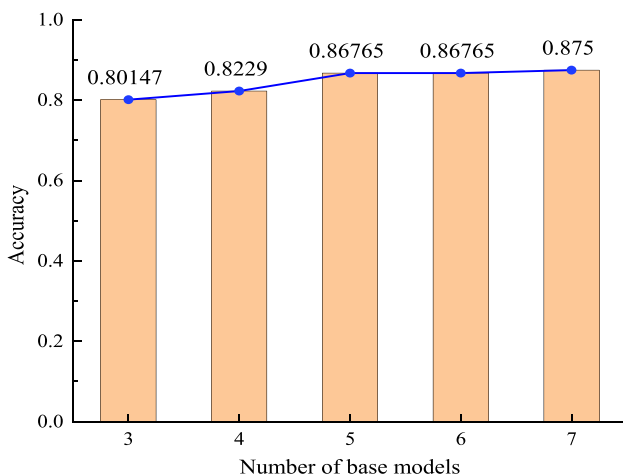

Fig. 13 Accuracy of Voting-Soft-AOA models with different classifier combinations in the test set
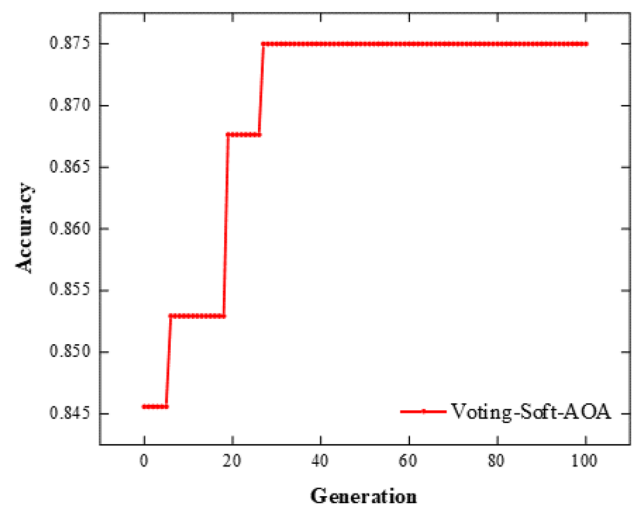

Fig. 14 Schematic diagram of Voting-Soft-AOA iteration

4.4 Voting Weight Tuning for Base Classifiers

Figure 14 displays the iterative process in which AOA finds the maximum accuracy. It can be seen from Fig. 14 that the accuracy increases gradually as AOA iterates, which means AOA is effective in optimizing weights of base classifiers. The first iteration witnesses the lowest accuracy of 0.8459, and it rises to 0.875 in the 27th iteration. Hyper-parameters of Voting-Soft are listed in Table 6.

4.5 Prediction Performance

4.5.1 Performance Comparison Between Ensemble and Individual Classifiers

Table 7 reveals F1-score, recall, and accuracy of the ensemble classifier (Voting-Soft-AOA), and individual classifiers at three rockburst intensity grades (grades 0–3). In prediction of none rockburst, Voting-Soft-AOA has the highest F1-score, recall, and accuracy of 0.96, 0.91, and 1, respectively. In prediction of weak rockburst, Voting-Soft-AOA has the highest F1-score, recall, and accuracy of 0.79, 0.88, and 0.72, respectively. In prediction of moderate rockburst,

Table 6 Voting weights of base classifiers

Base classifier	Weight
XGBoost	0.199
RF	0.199
GBDT	0.164
Bayesian	0.011
KNN	0.199
SVM	0.199
MLP	0.029

Table 7 Performance comparison of Voting-Soft-AOA and other ML models

Model	Metrics	0 (None)	1 (Weak)	2 (Moderate)	3 (Strong)
Voting-Soft-AOA	F1	0.96	0.79	0.77	0.88
	Recall	0.91	0.88	0.72	0.91
	Accuracy	1	0.72	0.82	0.86
GBDT	F1	0.9	0.77	0.76	0.87
	Recall	0.82	0.88	0.75	0.88
	Accuracy	1	0.68	0.77	0.85
XGBoost	F1	0.89	0.72	0.71	0.84
	Recall	0.8	0.81	0.75	0.82
	Accuracy	1	0.64	0.67	0.87
SVM	F1	0.92	0.72	0.62	0.85
	Recall	0.91	0.77	0.59	0.85
	Accuracy	0.93	0.67	0.66	0.85
MLP	F1	0.94	0.71	0.51	0.87
	Recall	0.91	0.88	0.41	0.91
	Accuracy	0.98	0.59	0.68	0.83
KNN	F1	0.86	0.7	0.61	0.83
	Recall	0.8	0.81	0.53	0.91
	Accuracy	0.92	0.62	0.71	0.77
RF	F1	0.89	0.64	0.62	0.791
	Recall	0.89	0.69	0.62	0.76
	Accuracy	0.91	0.6	0.62	0.83
Bayesian	F1	0.85	0.57	0.51	0.83
	Recall	0.91	0.58	0.41	0.91
	Accuracy	0.8	0.56	0.68	0.77

Fig. 15 Accuracy of Voting-Soft-AOA and other ML models

Fig. 16 ROC curves and AUC values of seven individual classifiers and Voting-Soft-AOA on the test set
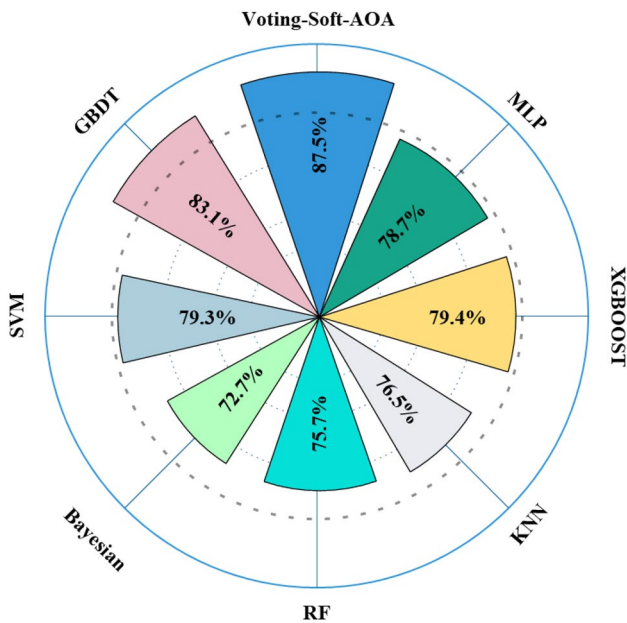
GBDT and XGBoost have the highest recall of 0.75, while Voting-Soft-AOA has the highest F1-score and accuracy of 0.77 and 0.82, respectively. In prediction of strong rockburst,
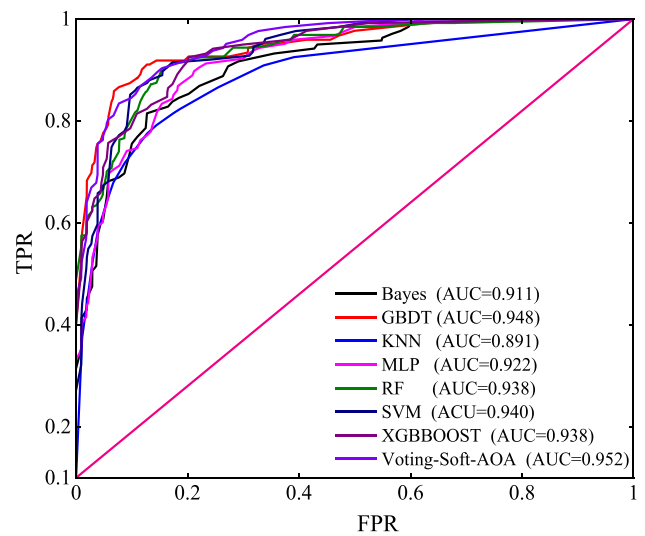
XGBoost has the highest accuracy of 0.87, while Voting-Soft-AOA has the highest F1-score and recall of 0.91 and 0.86, respectively. Overall, Voting-Soft-AOA shows the best predictive performance at all the three rockburst intensity grades.
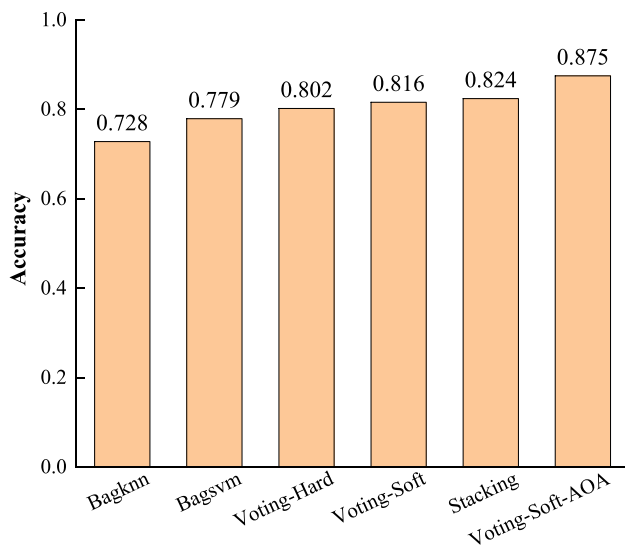
Fig. 17 Accuracy of different ensemble algorithms on the test set

Fig. 18 F1-scores of different ensemble algorithms on the test set

Figure 15 exhibits the overall accuracy of seven individual classifiers and Voting-Soft-AOA. Among the seven individual classifiers, GBDT has the highest accuracy of 83.1%, followed by XGBoost, SVM, MLP, KNN, RF, and Bayesian in turn. The Voting-Soft-AOA model has the highest overall accuracy of 87.5%, 4.4% higher than that of the ensemble learning algorithm GBDT. It suggests that Voting-Soft-AOA is superior in rockburst intensity grade prediction.

Figure 16 presents ROC curves and AUC values of seven individual classifiers and Voting-Soft-AOA. ROC curves of all the prediction models are on the upper left. In general, the ROC curve of Voting-Soft-AOA is the closest to the upper left corner, which proves the best predictive performance of Voting-Soft-AOA. Among the seven individual learners, GBDT has the highest AUC value of 0.948, while KNN has the lowest AUC value of 0.891. Voting-Soft-AOA has the highest AUC value of 0.952, 0.004 higher than that of GBDT. The results demonstrate that Voting-Soft-AOA achieves the best performance in predicting rockburst intensity grades.

4.5.2 Performance Comparison Between Voting-Soft-AOA and Other Ensemble Algorithms

To compare the predictive performance of Voting-Soft-AOA with other ensemble algorithms, Voting-Hard, Voting-Soft, Stacking, Bagging SVM (BagSVM), and Bagging KNN (BagKNN) were selected as comparative models. Figure 17 displays the overall accuracy of different ensemble algorithms on the test set. Voting-Soft-AOA has the highest overall accuracy, followed by Stacking, and BagKNN performs the worst. In addition, Voting-Soft performs better than Voting-Hard, which demonstrates that giving different
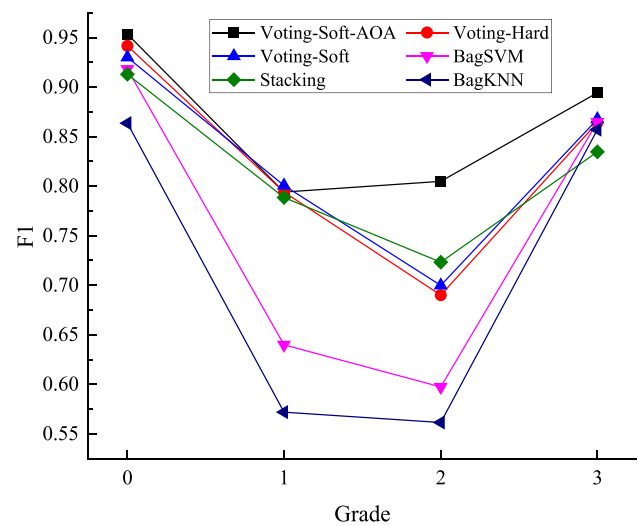
weights to different individual learners can improve the predictive performance of Voting. After hyper-parameter tuning for Voting-Soft with the aid of AOA, the accuracy of Voting-Soft model is 0.875, better than that of other ensemble learning models.

In the hope of further testing the predictive performance of different ensemble algorithms, F1-score was regarded as an evaluation index here. The F1-score values of different ensemble algorithms on the test set are illustrated in Fig. 18. Voting-Soft-AOA is superior to other ensemble models, while BagKNN is inferior to other ensemble models in terms of predictive capability at all the rockburst intensity grades.

4.6 Variable Importance

To calculate the relative importance of rockburst characteristic variables, Voting-Soft-AOA was taken as the objective function, and sensitivity analysis was performed on characteristic variables by the random balance design Fourier amplitude sensitivity test (RBD-FAST). RBD-FAST is a method that achieves the latest development in FAST by RBD, so as to reduce computational costs (Mara 2009). All parameters are set to the same frequency, and they are randomly recombined after sampling. Then, Fourier decomposition is performed with fast Fourier transform (FFT) on the model output based on the order of the previous recombination to obtain the first-order sensitivity analysis results of parameters (Gao et al. 2020).

In RBD-FAST, changes in the results are decomposed into:
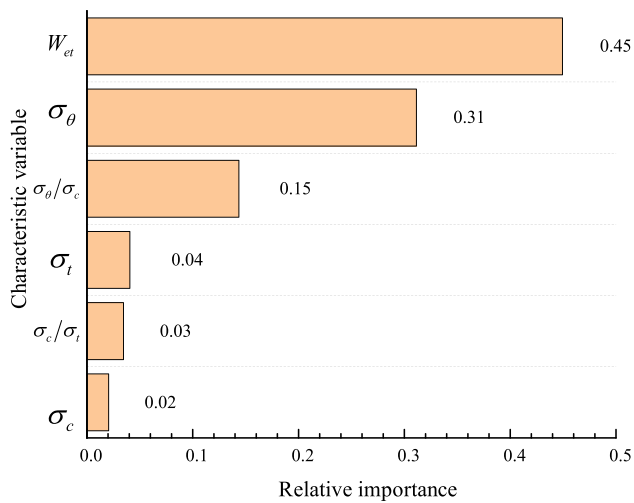
$$S_i = V_{x_i}/V(Y), \tag{14}$$

Fig. 19 Relative importance of characteristic variables

where V_{x_i} is the variance-based first-order influence of input factor x_i; and $V(Y)$ is the total variance output by Voting-Soft-AOA.

The relative importance of each input variable was calculated (Fig. 19). It can be seen from Fig. 19 that W_{et} is the most important input variable with a relative importance score of 0.45, followed by σ_θ (0.31), σ_θ / σ_c (0.15), σ_t (0.04), σ_c / σ_t (0.03), and σ_c (0.02) in turn.

The calculation results show that W_{et} is the most important factor affecting the intensity grade of rockburst. As one of the most commonly used evaluation indicators for the intensity grade of rockburst, W_{et} is often used in research on rockburst empirical criteria. The larger its value is, the more energy is released during rockburst. Thus, it can effectively reflect the occurrence and intensity of rockburst. In addition, the calculation results in this study are consistent with the research results of many scholars (Sun et al. 2021; Xue et al. 2022; Zhang et al. 2020). It is noteworthy that although the calculation results show that σ_c has the lowest relative importance, it does not mean that σ_c is unimportant, because the results were observed by comparing all the influencing factors together. Meanwhile, the calculation results in this study are also different from those of some other scholars

(Guo et al. 2022; Li et al. 2022b), mainly for the following reasons: (1) Different datasets can result in different degrees of variation, extremum values, input parameters, and rockburst grades of each variable, all of which can lead to different final calculation results. Besides, different data preprocessing methods may also yield different calculation results. (2) Different prediction models may lead to different nonlinear relationships between input and output variables, thereby producing different calculation results. Therefore, in future work, the authors will collect more samples, construct larger databases, and establish models with stronger generalization ability to make the calculation results more accurate.

A larger W_{et} means that more energy is stored in the surrounding rock, and thus the risk of rockburst is higher. Scholars have put forward various measures (He et al. 2020; Zhang et al. 2023, 2019; Zhao et al. 2016) to reduce the impact of W_{et} on rockburst intensity grades: (1) Roof pressure relief technology. It destroys those rock strata with large energy storage ahead of time by virtue of technologies including blasting, hydraulic fracturing, and surface fracturing, thus making rock strata less intact and releasing the stored energy. In this way, the value of W_{et} can be decreased to make engineering less prone to rockburst. (2) Floor pressure relief technology. It damages floor structures, and thus releases stored elastic energy in a timely manner by methods including deep hole floor-break blasting and floor grooving.

5 Case Application

For the purpose of testing the predictive performance of Voting-Soft-AOA in practical engineering, on-site data were collected from five different tunnels and mining projects, and six parameters obtained on-site were taken as input into this model to predict the on-site rockburst intensity grade. Besides, the prediction results were compared with those of the empirical prediction method based on the Russenes criterion (Russenes 1974). The prediction results revealed in Table 8 show that predictions of Voting-Soft-AOA are in line with actual situations of all these projects. Meanwhile, the

Table 8 Engineering application of the proposed Voting-Soft-AOA model

Group	σ_θ/MPa	σ_c/MPa	σ_t/MPa	σ_c / σ_t	σ_θ / σ_c	W_{et}	Actual	Voting-Soft-AOA predicted	Russenes criterion predicted
1 (Zhang et al. 2019)	35	133.4	9.3	14.34	0.26	2.89	1	1	0
2 (Tang and Xu 2020)	87.31	137.7	9.62	14.31	0.63	7.14	3	3	3
3 (Tang and Xu 2020)	87.31	94.4	9.16	10.31	0.92	3.57	1	1	3
4 (Wang et al. 2013)	121	87.5	8.73	13.9	0.72	9.05	3	3	3
5 (Wang et al. 2013)	79.1	124	8.64	14.4	0.64	7.74	3	3	3

overall prediction accuracy of Voting-Soft-AOA is superior to that of the Russell criterion. This proves that the model has great generalization ability and thereby can be applied in practical engineering. Moreover, these new rockburst data can enrich the rockburst database to improve the predictive ability of models.

6 Conclusions

A Voting-Soft-AOA ML model for rockburst data preprocessing and rockburst intensity grade prediction was proposed in this study. Besides, multiple data preprocessing methods were compared to verify the superiority of DBSCAN and Meancure-SMOTE in data prediction, as well as the accuracy of Voting-Soft-AOA in rockburst intensity grade prediction. Conclusions are summarized as follows:

1. The data were preprocessed by eliminating Outliers in the rockburst database through DBSCAN and then making the dataset balanced through Meancure-SMOTE. The predictive abilities of seven prediction models on different datasets were compared, and the distribution of these datasets in three-dimensional space was observed. It is drawn from the results that methods proposed in this study show better predictive performance than Kmeans-SMOTE, and SMOTE.
2. Hyper-parameters and voting weights for base classifiers in Voting were determined by means of AOA and fivefold CV. In addition, the optimal combination of base classifiers in Voting-Soft-AOA was determined by the combination weighting method of game theory and the exhaustive search method.
3. Voting-Soft-AOA outperforms individual learners and other ensemble models in terms of prediction at all the four rockburst intensity grade and overall prediction.
4. Sensitivity study was conducted on input variables with RBD-FAST, and the results suggest that W_{et} is the most important input variable, with a relative importance score of 44.94%. Hence, emphasis should be placed on W_{et} in practical underground engineering to prevent rockburst.
5. The application of Voting-Soft-AOA to practical engineering proves that it can provide reference for rockburst warning in actual underground engineering.

Declarations

References

Abualigah L, Diabat A, Mirjalili S, Abd Elaziz M, Gandomi AH (2021) The arithmetic optimization algorithm. Comput Meth Appl Mech Eng 376:38. https://doi.org/10.1016/j.cma.2020.113609

Arafa A, El-Fishawy N, Badawy M, Radad M (2022) RN-SMOTE: reduced noise SMOTE based on DBSCAN for enhancing imbalanced data classification. J King Saud Univ-Comput Inf Sci 34(8):5059–5074. https://doi.org/10.1016/j.jksuci.2022.06.005

Askaripour M, Saeidi A, Rouleau A, Mercier-Langevin P (2022) Rockburst in underground excavations: a review of mechanism, classification, and prediction methods. Undergr Space 7(4):577–607. https://doi.org/10.1016/j.undsp.2021.11.008

Asniar NU, Maulidevi NU, Surendro K (2022) SMOTE-LOF for noise identification in imbalanced data classification. J King Saud Univ-Comput Inf Sci 34(6):3413–3423. https://doi.org/10.1016/j.jksuci.2021.01.014

Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 16(1):321–357

Chen PF, Huang HB, Shi WZ (2022) Reference-free method for investigating classification uncertainty in large-scale land cover datasets. Int J Appl Earth Obs Geoinf 107:10. https://doi.org/10.1016/j.jag.2021.102673

Dong LJ, Li XB, Peng K (2013) Prediction of rockburst classification using random forest. Trans Nonferrous Met Soc China 23(2):472–477. https://doi.org/10.1016/s1003-6326(13)62487-5

Douzas G, Bacao F, Last F (2018) Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. Inf Sci 465:1–20. https://doi.org/10.1016/j.ins.2018.06.056

Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the second international conference on knowledge discovery and data mining. AAAI Press, Portland, pp 226–231

Feng X, Zhang HZ, Li LL, Zhang K, Wang TL (2019) The application of expectation and standard deviation calculations in the evaluation of dissolved arsenic in the Pu River, Liaoning Province, Northeastern China. Bull Environ Contam Toxicol 102(1):84–91. https://doi.org/10.1007/s00128-018-2503-5

Feng S, Keung J, Yu X, Xiao Y, Zhang M (2021) Investigation on the stability of SMOTE-based oversampling techniques in software defect prediction. Inf Softw Technol 139:14. https://doi.org/10.1016/j.infsof.2021.106662

Fernandez A, Garcia S, Herrera F, Chawla NV (2018) SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. J Artif Intell Res 61:863–905. https://doi.org/10.1613/jair.1.11192

Gao B, Yang Q, Peng ZJ, Xie WH, Jin H, Meng SH (2020) A direct random sampling method for the Fourier amplitude sensitivity test of nonuniformly distributed uncertainty inputs and its application

in C/C nozzles. Aerosp Sci Technol 100:8. https://doi.org/10.1016/j.ast.2020.105830

Ghasemi E, Gholizadeh H, Adoko AC (2020) Evaluation of rockburst occurrence and intensity in underground structures using decision tree approach. Eng Comput 36(1):213–225. https://doi.org/10.1007/s00366-018-00695-9

Gong FQ, Yan JY, Li XB, Luo S (2019) A peak-strength strain energy storage index for rock burst proneness of rock materials. Int J Rock Mech Min Sci 117:76–89. https://doi.org/10.1016/j.ijrmms.2019.03.020

Gong FQ, Wang YL, Luo S (2020) Rockburst proneness criteria for rock materials: Review and new insights. J Cent South Univ 27(10):2793–2821. https://doi.org/10.1007/s11771-020-4511-y

Guo DP, Chen HM, Tang LB, Chen ZX, Samui P (2022) Assessment of rockburst risk using multivariate adaptive regression splines and deep forest model. Acta Geotech 17(4):1183–1205. https://doi.org/10.1007/s11440-021-01299-2

Hao SX, Zhou XF, Song H (2015) A new method for noise data detection based on DBSCAN and SVDD. In: IEEE international conference on cyber technology in automation, control, and intelligent systems (CYBER). IEEE, Shenyang, pp 784–789

He SQ, Song DZ, Li ZL, He XQ, Chen JQ, Zhong TP, Lou Q (2020) Mechanism and prevention of rockburst in steeply inclined and extremely thick coal seams for fully mechanized top-coal caving mining and under gob filling conditions. Energies 13(6):26. https://doi.org/10.3390/en13061362

Hu L, Feng XT, Yao ZB, Zhang W, Niu WJ, Bi X, Feng GL, Xiao YX (2023a) Rockburst time warning method with blasting cycle as the unit based on microseismic information time series: a case study. Bull Eng Geol Environ 82(4):24. https://doi.org/10.1007/s10064-023-03141-3

Hu Q, Yuan Z, Qin KY, Zhang J (2023b) A novel outlier detection approach based on formal concept analysis. Knowl-Based Syst 268:13. https://doi.org/10.1016/j.knosys.2023.110486

Ji B, Xie F, Wang XP, He SQ, Song DZ (2020) Investigate contribution of multi-microseismic data to rockburst risk prediction using support vector machine with genetic algorithm. IEEE Access 8:58817–58828. https://doi.org/10.1109/access.2020.2982366

Kadkhodaei MH, Ghasemi E (2022) Development of a semi-quantitative framework to assess rockburst risk using risk matrix and logistic model tree. Geotech Geol Eng 40(7):3669–3685. https://doi.org/10.1007/s10706-022-02122-9

Kadkhodaei MH, Ghasemi E, Sari M (2022) Stochastic assessment of rockburst potential in underground spaces using Monte Carlo simulation. Environ Earth Sci 81(18):15. https://doi.org/10.1007/s12665-022-10561-z

Kidybiński A (1981) Bursting liability indices of coal. Int J Rock Mech Min Sci Geomech Abstr 18(4):295–304. https://doi.org/10.1016/0148-9062(81)91194-3

Leveille P, Sepehri M, Apel DB (2017) Rockbursting potential of kimberlite: a case study of Diavik diamond mine. Rock Mech Rock Eng 50(12):3223–3231. https://doi.org/10.1007/s00603-017-1294-z

Li TZ, Li YX, Yang XL (2017) Rock burst prediction based on genetic algorithms and extreme learning machine. J Cent South Univ 24(9):2105–2113. https://doi.org/10.1007/s11771-017-3619-1

Li ZR, Wang YJ, Zhao GH, Cheng L, Ma XK (2018) FROD: Fast and robust distance-based outlier detection with active-inliers-patterns in data streams. In: 27th international conference on artificial neural networks (ICANN). Springer, Rhodes, pp 626–636

Li DY, Liu ZD, Armaghani DJ, Xiao P, Zhou J (2022a) Novel ensemble intelligence methodologies for rockburst assessment in complex and variable environments. Sci Rep 12(1):23. https://doi.org/10.1038/s41598-022-05594-0

Li DY, Liu ZD, Armaghani DJ, Xiao P, Zhou J (2022b) Novel ensemble tree solution for rockburst prediction using deep forest. Mathematics 10(5):23. https://doi.org/10.3390/math10050787

Li GK, Xue YG, Qu CQ, Qiu DH, Wang P, Liu QS (2023) Intelligent prediction of rockburst in tunnels based on back propagation neural network integrated beetle antennae search algorithm. Environ Sci Pollut Res 30(12):33960–33973. https://doi.org/10.1007/s11356-022-24420-8

Liang WZ, Sari YA, Zhao GY, McKinnon SD, Wu H (2021) Probability estimates of short-term rockburst risk with ensemble classifiers. Rock Mech Rock Eng 54(4):1799–1814. https://doi.org/10.1007/s00603-021-02369-3

Lin Y, Zhou KP, Li JL (2018) Application of cloud model in rock burst prediction and performance comparison with three machine learnings algorithms. IEEE Access 6:30958–30968. https://doi.org/10.1109/access.2018.2839754

Liu YR, Hu SK (2019) Rockburst prediction based on particle swarm optimization and machine learning algorithm. In: 3rd international conference on information technology in geo-engineering (ICITG). Guimaraes, pp 292–303

Lu A, Yan P, Lu WB, Chen M, Wang GH, Luo S, Liu X (2021) Numerical simulation on energy concentration and release process of strain rockburst. KSCE J Civ Eng 25(10):3835–3842. https://doi.org/10.1007/s12205-021-2037-y

Mara TA (2009) Extension of the RBD-FAST method to the computation of global sensitivity indices. Reliab Eng Syst Saf 94(8):1274–1281. https://doi.org/10.1016/j.ress.2009.01.012

Mark C (2016) Coal bursts in the deep longwall mines of the United States. Int J Coal Sci Technol 3(1):1–9. https://doi.org/10.1007/s40789-016-0102-9

Motsinger AA, Ritchie MD (2006) The effect of reduction in cross-validation intervals on the performance of multifactor dimensionality reduction. Genet Epidemiol 30(6):546–555. https://doi.org/10.1002/gepi.20166

Nnamoko N, Korkontzelos I (2020) Efficient treatment of outliers and class imbalance for diabetes prediction. Artif Intell Med 104:12. https://doi.org/10.1016/j.artmed.2020.101815

Rojarath A, Songpan W (2021) Cost-sensitive probability for weighted voting in an ensemble model for multi-class classification problems. Appl Intell 51(7):4908–4932. https://doi.org/10.1007/s10489-020-02106-3

Russenes BF (1974) Analysis of rock spalling for tunnels in steep valley sides (in Norwegian). Dissertation/Thesis, Norwegian Institute of Technology

Sepehri M, Apel DB, Adeeb S, Leveille P, Hall RA (2020) Evaluation of mining-induced energy and rockburst prediction at a diamond mine in Canada using a full 3D elastoplastic finite element model. Eng Geol 266:17. https://doi.org/10.1016/j.enggeo.2019.105457

Shukla R, Khandelwal M, Kankar PK (2021) Prediction and assessment of rock burst using various meta-heuristic approaches. Mining Metall Explor 38(3):1375–1381. https://doi.org/10.1007/s42461-021-00415-w

Simser BP (2019) Rockburst management in Canadian hard rock mines. J Rock Mech Geotech Eng 11(5):1036–1043. https://doi.org/10.1016/j.jrmge.2019.07.005

Starczewski A, Goetzen P, Er MJ (2020) A new method for automatic determining of the DBSCAN parameters. J Artif Intell Soft Comput Res 10(3):209–221. https://doi.org/10.2478/jaiscr-2020-0014

Sun YT, Li GC, Zhang JF, Huang JD (2021) Rockburst intensity evaluation by a novel systematic and evolved approach: machine learning booster and application. Bull Eng Geol Environ 80(11):8385–8395. https://doi.org/10.1007/s10064-021-02460-7

Wenkan TAN, Nanyan HU, Yicheng YE, Menglong WU, Zhaoyun HUANG, Xianhua WANG (2022) Rockburst intensity

classification prediction based on four ensemble learning. Chin J Rock Mech Eng 41(S02):10

Tan W, Ye Y, Hu N, Wu M, Huang Z (2021) Severe rock burst prediction based on the combination of LOF and improved SMOTE algorithm. Chin J Rock Mech Eng 40(6):9

Tang L, Wang W (2002) New rock burst proneness index. Chin J Rock Mech Eng 21(6):874–878

Tang Z, Xu Q (2020) Rockburst prediction based on nine machine learning algorithms. Chin J Rock Mech Eng 39(4):9

Trajdos P, Kurzynski M (2018) Weighting scheme for a pairwise multi-label classifier based on the fuzzy confusion matrix. Pattern Recognit Lett 103:60–67. https://doi.org/10.1016/j.patrec.2018.01.012

Ullah B, Kamran M, Rui YC (2022) Predictive modeling of short-term rockburst for the stability of subsurface structures using machine learning approaches: t-SNE, K-means clustering and XGBoost. Mathematics 10(3):20. https://doi.org/10.3390/math10030449

Wang Y, Xu Q, Chai H, Liu L, Xia Y, Wang X (2013) Rock burst prediction in deep shaft based on RBF-AR model. J of Jilin Univ (earth Sci Ed) 43(6):1943–1949

Wang KK, Liu XD, Zhao JM, Gao HW, Zhang Z (2020) Application research of ensemble learning frameworks. In: Chinese automation congress (CAC). IEEE, Shanghai, pp 5767–5772

Wang JC, Ma HJ, Yan XH (2023) Rockburst intensity classification prediction based on multi-model ensemble learning algorithms. Mathematics 11(4):29. https://doi.org/10.3390/math11040838

Wong TT, Yeh PY (2020) Reliable accuracy estimates from k-fold cross validation. IEEE Trans Knowl Data Eng 32(8):1586–1594. https://doi.org/10.1109/tkde.2019.2912815

Xue YG, Li ZQ, Li SC, Qiu DH, Tao YF, Wang L, Yang WM, Zhang K (2019) Prediction of rock burst in underground caverns based on rough set and extensible comprehensive evaluation. Bull Eng Geol Environ 78(1):417–429. https://doi.org/10.1007/s10064-017-1117-1

Xu G, Li KG, Li ML, Qin QC, Yue R (2022) Rockburst intensity level prediction method based on FA-SSA-PNN model. Energies 15(14):19. https://doi.org/10.3390/en15145016

Xue YG, Bai CH, Kong FM, Qiu DH, Li LP, Su MX, Zhao Y (2020a) A two-step comprehensive evaluation model for rockburst prediction based on multiple empirical criteria. Eng Geol 268:11. https://doi.org/10.1016/j.enggeo.2020.105515

Xue YG, Bai CH, Qiu DH, Kong FM, Li ZQ (2020b) Predicting rockburst with database using particle swarm optimization and extreme learning machine. Tunn Undergr Space Technol 98:12. https://doi.org/10.1016/j.tust.2020.103287

Xue RX, Liang ZZ, Xu NW (2021) Rockburst prediction and analysis of activity characteristics within surrounding rock based on microseismic monitoring and numerical simulation. Int J Rock Mech Min Sci 142:12. https://doi.org/10.1016/j.ijrmms.2021.104750

Xue YG, Li GK, Li ZQ, Wang P, Gong HM, Kong FM (2022) Intelligent prediction of rockburst based on Copula-MC oversampling architecture. Bull Eng Geol Environ 81(5):14. https://doi.org/10.1007/s10064-022-02659-2

Yang X, Pei Y, Cheng H, Hou X, Lv J (2021) Prediction method of rockburst intensity grade based on SOFM neural network model. Chin J Rock Mech Eng 40(S01):8

Yang FJ, Hui Z, Xiao HB, Azhar MU, Yong Z, Chi FD (2022) Numerical simulation method for the process of rockburst. Eng Geol 306:16. https://doi.org/10.1016/j.enggeo.2022.106760

Yi XK, Xu YY, Hu Q, Krishnamoorthy S, Li W, Tang ZZ (2022) ASN-SMOTE: a synthetic minority oversampling method with adaptive qualified synthesizer selection. Complex Intell Syst 8(3):2247–2272. https://doi.org/10.1007/s40747-021-00638-w

Yu SY, Ren XH, Zhang JX, Sun ZH (2023) Numerical simulation on the excavation damage of Jinping deep tunnels based on the SPH method. Geomech Geophys Geo-Energy Geo-Resour 9(1):18. https://doi.org/10.1007/s40948-023-00545-z

Zeng C, Wang RD, Zuo QH (2022) Analysis of abnormal flight and controllers data based on DBSCAN method. Secur Commun Netw 2022:8. https://doi.org/10.1155/2022/7474270

Zhang H, Chen L, Chen SG, Sun JC, Yang JS (2018) The spatiotemporal distribution law of microseismic events and rockburst characteristics of the deeply buried tunnel group. Energies 11(12):21. https://doi.org/10.3390/en11123257

Zhang SC, Li YY, Shen BT, Sun XZ, Gao LQ (2019) Effective evaluation of pressure relief drilling for reducing rock bursts and its application in underground coal mines. Int J Rock Mech Min Sci 114:7–16. https://doi.org/10.1016/j.ijrmms.2018.12.010

Zhang JF, Wang YH, Sun YT, Li GC (2020) Strength of ensemble learning in multiclass classification of rockburst intensity. Int J Numer Anal Methods Geomech 44(13):1833–1853. https://doi.org/10.1002/nag.3111

Zhang AM, Yu HL, Huan ZJ, Yang XB, Zheng S, Gao S (2022a) SMOTE-RkNN: a hybrid re-sampling method based on SMOTE and reverse k-nearest neighbors. Inf Sci 595:70–88. https://doi.org/10.1016/j.ins.2022.02.038

Zhang Y, Yan JL, Qiao L, Gao HB (2022b) A novel approach of data race detection based on CNN-BiLSTM hybrid neural network. Neural Comput Appl 34(18):15441–15455. https://doi.org/10.1007/s00521-022-07248-8

Zhang J, Liu JN, Wang YJ, Yang G, Hou SL, Wang YJ, He MC, Yang J (2023) Study on pressure relief mechanism of hydraulic support in working face under directional roof crack. Arch Min Sci 68(1):103–123. https://doi.org/10.24425/ams.2023.144320

Zhao HB, Chen BR (2020) Data-driven model for rockburst prediction. Math Probl Eng 2020:14. https://doi.org/10.1155/2020/5735496

Zhao SK, Deng ZG, Qi QX, Li HY (2016) Theory and application of deep hole floor-break blasting in floor rock burst coal mine. In: ISRM international symoposium on rock mechanics and rock engineering—from the past to the future. CRC Press-Balkema, pp 511–515

Zhou J, Li XB, Shi XZ (2012) Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. Saf Sci 50(4):629–644. https://doi.org/10.1016/j.ssci.2011.08.065

Zhou K, Lei T, Hu J (2013) RS-TOPSIS model of rockburst prediction in deep metal mines and its application. Chin J Rock Mech Eng S2:7

Zhou J, Li XB, Mitri HS (2016) Classification of rockburst in underground projects: comparison of ten supervised learning methods. J Comput Civil Eng 30(5):19. https://doi.org/10.1061/(asce)cp.1943-5487.0000553

Zhu WB, Webb ZT, Mao KT, Romagnoli J (2019) A deep learning approach for process data visualization using t-distributed stochastic neighbor embedding. Ind Eng Chem Res 58(22):9564–9575. https://doi.org/10.1021/acs.iecr.9b00975