RESEARCH ARTICLE

Construction of a Two-Stage Rockburst Warning Model Based on Multi-Source Rockburst Case Studies

JIN SHANG , QINGWANG LIAN , XINLIN CHEN , AND HAORU YANG

Key Laboratory of In-Situ Property-Improving Mining of Ministry of Education, School of Mining Engineering, Taiyuan University of Technology, Taiyuan 030024, China

Corresponding author: Qingwang Lian (lianqingwang@tyut.edu.cn)

ABSTRACT Rock burst is a sudden disaster that is influenced by various factors. Accurately identifying and predicting rock burst risks is of great significance for improving mine safety. In order to improve the accuracy of rock burst prediction from the perspective of data structure, this paper constructs a two-level prediction model based on six rock burst features. The first-level model uses Box-Cox, Yeo-Johnson, and uniform transformations for data scaling and extends the data using the CTGAN architecture, followed by feature dimension optimization. The second-level model uses the K-Means algorithm to reconstruct labels and enhance inter-class differences, and visualizes clustering effects using ISOMAP. At the algorithm optimization level, an ensemble model stacked with 8 algorithms and a deep forest are used for prediction. The results show that data transformation, increasing model complexity, and appropriate feature expansion can effectively improve prediction accuracy. The single model achieved a maximum accuracy of 81.25%, and the established two-level model outperformed a single machine learning method, with an accuracy improvement of 17.3%. Feature dimension optimization had the highest accuracy improvement of 6.3%. Through comparison, it was found that the deep forest has a prediction accuracy of 98.6%, which is superior to other models such as Gradient Boosting and Multilayer Perceptron. In addition, the SHAP value and 7 evaluation indicators were used to evaluate the model and further explain the prediction results. The proposed two-level rock burst prediction model provides a certain reference value for accurately predicting rock bursts.

INDEX TERMS Rock burst, two-level prediction model, Yeo-Johnson, Box-Cox, uniform transformation, CTGAN, K-means, ISOMAP, deep forest, SHAP value.

I. INTRODUCTION

Rockburst is a common sudden disaster, manifested by the spalling, ejection, and large deformation of the underground space structure and accompanied by damage noise, in the deep engineering excavation process, mostly occurring in high-stress areas [1], [2]. It is generally believed that the rock's high strength and the fact that it is harder, brittle, and more elastic are the main causes of the rock explosion phenomenon [3]. As the rock is disturbed, causing the original rock stress to be redistributed to form a stress concentration region, the elastic strain energy accumulated inside the rock is rapidly released, causing the deep space near the high-stress region to be dynamically unstable, resulting in a rock burst [4], [5], [6]. This has caused great damage to countries and regions including the USA, South Africa, Germany, Canada, and China [7], [8]. Rock explosions have become a major challenge threatening the safety of underground spaces [9], [10], [11], and advanced identification, risk prediction, and evaluation are major requirements to risk prediction and evaluation are major requirements to ensure the safety of personnel and equipment.

The associate editor coordinating the review of this manuscript and approving it for publication was Michele Magno .

Most of the traditional means of rockburst prediction are based on physical methods that analyze the mechanistic properties [12], [13] and many scholars have used the physical properties of rocks to characterize the rupture evolution of rocks to achieve control over rockburst. Previously, Zhao et al. [14] experimentally explored changes in the rockburst propensity index in a coal-rock complex and applied DHPB to structurally realize the energy transfer of a rockburst to reduce high-stress accumulation. Moreover, Wang et al. [15] experimentally explored the deformation characteristics of granite combined with the acoustic emission characteristic index to construct a prediction model to reveal the rock crack expansion law during a rockburst. Liu et al. [16] revealed the microscopic evolution of self-potential (SP) by studying the internal electrical point discharge characteristics of the rock damage process. Jiang et al. [17] used a microseismic detection system (MS) to study the mining stress rock fracture.Pattern and fracture zone distance of a strong impact on a coal seam when subjected to mining stress near the coal seam, which plays a role in coal mine waterproofing, as based on the MS visualization results.

However, due to the complex characterization indicators and numerous features associated with rock bursts, as well as the highly nonlinear relationships between potential triggering factors, traditional prediction methods based on mechanism properties are unable to provide reliable prediction accuracy [18]. Moreover, a general understanding of the mechanism behind rock burst occurrences has yet to be established.

In recent years, with the development of artificial intelligence, intelligent algorithms have facilitated solutions to rock blast problems [19], and machine learning and deep learning have been widely used in engineering application scenarios, such as signal recognition processing, anomaly detection, image processing, digital modeling, long- and short-term rock blast prediction, and hazard evaluation. Yang et al. [20] used the K-mean algorithm to identify and detect weak tectonic signals to characterize six types of non-tectonic source noise field signals. Wang et al. [21] utilized deep learning network frameworks to recognize and process precursor time series signals of rockburst, thus intelligently capturing hazard information for rockburst prevention. Li et al. [22] used fully convolutional neural networks to construct 3D mathematical models of CT (Computed Tomography) images and reconstructed metrics such as mean curvature and porosity to compare the accuracy of the constructed models from a microscopic perspective, and the advent of these methods has largely reduced the "empirical threshold" and made the decision-making process more scientific.

In terms of rock burst prediction, machine learning (as shown in Figure 1) has strong capabilities in dealing with nonlinear problems [23], [24], relying on data-driven methods to make autonomous decisions about classification boundaries [25]. Scholars have used machine learning algorithms to develop prediction methods in order to obtain
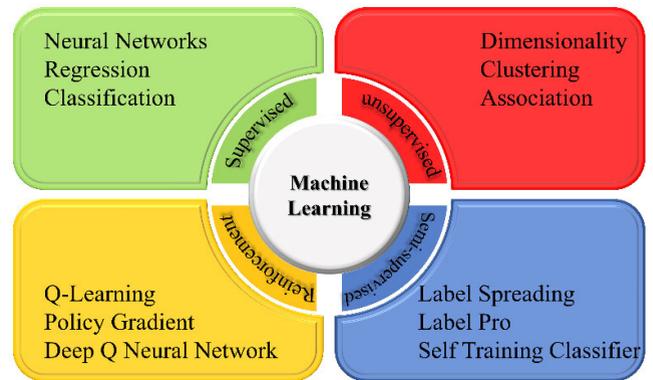

FIGURE 1. Machine learning foundation classification chart.

the best prediction models. Zhou et al. [26] compared ten single-model supervised learning methods and selected the best prediction model based on different metrics. In addition, the ensemble strategy of combining multiple basic algorithm models has been increasingly applied to rock burst prediction. Li et al. [27], [28] constructed an ensemble learning framework based on models such as random forests and gradient boosting, using three combination methods: voting, stacking, and bagging, all of which showed better prediction results than single models. They also applied the deep forest model, which combines the characteristics of deep learning and ensemble models, to rock burst prediction, achieving even more outstanding prediction capabilities. Although scholars have made certain achievements in using machine learning methods to study rock burst prediction, there are still many problems in this area:

1) Data level: Machine learning algorithms are highly dependent on data, but there are many challenges in collecting and organizing data for rockburst prediction in practical scenarios, such as difficulties in data acquisition, data missing, poor data quality, imbalanced data types, and limited precursor information parameters under different working conditions, leading to poor performance of machine learning algorithms in prediction.

2) Model level: The improvement of prediction performance depends on model complexity and hyperparameter selection. Machine learning algorithms often rely on experience and require a large amount of computing resources to optimize parameters. The prediction results brought by different algorithm combination strategies are difficult to control, which brings great challenges to the application of algorithms.

3) Algorithm level: Different decision-making logic of machine learning algorithms affects the accuracy, speed, stability, and interpretability of machine learning decisions. Different algorithms have different sensitivity to data, which may cause problems of accuracy reduction during training.

4) Interpretability level: Machine learning algorithms have black box characteristics and are difficult to interpret. The prediction process and results are difficult to visualize, so the discrimination mechanism inside the model is not clear

enough, affecting the credibility and reliability of the model application.

In summary, this paper builds upon previous research to investigate long-term prediction of rockburst. Through literature analysis, 300 sets of original rockburst data were collected, containing six sets of rockburst characteristics. To address the dependency of machine learning algorithms on data, which can affect the final prediction results, three non-linear transformation methods and the CTGAN [29] (Conditional Tabular GAN) architecture were selected to construct five data sets at the data level, addressing the issue of inter-class dispersion and the presence of outliers in the original data. In addition, the rockburst characteristics were extended at the feature level by constructing a secondary model to increase model complexity. The prediction model's accuracy, precision, F1 score, recall rate, and other metrics were evaluated. To address the issue of model interpretability, the SHAP (Shapley Additive Explanations) explainer was used to explain the model's decision logic. The findings of this study enrich the results of rockburst prediction and provide useful guidance for practical engineering applications.

II. ALGORITHMIC FRAMEWORK AND UNDERLYING THEORY

A. NONLINEAR TRANSFORMATIONS

The appropriate replacement or removal of outliers will enhance data consistency but, from the perspective of the structure of the data, will interfere with the objective distribution of the data, causing a loss of feature information. The appropriate introduction of noise can enhance model robustness [30]. In this study, three nonlinear mapping methods were used to improve the data structure by reasonably deflating the data, the original data information is shown in Table 2.

Since original data often have long-tailed features and cluttered stacking, distribution features generally lack significant features, reducing the performance of many machine learning predictions and slowing down the convergence of gradient-based learners, and the normal nature of the probability distribution helps to improve the performance of the classifiers [31].

The Yeo-Johnson algorithm [32], Box-Cox transformation, and quantile uniform distribution mapping are commonly used nonlinear transformation methods to map data to a Gaussian distribution, reduce the heteroscedasticity of data features, minimize bias, and improve the normality of data. Among them, the quantile uniform distribution mapping method (Equation 1) can smooth out anomalous distributions, correct skewness and kurtosis, and is more robust, which can reduce the impact of outliers on the results by mapping the data according to the cumulative probability density function. The Box-Cox transformation is a generalized power transformation method (Equation 2), which can be adjusted by the parameter lambda to improve the normality and homoscedasticity of the data. The Yeo-Johnson algorithm

(Equation 3) introduces an offset parameter based on the Box-Cox transformation, which can handle positive and negative numerical values and zero values. The lambda values in the latter two transformations can be determined by maximizing the likelihood function, making the transformed data more consistent with the assumption of normal distribution.

$$f(x) = \begin{cases} 0 & else \\ 1/(b-a) & if\, a \leq x \leq b \end{cases} \tag{1}$$

where a, b denote boundary value, $f(x)$ represents the probability density function of a uniform distribution.

$$y = \begin{cases} \{x^{\lambda} - 1\}/\lambda & if\, \lambda \neq 0 \\ \log(x) & if\, \lambda = 0 \end{cases} \tag{2}$$

where λ is the transformation parameter and x is the original dependent variable.

$$\psi(y_i, \lambda) = \begin{cases} \{(1+y_i)^{\lambda} - 1\}/\lambda & if\, \lambda \neq 0, y_i \geq 0 \\ \log(1+y_i) & if\, \lambda = 0, y_i \geq 0 \\ -\{(1-y_i)^{\lambda} - 1\}/\lambda & if\, \lambda \neq 0, y_i < 0 \\ -\log(1-y_i) & if\, \lambda = 0, y_i < 0 \end{cases} \tag{3}$$

where y_i is the feature data, λ is a tunable parameter.

B. CONDITIONAL TABULAR GAN

Conditional Tabular GAN (CTGAN) [33] is a type of generative adversarial network used to generate synthetic data with a similar structure to a given dataset. Its main principle is to use the conditional GAN approach to learn the underlying distribution of the original dataset and then generate synthetic data based on that distribution. The generator of CTGAN consists of two sub-networks, an embedding network to learn the distribution and patterns of the input data, and a generator network to transform the learned embedding into synthetic data. During the learning process, CTGAN introduces a conditioning mechanism to use conditional attributes to generate synthetic data, ensuring consistency between the generated data and the original data on the conditional attributes.

The goal of CTGAN is to minimize the GAN loss function, which includes the losses of the generator and discriminator. The generator loss measures the degree to which the generator misleads the discriminator, while the discriminator loss measures the degree of difference between the generated data and the original data. Through iterative training, the generator of CTGAN gradually learns the distribution and features of the data and generates synthetic data, while the discriminator gradually learns the features of the data and makes more accurate judgments of the generated data.

In addition, CTGAN uses the cross-entropy between the conditional vector and the new sample as the loss term and adds the PacGAN [34] (Progressive Attentional Context GAN)
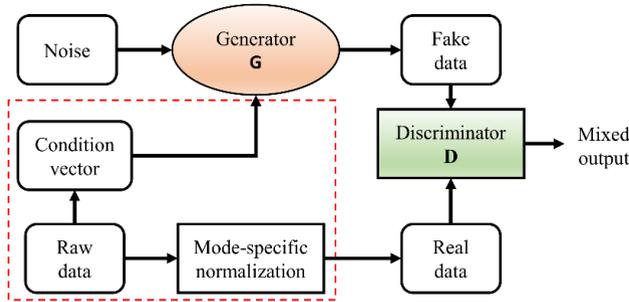
FIGURE 2. CTGAN schematic.

discriminator architecture to improve pattern stability and data quality. The CTGAN model is shown in Figure 2, and the key steps of the algorithm implementation are as follows:

1) The generator of CTGAN uses maximum likelihood estimation as the loss function, where the similarity between the samples generated by the generator and the original dataset is measured by the KL divergence. The formula for the loss function is:

$$L_G = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{M} \left[P_m(x_j^i) log \frac{P_m(x_j^i)}{P_d(x_j^i)} \right] \quad (4)$$

where N is the batch size, M is the number of features, $P_m(x_j^i)$ is the probability of the generator generating the sample x_j^i under the current condition, and $P_d(x_j^i)$ is the probability of x_j^i in the original dataset.

2) The discriminator's loss function uses binary cross-entropy, and the calculation formula for the loss function is:

$$L_D = -\frac{1}{N} \sum_{i=1}^{N} [logD(x_i) + log(1 - D(G(z_i)))] \quad (5)$$

where x_i is a real sample from the original data set, z_i is a random vector from the noise distribution, $G(z_i)$ is a sample generated by the generator using z_i as a condition, and $D(x_i)$ and $D(G(z_i))$ are the discrimination probabilities of the discriminator for x_i and $G(z_i)$, respectively.

3) CTGAN uses a neural network to generate the conditional vector, which takes the original dataset as input and outputs a vector used as a condition for both the generator and the discriminator.

C. ISOMETRIC MAPPING

ISOMAP [35] (Isometric Mapping) is a nonlinear manifold learning algorithm that finds its low-dimensional embedding for a given high-dimensional manifold so that the nearest-neighbor structure on the high-dimensional manifold is maintained in the low dimension, uses the minimum path approximation geodesic distance to characterize the data point spacing, and the essential dimensionality of the implied low-dimensional embedding is determined by the residual variance to ensure the stability and global optimality of the results, which is achieved as follows:

1) Construct a distance matrix for the points in the high-dimensional space, keeping the points at a distance of their k-neighborhood, and treating the other points as infinity.

2) Calculate the shortest path between point pairs, i.e., for all point pairs $i{\rightarrow}j$, find the shorter path $i{\rightarrow}k{\rightarrow}j$ for substitution, which is expressed as follows:

$$d_{ij} = d_{ik} + d_{kj} \quad (6)$$

where the above equation is satisfied when $d_{ij} > d_{ik} + d_{kj}$, k is the intermediate node.

3) The reconstructed distance matrix is combined with the MDS (Multidimensional Scaling) algorithm for dimensionality reduction, the correlation matrix B of the reduced vector Z is calculated, and eigenvalue decomposition is performed to obtain a larger number of eigenvalues and eigenvectors to obtain Z.

$$b_{ij} = -\frac{1}{2}(d_{ij}^2 - \frac{1}{m}\sum_{i=1}^{m}d_{ij}^2 - \frac{1}{m}\sum_{j=1}^{m}d_{ij}^2 + \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m}d_{ij}^2) \quad (7)$$

$$\mathbf{B} = U\Lambda U^T = \left(\Lambda^{\frac{1}{2}}U^T\right)^T \left(\Lambda^{\frac{1}{2}}U^T\right) = Z^T Z \quad (8)$$

where U denotes the eigenvector, Λ denotes the eigenvalue, b_{ij} denotes the element in the intercorrelation matrix B, and d_{ij} denotes the ISOMAP distance metric.

D. K-MEANS CLUSTERING ALGORITHM

Clustering is a process of dividing a dataset into different clusters according to specific criteria, making sure that the data points within the same cluster have high similarity and the clusters are separated by distinct boundaries. K-Means is a classic unsupervised clustering algorithm. When using the K-Means clustering algorithm, the first step is to determine the number of clusters. Then, K initial centroids are randomly chosen, and each data point is assigned to the cluster whose centroid is the closest. After that, the centroids of each cluster are recalculated, and the process is iterated until a certain number of iterations or when the assignment of the data points to clusters no longer changes. The steps are as follows:

1) Randomly select k centers and mark: $\mu_1^{(0)}, \mu_2^{(0)}, \ldots, \mu_k^{(0)}$, $k = 0, 1, 2, \ldots$

2) Define the loss function:

$$J(c, \mu) = min \sum_{i=1}^{M} \|x_i - \mu_{c_i}\|^2 \quad (9)$$

where x_i represents the ith sample, c_i represents the cluster to which x_i belongs, μ_{c_i} represents the center of mass, and M represents the total number of samples.

3) Assign the nearest center to x_i and iterate:

$$c_i^t < -argmin_k \|x_i - \mu_{c_i}\|^2 \quad (10)$$

4) Iterate over the class center k:

$$\mu_k^{(t+1)} < -argmin_\mu \sum_{c_i^t=k}^{b} \|x_i - \mu\|^2 \quad (11)$$
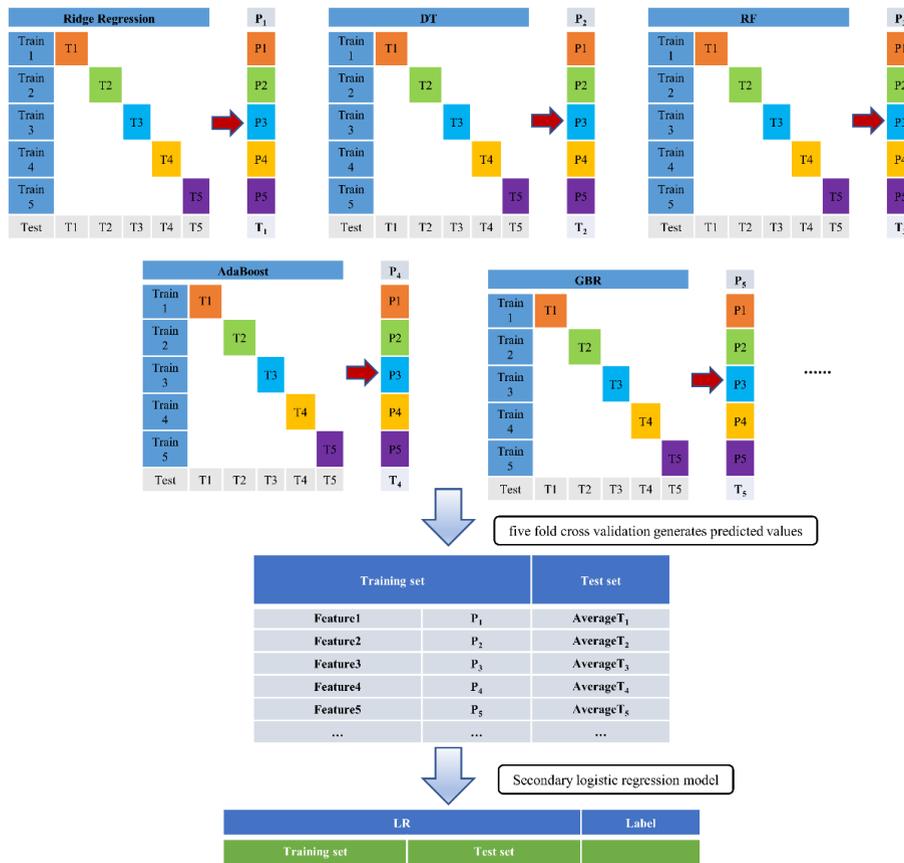
where $t = 0, 1, 2, \ldots$

FIGURE 3. Stacking schematic.

E. INTEGRATED STACKING

The upper prediction limit of a single model is often difficult to improve further, and the dependence of a single model on features affects the prediction accuracy. The base combiner and learning combiner methods are the two core methods of model fusion. The accuracy of the model can be further improved by adjusting the model weights and fusing multiple learner models for weighted optimization, such as through stacking, model blending, and cascade optimization.

Layer-by-layer processing, feature variation within the model, and model complexity determine the integrated learning effect. The stochastic nature of the integrated learning [36] framework effectively improves the training effect by rotating the feature space [37], and stacking [27] is an important method for improving prediction scores in integrated learning. The principle of stacking is shown in Figure 3. The steps are:

1) The primary trainer first splits the training set into folds, assembles multiple classifiers for feature extraction, selects a five-fold cross-validation for data normalization, sets the model parameter pool using a grid search method to output the best parameters for each base model, and outputs the sub-validation set.

2) Secondly, the output of the primary trainer is imported into the secondary regression model trainer, and the combined

sub-validation sets are averaged to output the hierarchical prediction results.

F. DEEP FOREST

Deep learning frameworks are mostly built on neural networks with multi-layer parameterized differentiable nonlinear modules, which outperform traditional machine learning in terms of classification and regression performance and applicability and have sufficient model complexity, usually leading to good prediction results. However, they rely on hyperparameter tuning and do not apply to small-scale datasets. The deep forest model [38], a non-neural network-based supervised integrated learning model with a cascade structure, can adaptively adjust the network structure for feature learning through the forest, reevaluate the K-fold cross-validation dataset at each layer, and automate the iterative training network layers to avoid the risk of overfitting, as shown in Figure 4. The steps are as follows:

1) To enhance model generalization, two CRF (Completely Random Forest) and RF (Random Forest) structures are selected for each layer to learn information about the input feature vectors.

2) The decision process involves feeding the samples into the deep forest, which obtains a d-dimensional class vector at the leaf nodes, representing the probability distribution of the samples belonging to different categories. Then, all the mean
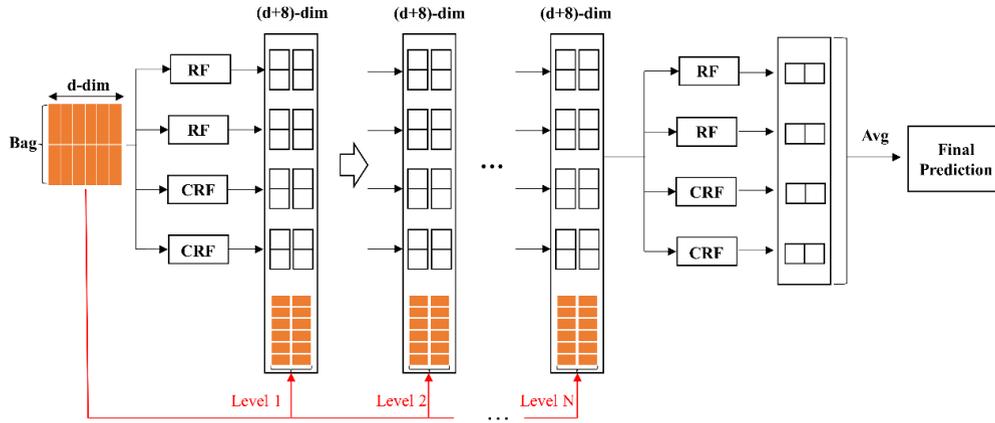
FIGURE 4. Deep forest schematic.

TABLE 1. Interpretation of the meaning of selected indicators.

Rockburst indicators	Indicates meaning	Rock blast classification guidelines			
		None(I)	Light(II)	Moderate(III)	Strong(IV)
$\sigma_{\theta}(MTS)$	the maximum tangential stress of the surrounding rock, MPa	0-24.0	24.0-60.0	60.0-126.0	126.0-200.0
$\sigma_{c}(UCS)$	the uniaxial compressive strength of the rock, MPa	0-80.0	80.0-120.0	120.0-180.0	180.0-320.0
σ_{t}	the uniaxial tensile strength of the rock, MPa	0-5.0	5.0-7.0	7.0-9.0	9.0-30.0
$W_{et}(EEI)$	rock elastic strain energy index	0-2.0	2.0-3.5	3.5-5.0	5.0-6.5
$SCF(SCF = \sigma_{\theta}/\sigma_{c})$ or SR	rock stress coefficient	0.1-0.3	0.3-0.5	0.5-0.7	0.7-0.9
$B_{1}(B_{1} = \sigma_{c}/\sigma_{t})$ or BR	rock brittleness coefficient	40.0-53.0	26.7-40.0	14.5-26.7	0-14.5

class vectors for each category are concatenated to form a 4d-dimensional vector, and all the vectors are concatenated to form a k × 4d-dimensional matrix. The category with the highest score in the k-dimensional vector is taken as the final prediction result for each sample.

G. DATA BINNING
Data binning can optimize the data structure, reduce the risk of model overfitting, speed up training, and be more robust with outliers, which is necessary for certain logistic regression models (e.g., LR, Logistic Regression, etc.). For the discrete features, similar features can be merged and ordered numerically through the binning operation. In this paper, we use an unsupervised approach to equifrequency discrete binning of features, calculate the quantile, and map the data to the quantile bins.

III. TWO-LEVEL PREDICTIVE MODEL CONSTRUCTION
A. ROCKBURST FEATURE EXTRACTION
The precursory process of rock explosion is often accompanied by changes in the mechanical parameters of the rock, and the accumulated elastic deformation energy of brittle rocks in the high stress zone is released instantaneously. In addition to this, on the basis of defining the material and structural damage of rocks under high stress, there is a division into

dynamic superposition, deep high static load loading, and deep high static load unloading rockbursts. Rock bursts are therefore closely related to the mechanical properties of the rock, the state of the energy deposit, and the properties of the rock itself.

In this paper, based on rockburst characteristics in relation to intensity, six rockburst indicators are selected based on the meta-literature example study method. Depending on the source of the data, the rockbursts are classified into four categories: None, Light, Medium, and Strong, according to the classification of rockburst intensity. The results of the rockburst feature selection and classification basis are shown in Table 1.

B. FIRST-LEVEL MODEL BUILDING
The first level model of rockburst data (Figure 5) is established by collecting real rockburst data through meta literature analysis method in the data preparation stage, followed by the Box-Cox transform, Yeo-Johnson power transform, uniform transform, and CTGAN framework in the data pre-processing stage. After that, the dataset combined with different nonlinear transformations and extensions is split according to a 30% test set and a 70% training set. The training process starts with KNN (K-Nearest Neighbor) and CatBoost (Categorical Boosting) for prediction using a grid
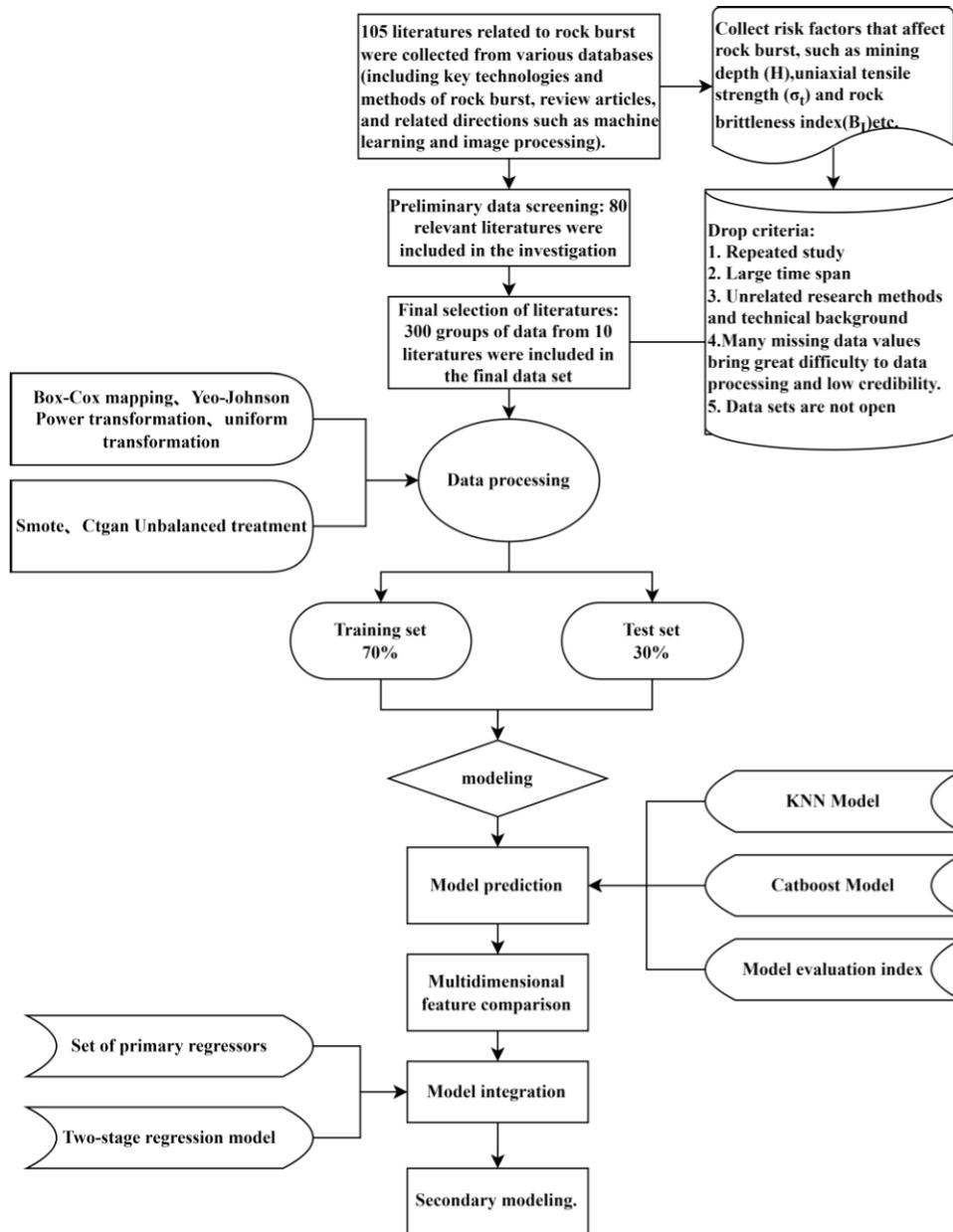
FIGURE 5. First-level model flowchart.

search approach and a 5-fold cross-validation strategy for parameter optimization and overfitting prevention, followed by optimization at the RockBurst feature level to find the optimal solution by increasing the number of feature dimensions. Finally, the two-stage stacking model is introduced.

C. SECONDARY MODEL BUILDING

The rockburst secondary modeling (Figure 6) process uses the K-Means algorithm to reduce intra-class variability in the Yeo-Johnson-CTGAN dataset, while the clustering results are visualized using ISOMAP and PCA (Principal Component Analysis) dimensionality reduction. The classification decision process of the K-Means algorithm is visualized through

the KElbowVisualizer tool. Finally, the prediction results of deep forest, integrated stacking, and a single model are compared.

D. MODEL EVALUATION STRATEGIES

To improve the accuracy of rockburst prediction, this study has established a two-level modeling process. For the first-level model, the focus is on data structure transformation. To highlight the transformation effect, five datasets were constructed, including the original dataset, Uniform Transformation-CTGAN, Box-Cox-CTGAN, Yeo-Johnson-CTGAN, and Yeo-Johnson-Smote. KNN and CatBoost base models were used for preliminary prediction in the
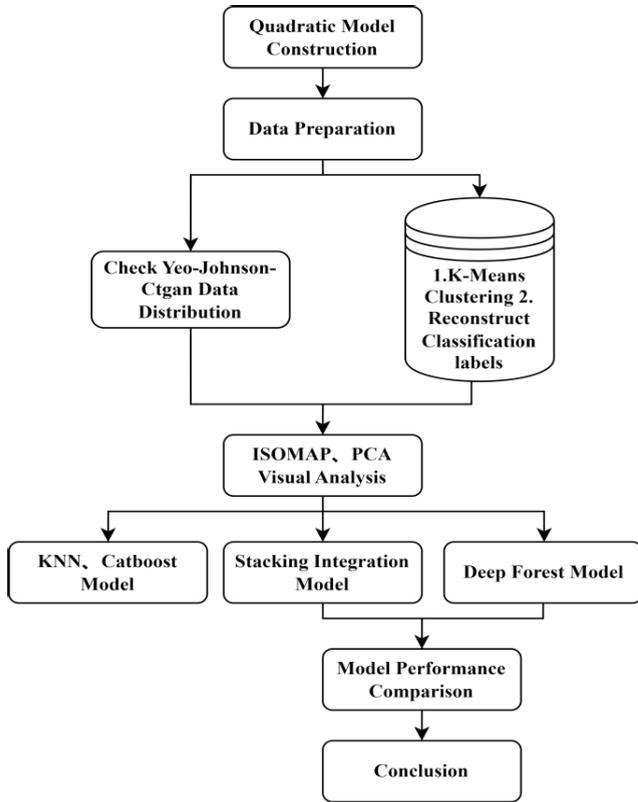
FIGURE 6. Secondary model construction flow chart.

TABLE 2. Original data information.

	σ_θ(MPa)	σ_c(MPa)	σ_t(MPa)	SCF	B_1	W_{et}
Max	297.80	304.20	22.60	4.87	80.00	30.00
Min	2.60	20.00	1.30	0.10	0.15	0.81
Mean	58.72	115.24	7.09	0.57	20.93	5.05
Std	50.65	44.89	4.15	0.61	13.22	3.85
Cov	0.86	0.39	0.58	1.08	0.63	0.76

Note: 'Std' tabulates the standard deviation of each column, 'Cov' indicates the coefficient of variation ($C_v = \sigma/\mu$, σ indicates the standard deviation, μ indicates the mean).

TABLE 3. Information on collected rockburst data.

Input Parameters	Data selection	Reference
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}$	129	Dong et al. [40]
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}, H$	85	Zhang et al. [41]
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}, \sigma_\theta/\sigma_c, \sigma_c/\sigma_t$	44	Zhou et al. [26]
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}, \sigma_\theta/\sigma_c, \sigma_c/\sigma_t$	15	Xue et al. [42]
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}, \sigma_\theta/\sigma_c, \sigma_c/\sigma_t$	52	Pu et al. [43]
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}, \sigma_\theta/\sigma_c, \sigma_c/\sigma_t$	29	Liu et al. [44]
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}, \sigma_\theta/\sigma_c, \sigma_c/\sigma_t, B_2$	37	Jia et al. [45]
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}, \sigma_\theta/\sigma_c, \sigma_c/\sigma_t, B_2$	45	Wu et al. [9]
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}, \sigma_\theta/\sigma_c, \sigma_c/\sigma_t, B_2$	41	Xue et al. [46]
$\sigma_\theta, \sigma_c, \sigma_t, W_{et}, \sigma_\theta/\sigma_c, \sigma_c/\sigma_t, H$	37	Wang et al. [47]

Note: $B_2(B_2 = (\sigma_c - \sigma_t)/(\sigma_c + \sigma_t))$ indicates rock brittleness index; H indicates depth.

verification and evaluation process, with Precision, Recall, FPR, F1 score, and overall accuracy selected as prediction evaluation indicators. Confusion matrix and ROC curve were analyzed. Then, the feasibility of expanding the model prediction effect at the feature level was explored through polynomial derivation, and the necessity of establishing a second-level prediction model was verified through feature decision analysis using shap graphs.

The second-level prediction model considered the actual engineering application scenario and mainly worked on model optimization. By reconstructing the classification labels of the data, the model complexity was increased and the prediction accuracy was improved. The effectiveness of the second-level model established in this study was verified through comparison with the base model and the integrated model.

IV. RESULTS AND DISCUSSION
A. DATA SOURCES AND PROCESSING
1) DATA COLLECTION
Rockburst data were collected using a literature analysis [39] approach, with WOS (Web of Science), CSCD (Chinese Science Citation Database), EI (Engineering Village), and other databases featuring, chosen as the data source, subject terms limited to rockburst prediction, machine learning, etc. for precise retrieval. A total of 105 relevant papers were retrieved, collecting both Chinese and foreign research results from different engineering backgrounds and evaluation indexes. Overall, 25 duplicate papers and 11 review-type articles were excluded as they did not contain the required terms, and 59 articles were excluded after a close reading of their abstracts, technical backgrounds, and methods used. Finally, 10 papers were selected as the final research objects in terms of data accessibility (see Figure 5 for the screening process). These data were derived from real engineering cases (hydropower station caverns, mines, and tunnels) and so the authenticity can be guaranteed.

A total of 514 case studies with corresponding data were collected from the literature (Table 3), while considering that some of the data would have had missing values [40], [41] if interpolation and partial elimination methods had been used. As this would have introduced bias to the original features of the data set, leading to the problem of order imbalance between regions, these incomplete data samples were eliminated in this paper. The final 300 sets of complete data were obtained. The maximum, minimum, mean, standard deviation, and coefficient of variation within the original data characteristics are shown in Table2.

Nonlinear transfer 300 raw rockburst classification records collected had six characteristics (Table 1), which were classified into four categories based on rockburst intensity: None, Light, Moderate, and Strong, with a ratio of 1.4:2.7:4.0:1.9, respectively (Figure 7). It can be seen that the data are in a skewed state, the I class samples and III class samples have large differences in the number of samples while the I class sample data feature less.
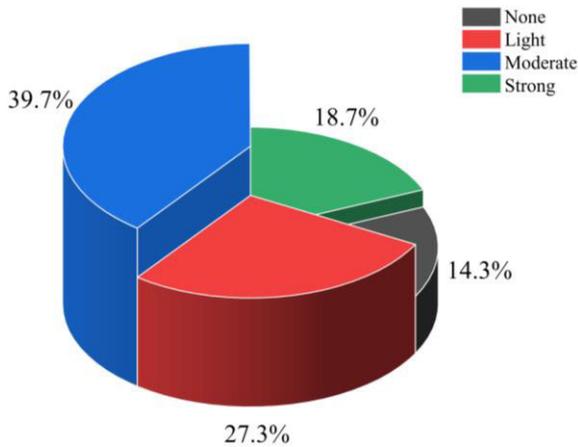
FIGURE 7. Raw rockburst data proportion of each grade.

The data distributions of the two original features (maximum and minimum cov) are illustrated in Figure 8(a), which shows that outliers are prevalent in the features and that the data distributions are characterized by skewed fronts and long-tailed distributions. In addition to this, by calculating the Pearson correlation coefficient (pcc) between the features, it can be seen that there is a strong correlation between SCF and σ_θ (pcc= 0.9), W_{et}, σ_θ (pcc= 0.47), σ_t, σ_c (pcc= 0.48) and B_1, σ_c (pcc= 0.63) had a strong correlation between them. It is therefore necessary to perform a non-linear transformation of the raw data to reduce the interference of outliers.

2) NONLINEAR TRANSFORMATIONS
The data set after the initial sieving does not contain missing values and does not require deletion and filling operations; each feature column has the same amount of data; there are no large outliers or duplicate data; and the data integrity is good. The rockburst classification data are numerically processed, and the rockburst intensities of None, Light, Moderate, and Strong correspond to the classification values (0, 1, 2, 3), respectively.

The process of data collection will result in outliers due to random errors, human errors, and variances in the original data. To fit the original dataset, one is used to divide the dataset into different groups, and another is used to build a low-rank approximation of the covariance matrix based on the induced variables [31]. Gou et al. [3] plotted box line plots to scale outliers within the median and interquartile range. Yin et al. [1] visualized outliers using data dimensionality reduction algorithms, such as the local outlier factor (LOF), and replaced the outliers with expectation maximization (EM).To highlight the effect of the transformation, the feature of σ_c (cov=0.39) with a smaller Coefficient of Variation (cov) and the larger feature of SCF (cov=1.08) were selected to construct the data distribution and compared with the above nonlinear transformation results, Figure 8 represents the comparison between the transformed results and the original data: (a), (b), (c), and (d)

are represented as the Box-Cox mapping of the original data, the Yeo-Johnson power transformation, the histogram of the frequency distribution of the uniformly distributed feature σ_c with SCF, the Q-Q plot, and the scatter plot of the distribution between the two features, respectively.

To quantify the distribution of the data after the transformation introduced Skewness (Skew), measure the asymmetry of the probability distribution of the random variables and the size of the value indicates the degree of skewness within the normal distribution. Kurtosis (BK) indicates the characteristic number of the peak height of the probability density distribution curve at the mean, a statistic describing the steepness of the distribution pattern of all the values taken. The data set with a high peak tends to have heavy-tailed distribution characteristics or outliers, and the formula is as follows:

$$Kurt = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^4$$

$$Skew = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^3 \quad (12)$$

where μ denotes the mean value and σ is the standard deviation.

A quantile-quantile plot (Q-Q plot) was used to visualize the normal distributivity of the features, with the horizontal axis indicating the actual data and the vertical axis indicating the quantile of the assumed normal data. Normality analysis of the data after several transformations, with the raw data in (a) Skew, the BK of σ_c and SCF are (170, 1.01), (0.53, 20.46), respectively. A very poor SCF normality was observed, and the outliers on the scatter plot were mainly concentrated on this feature.

In (a), the original data have long-tailed data distribution characteristics due to the presence of sparse data that may be subjectively considered outliers. The poor data peak of the SCF normality state skewed the normal distribution.

From (b) and (c), it can be seen that the normality is substantially improved after the Box-Cox and Yeo-Johnson transformations, and Skew and BK are substantially reduced. When comparing the results of the effects of the two treatments on the data, they are not very different.

(d) After uniform transformation, the data are uniformly distributed in the interval [0, 1]. The σ_c and Skew of SCF are slightly reduced compared to the above two transformations, but BK is slightly increased, and the kurtosis value of both features after uniform transformation is - 1.2. The kurtosis is flat and the same.

The 0-1 distribution (Equation 2) distorts the correlation and distance between and within the features, while the power transformation method arbitrarily maps distributions to locations close to the Gaussian distribution for different data sets. Examining the scatter plot shows that all three transformations narrow the gap between the outliers and the aggregated region and reduce the interference of the outliers, and the original Skew changes significantly after the three
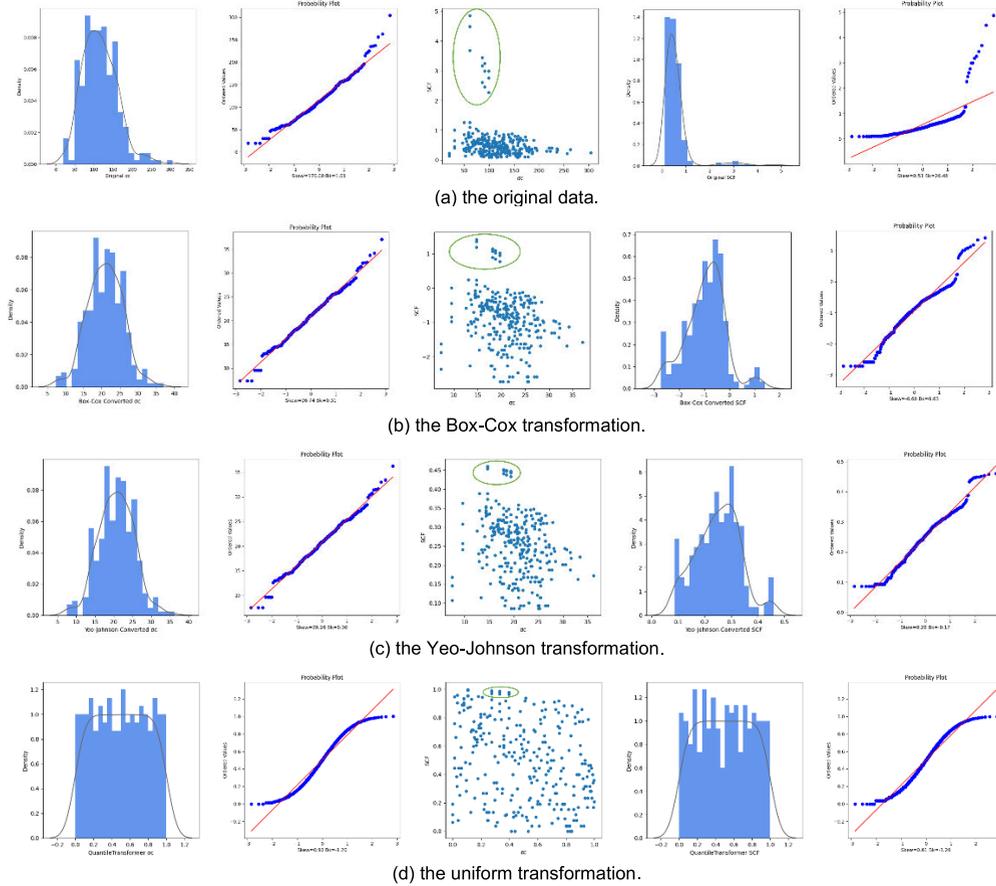
(a) the original data.

(b) the Box-Cox transformation.

(c) the Yeo-Johnson transformation.

(d) the uniform transformation.

FIGURE 8. Non-linear transformed data distribution comparison chart.

transformations. The variability with the normal distribution is greatly reduced, and the Q-Q plot reveals that the three transformations have the greatest impact on the SCF, and the normality of the characteristic distribution is corrected.

3) BIAS DATA PROCESSING

False accuracy is caused when the data set is skewed toward a certain classification result, and it is difficult to determine the distribution of the minority class data due to the limited data information contained in the minority classification sample, resulting in a low recognition rate [1]. There are two main approaches to addressing this issue, as follows:

Data approaches, such as k-means and SVM resampling methods based on the Smote algorithm, the fusion of Bayesian posterior probability and distribution density processing categories overlap downsampling methods combined with up-sampling denoising, a multi-model fusion of integrated sampling methods, etc.

Algorithmic approaches, which consider the difference in the cost of different misclassification cases for optimization, e.g., based on cost-sensitive learning algorithms such as meta-learning process and using insensitive algorithmic models such as CatBoost, RF, etc.

For the biased data structure of this paper, an oversampling technique that generates new data is more appropriate.

Smote, which is based on the concept of generative clusters, is the classic oversampling algorithm [48] for handling class-imbalanced data, but the newly generated data samples are based within the range of the native data samples; therefore, the generalization ability is poor. The GAN model learns data distribution through an adversarial training process but has weak performance concerning distributed modeling; therefore, this paper uses the improved model CTGAN [29] to generate new sample data to cope with different real data distributions and solve the sample imbalance problem.

The new dataset consists of 499 sets of data, including 144 sets of native test samples and 355 sets of hybrid training samples. To highlight the model performance of CTGAN, the Yeo-Johnson transformed data are Smote oversampled to generate 332 sets of hybrid training samples, and 144 sets of native test samples are retained for model comparison.

Figure 9 compares the data distributions of the three single-feature variables, with the x-axis indicating the data distribution interval and the y-axis indicating the frequency of the variable distribution. The data distributions generated in both ways are similar to the original distributions. Similarity analysis of univariate kernel density functions: the KL scatter was applied to compare the expectation of the log difference between the original and approximate distribution
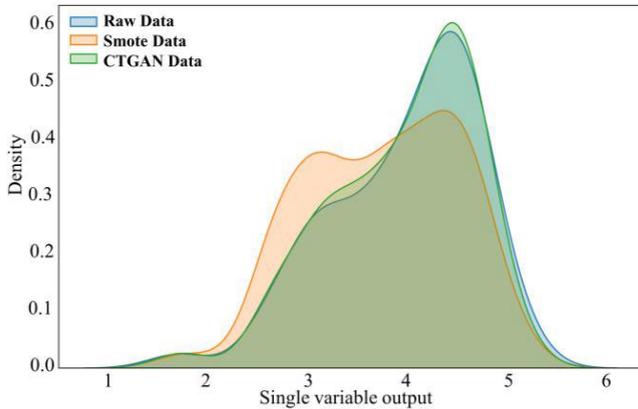
FIGURE 9. CTGAN, smote data distribution comparison.

TABLE 4. Model evaluation indexes.

Evaluation Metrics	Meaning and expression
Precision	TP / (TP + FP)
FPR (False Positive Rate)	FP / (FP + TN)
Recall	TP / (TP + FN)
F1 Score	P*Recall*2/ (P + Recall)
Accuracy	(TP + TN) / (TP + TN + FP + FN)

Note: TP means the true case, FN means positive case predicted to be the negative case; FP means negative case predicted to be the positive case; and TN means a true negative case.

probabilities. Smote and CTGAN calculated the KL scatter values of 44.95 and 13.21, respectively, as follows:

$$D_{KL}(P||Q) = \sum_i^N p(i) \ln \left(\frac{P(i)}{Q(i)} \right) \quad (13)$$

where D(P||Q) denotes the information loss of the probability distribution, Q, estimating the true distribution, P. A larger value for the KL scatter indicates a larger gap between the two distributions.

B. KNN AND CATBOOST MODEL PREDICTION

To compare different algorithms for data sensitivity differences, five different data sets, including the original data, were created for a comparison between the above three nonlinear transformation methods and two data expansion methods, and the generalization ability was analyzed using the K-nearest neighbor (KNN) and CatBoost models. KNN [49] is fast and insensitive to outliers, and the CatBoost [50] model is a gradient-boosting model relying on the GBDT framework, and both models can call the model interpreter and tend to. perform better for different training tasks with better generalization ability for analysis.

Predictions were made for five different combinations of rockburst datasets; the test set was split by 30%; and the original sorting was randomly disrupted and feature coded for the original classification labels (0, 1, 2, 3). For KNN, the CatBoost model uses a grid search to find the optima parameters, select five-fold cross-validation, and obtain the prediction evaluation index to obtain the precision, recall, FPR, F1 value, and accuracy. Table 4 lists the model prediction evaluation index and its meaning.

Through the analysis of the prediction results of the KNN and CatBoost models in Tables 5 and 6, it can be observed that the Yeo-Johnson-Ctgan dataset has the best overall performance in both models, with the highest accuracy reaching 81.25%. The uniform transformation-Ctgan and Box-Cox-Ctgan have similar prediction performance, and the Yeo-Johnson-Smote dataset performs equally well in both models, with an accuracy of about 77%. The original dataset has the worst performance, with only a maximum accuracy of 58.89%.

Both models have good accuracy in Precision for the no explosion and strong explosion categories, while in terms of Recall rate, both models have the best recall rate for the Yeo-Johnson-Ctgan dataset. The classifiers are able to better capture data for the no explosion and strong explosion categories, and the KNN model exhibits better ability to predict positive class samples than CatBoost. FPR is reduced compared to the original dataset, and the KNN model has significantly lower misjudgment rate for negative class samples.

The confusion matrix of the Yeo-Johnson-CTGAN dataset in the KNN model is shown in Figure 10(a). The class0 misclassification rate is 17.14%, the class1 misclassification rate is 21.05%, the class2 misclassification rate is 21.95%, and the class3 misclassification rate is 13.33%.

The Roc curve in Figure 11 depicts the variation process of KNN model performance with the change in the classifier threshold. The horizontal coordinate indicates the classifier performance and the area closer to 1 means stronger the recognition ability, the stronger the classifier performance. Observing the five class curves, it can be seen that the model has a better predictive identification ability for high-risk factors (class 3) and low-risk factors (class 0), and the overall identification performance is better.

C. FEATURE DERIVATION AND COMPARISON

The predictions for different numbers of rockburst features are documented in Figure 12, and datasets of different sizes exhibit large deviations in predictive power influenced by the model prediction accuracy and the number of features. However, in engineering practice, the indicators characterizing rock rupture are not limited to the uniaxial compressive strength of rock (R_c), the axial stress in the surrounding rock of the cavern (σ_L), the maximum initial ground stress perpendicular to the axis of the cavern (σ_{Max}), etc. Furthermore, the criteria of different criteria, such as the engineering criteria, depth criteria, etc., derived from the judgment indicators are very different, so the model relying on only six indicators cannot explain the occurrence of the rockburst phenomenon. The set of feature variables was enriched in the dimensional space through the crossover of multiple features with the help of the Polynomial Features tool. Using this method, a set of 28 features was constructed in Yeo-Johnson-CTGAN, and the model performed well. The newly constructed feature set mixes original features,

TABLE 5. KNN model prediction results.

Model		KNN				
		Precision	Recall rate	FPR	F1	Accuracy
Raw data	N	0.5000	0.6429	0.1184	0.5625	
	L	0.6897	0.6667	0.1500	0.6780	
	M	0.5000	0.6538	0.2656	0.5667	0.5889
	S	0.7778	0.3500	0.0286	0.4828	
Uniform Transformation-CTGAN	N	0.8000	0.7762	0.0648	0.7887	
	L	0.7778	0.7778	0.0741	0.7776	
	M	0.6829	0.7792	0.1204	0.7275	0.7778
	S	0.8750	0.7780	0.0370	0.8235	
Box-Cox-CTGAN	N	0.9667	0.8056	0.0093	0.8788	
	L	0.7000	0.7778	0.1111	0.7368	
	M	0.7073	0.8056	0.1111	0.7532	0.7847
	S	0.8182	0.7500	0.0556	0.7826	
Yeo-Johnson-CTGAN	N	0.8286	0.8056	0.0556	0.8169	
	L	0.7895	0.8333	0.0741	0.8108	
	M	0.7805	0.8889	0.0833	0.8312	0.8125
	S	0.8667	0.7222	0.0370	0.7879	
Yeo-Johnson-Smote	N	0.9487	0.8605	0.0200	0.9024	
	L	0.6316	0.6857	0.1296	0.6575	
	M	0.7059	0.7500	0.0901	0.7273	0.7762
	S	0.8125	0.7879	0.0545	0.8000	

TABLE 6. CatBoost model prediction results.

Model		CatBoost				
		Precision	Recall rate	FPR	F1	Accuracy
Raw data	N	1.0000	0.5000	0.0000	0.6667	
	L	0.6000	0.4000	0.1333	0.4800	
	M	0.4490	0.8462	0.4219	0.5867	0.5778
	S	0.7857	0.5500	0.0429	0.6471	
Uniform transformation-CTGAN	N	1.0000	0.6389	0.0000	0.7797	
	L	0.6444	0.8056	0.1481	0.7160	
	M	0.6275	0.8889	0.1759	0.7356	0.7569
	S	1.0000	0.6944	0.0000	0.8197	
Box-Cox-CTGAN	N	0.9355	0.8056	0.0185	0.7458	
	L	0.6444	0.8056	0.1481	0.6914	
	M	0.6222	0.7778	0.1574	0.7160	0.7500
	S	0.9565	0.6111	0.0093	0.8657	
Yeo-Johnson-CTGAN	N	1.0000	0.7500	0.0000	0.8571	
	L	0.6667	0.8333	0.1389	0.7407	
	M	0.6444	0.8056	0.1481	0.7160	0.7708
	S	0.9259	0.6944	0.0185	0.7937	
Yeo-Johnson-Smote	N	1.0000	0.8375	0.0000	0.9114	
	L	0.6875	0.6286	0.0926	0.6567	
	M	05532	0.8125	0.1892	0.6582	0.7622
	S	0.8929	0.7576	0.0273	0.8197	

quadratic features, and the product cross terms of arbitrary features.

The greater the number of features, the richer the feature information captured by the complex system and the clearer the decision boundary, but increasing the number of data features utilizing derived features will inevitably make the system redundant for features with strong correlations in the data set, causing a decrease in prediction accuracy. PCA [52] was used to structure the sample matrix by dimensionality reduction, covariance eigenvectors, and eigenvalues were calculated to identify principal components, and high contribution rate features were selected to reconstruct the
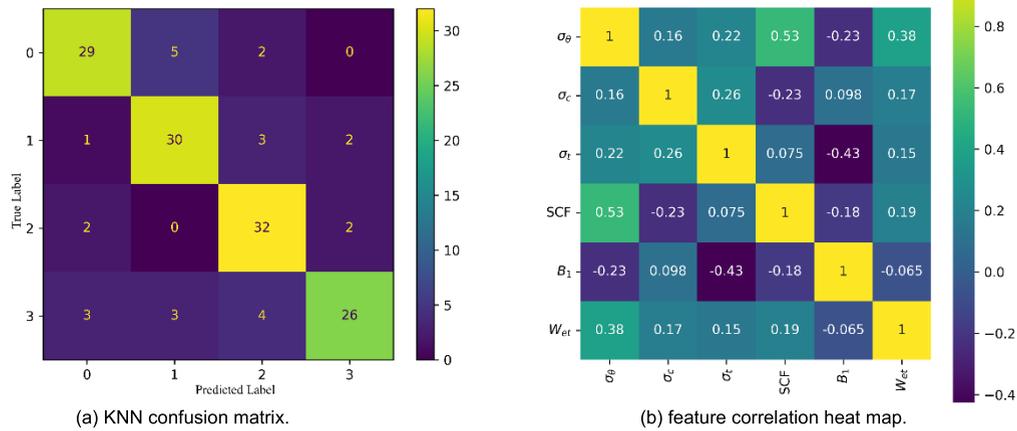
(a) KNN confusion matrix.

(b) feature correlation heat map.

FIGURE 10. Rockburst dataset feature relationships and prediction results.
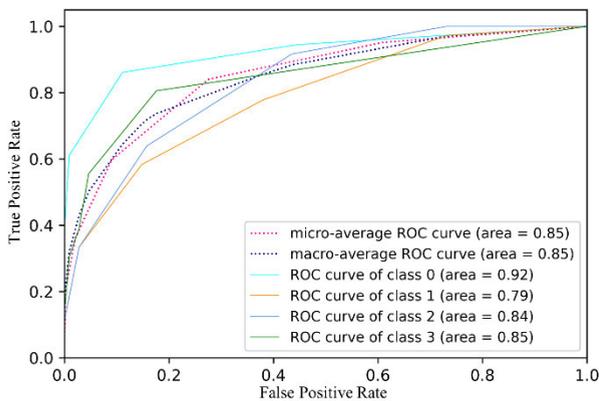

FIGURE 11. KNN ROC multiclassification curve.

data samples based on the cumulative contribution rate in descending order to reduce the interference of cross terms.

The set of features $\beta = \{B(i) | 1 < i < 28, i \in N^*\}$ is chosen to construct a new dataset $[B(i), B(j)]$, where $[j > i \& i, j \in (1, 28)]$, and Figure 12(a) represent the variation in the accuracy of the different sets of features in the KNN and CatBoost models. The KNN model reached a maximum accuracy of 79.86% when the number of features was equal to 4, after which the accuracy leveled off and remained at 78.47%. The CatBoost model achieves 83.33% accuracy when the

number of features equals 21, which is 8.11% higher than the original feature prediction accuracy and exceeds the original results in 73.08% of the datasets.

Figure 12(b) shows the change curve of the prediction accuracy when increasing the number of polynomial features in the dataset, and the KNN model selects the number of features with the best prediction results. it can be seen that the fluctuation in the prediction accuracy is not obvious around 79.86%. The CatBoost model takes into account the training time problem caused by the extended dataset and chooses a prediction accuracy of 81.25% with several features equal to 13. Performing the characteristic number derivation, it can

be seen that the prediction results fluctuate more overall, showing a decreasing trend.

D. LEVEL 1 MODEL ASSESSMENT

The machine learning process for high-complexity systems is often accompanied by multi-level decision-making tasks. Model algorithms achieve higher prediction accuracy through autonomous selection and abstraction judgments, but this decision complexity turns the model into a black box, making the predictions difficult to understand and reducing confidence in the results [51]. Therefore, a visual analysis of the model decision process is necessary. To verify the validity of the primary model and the necessity of constructing a secondary model, the Yeo-Johnson-CTGAN dataset with the best prediction results was subjected to model characterization using the KNN algorithm.

1) SHAP VALUE ANALYSIS

The SHAP framework is an interpretive framework for explainng model predictions that helps us understand the outputs and decisions of machine learning models, and provides explanations for both global and local results [50]. The feature density scatter plot in Figure 13 explains, from a macro perspective, the predictions of four types of rockburst intensity by a KNN model, given six different rockburst feature values after an initial transformation. This plot visually demonstrates the influence of each feature on the model prediction and the distribution of feature values. The horizontal axis represents SHAP values, and the vertical axis represents features in descending order. Each scatter plot represents the SHAP value of a sample, and the density of the scatter plot represents the degree to which the samples are clustered. The color represents the size of the feature value, with darker colors indicating larger feature values.

Observing the four figures, we can see that the features σ_c and σ_θ have the highest importance in the model, and their influence on the model decision is the greatest. The samples of the *SCF* feature are distributed around SHAP value =0, indicating that its impact on the model decision is minimal.
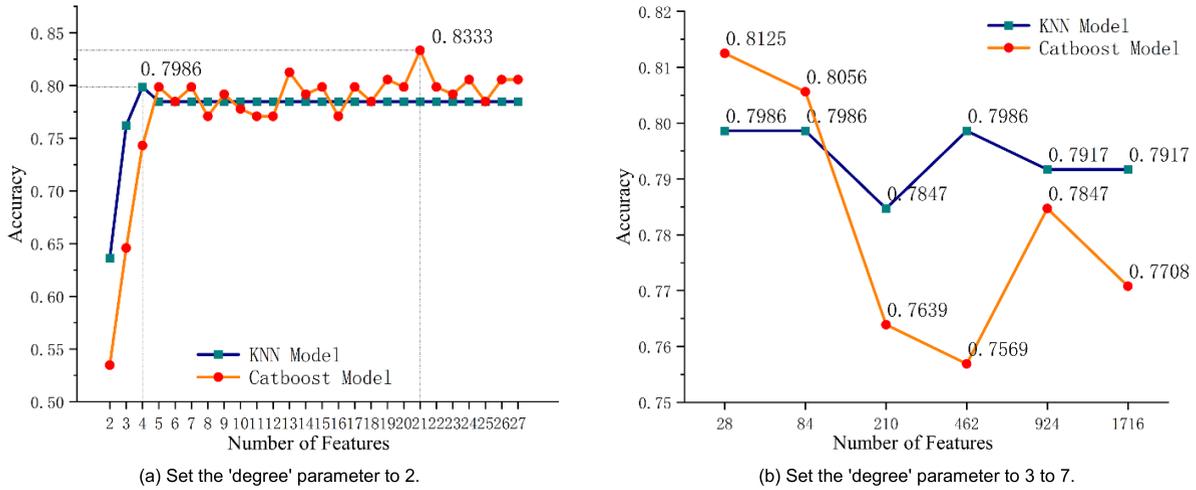
(a) Set the 'degree' parameter to 2.

(b) Set the 'degree' parameter to 3 to 7.

FIGURE 12. KNN, CatBoost feature number, and accuracy change curve.

In Figure (a), the values of the σ_c feature are relatively scattered and mainly cluster around the interval $[0.2, 0.0]$, which mixes a large number of feature values. Horizontally, the magnitude of the feature values does not significantly affect the decision of class0 rockbursts. Additionally, the boundary between σ_θ and W_{et} features is clearly distinguished. The positive region of the SHAP values of these features mainly includes samples with smaller feature values, while the negative region includes samples with larger feature values. The changes of these two features are somewhat correlated, but their importance to the model is relatively low.

Similar to Figures (b) and (c), the σ_c feature is mainly clustered in the negative region of the SHAP values, while a large number of samples are also gathered in the positive region. The horizontally extended samples mix different values and are not clearly distinguished. Other than the W_{et} feature, the other features do not have clear boundary distinctions. Therefore, this validates the model's low prediction accuracy for class1 and class2 rockbursts.

In Figure (d), we can see that the σ_θ feature has the highest importance for predicting class3 rockbursts, and the feature's influence is clearly divided between positive and negative SHAP values. A large number of samples are concentrated between $[-0.1, 0.0]$, with medium-sized values being the main trend. In the positive region, high feature values are dominant, while in the negative region, low values are dominant. This indicates that the rockburst strength increases with the maximum tangential stress of the surrounding rock, thus the model has the best prediction performance for class3 rockbursts.

A SHAP Dependence Plot is a type of visualization that shows the impact of a single feature on a model's prediction. By observing the macroscopic trend reflected in the plot, we can understand how changes in the feature's values affect the model's predictions. The Figure 14 shows the SHAP dependence plot of σ_θ and B_1 features on the prediction of

class3 rock burst. The x-axis represents the values of σ_θ feature, while the y-axis represents the size of SHAP values of σ_θ, which is reflected by the color depth of the comparison feature B_1. By observing the plot, we can find that the SHAP values show an overall fluctuating upward trend with the increase of σ_θ, and the same feature value corresponds to different SHAP values, indicating an interaction effect between σ_θ and B_1. When $\sigma_\theta > 5$, relatively small B_1 values have a positive effect on the prediction results, while when $\sigma_\theta < 2$, relatively small B_1 has a negative impact on the model prediction. However, overall, a large number of samples gather around SHAP $= 0$, and the overall trend is not clear.

2) CORRELATION ANALYSIS

Too high a degree of correlation between the features of the rockburst data can cause information redundancy and affect the prediction results. Figure 10(b) shows the feature heatmap in the Yeo-Johnson-CTGAN dataset, a darker color represents a stronger correlation between the two. Generally speaking, correlation values between $[0.2, 0.4]$ belong to weak correlations, $[0.4, 0.6]$ belong to medium correlations, and $[0.6, 0.8]$ indicate strong correlations. By calculating the spearman correlation coefficient, it was found that the highest degree of correlation between SCF and σ_θ is 0.53, which is a medium correlation, followed by the value of 0.38 between W_{et} and σ_θ, which is a weak correlation.

E. ANALYSIS OF SECONDARY MODEL PREDICTION RESULTS
1) NECESSITY ANALYSIS

From the KNN confusion matrix, Figure 10(a), it can be seen that the misclassification rate of low-intensity (class 0 and 1) labels reaches 19.18% and the misclassification rate of high-intensity (class 2 and 3) labels is 18.31%. From the correlation coefficient between the features in Figure 10(b), there are redundant features and multicollinearity.
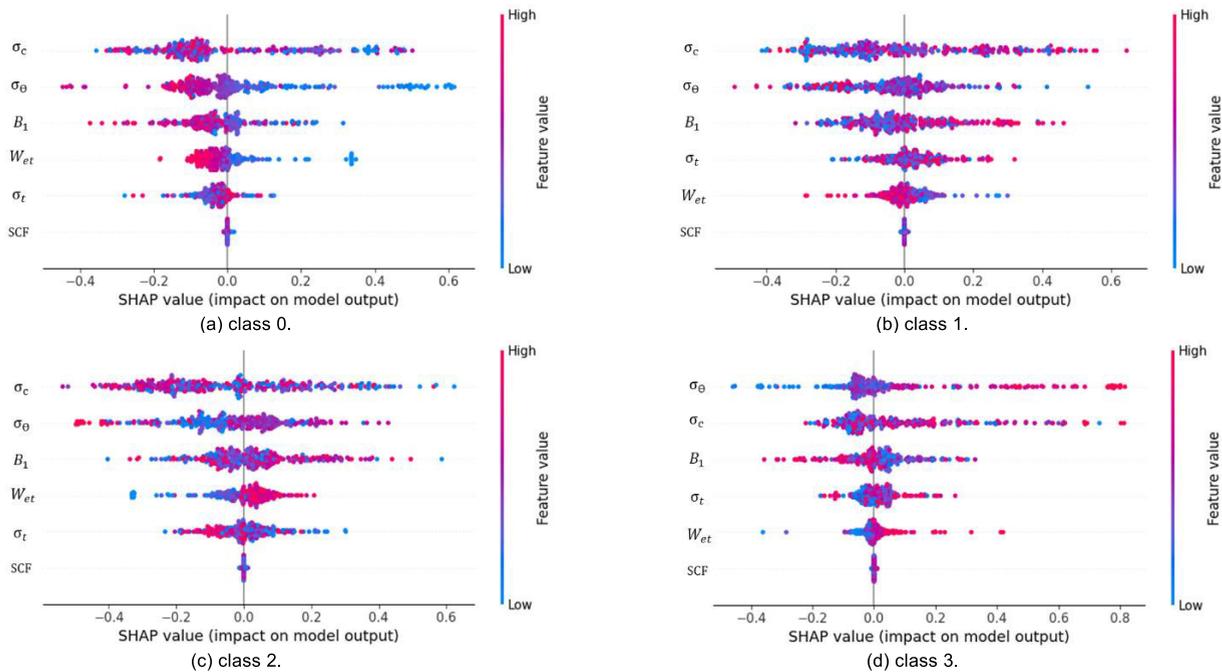
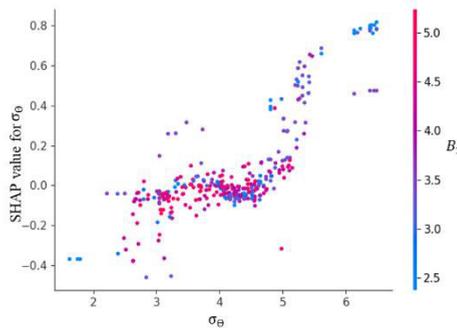**FIGURE 13. SHAP diagram of the impact of KNN model features.**

FIGURE 14. Feature dependency diagram for class 3.

Observing the SHAP values (Figure 13), we can see that the classification feature values are not obvious, the clustering area is confusing, and the classification effect is not good. Observing the misclassified data, we found that the prediction effect is not good due to the longitudinal expansion of the dataset or due to the inconsistent classification standard of the original labels, such that some data cannot match the original labels.

In practical engineering contexts, misclassification problems caused by low model recognition accuracy can lead to increased anti-impact costs, trigger impact events, reduce model reliability, and cause trouble in on-site work. Although the primary model in this study achieved an accuracy of 81.25%, there is still a gap from the expected error allowed in engineering. Therefore, it is necessary to construct a secondary model to further improve prediction accuracy. In this study, the secondary model used K-Means clustering

to reconstruct classification labels and the ISOMAP dimensionality reduction to visualize this process.

2) DIMENSIONALITY REDUCTION VISUALIZATION AND RELABELING

To visualize the data distribution characteristics of the Yeo-Johnson-CTGAN dataset, the ISOMAP nonlinear dimensionality reduction algorithm was used to filter the three most important features for normalization and compare them with the PCA linear reduction algorithm (Figure 15 and Figure 16). Figure 17 shows the clustering metric selection process, traversing the model classification effect with different values of the number of clusters. (a) The ratio of the sum of all inter-cluster dispersion to the sum of intra-cluster dispersion is calculated by the variance ratio criterion (dispersion denotes the sum of squared distances), and higher CH values (Calinski-Harabaz Index) represent better clustering (denser concerning observations, better inter-cluster separation). (b) is a measure used to calculate the merit of the clustering technique and takes a value in the range [-1, 1], where the larger the value represents the relative distance of clusters in the clustering space, which are clearly dispersed and distinguished. (c) With the help of KElbowVisualizer visualization, the optimal number of clusters is chosen by fitting a range of model values, and the dashed line (n= 4) on this figure indicates the most suitable value. By looking at the three images, n= 4 is the optimal number of clusters, with relatively good inter-cluster dispersion and achieving the best clustering, which is consistent with the number of labels that need to be reconstructed in the dataset, thus achieving good training results.
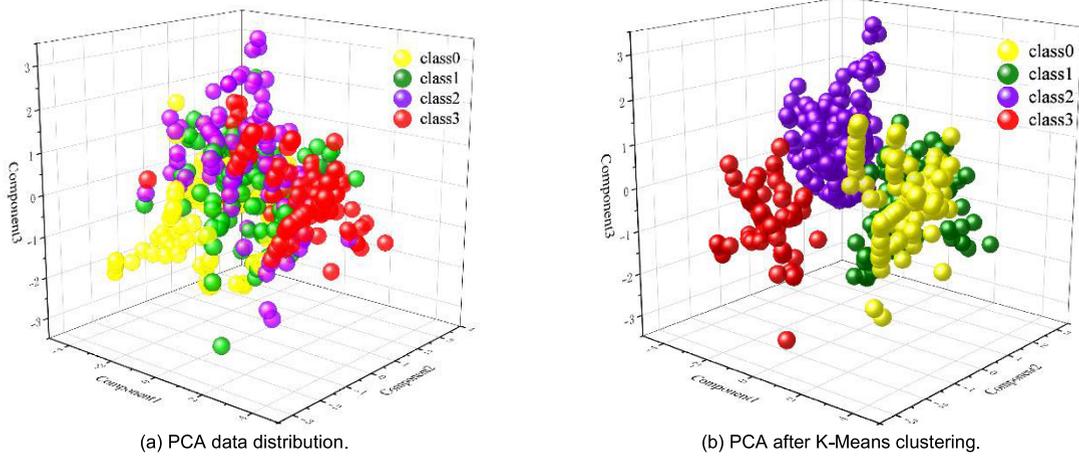
(a) PCA data distribution.

(b) PCA after K-Means clustering.

FIGURE 15. PCA algorithm K-Means clustering before and after effect comparison.

(a) ISOMAP data distribution.

(b) ISOMAP after K-Means clustering.
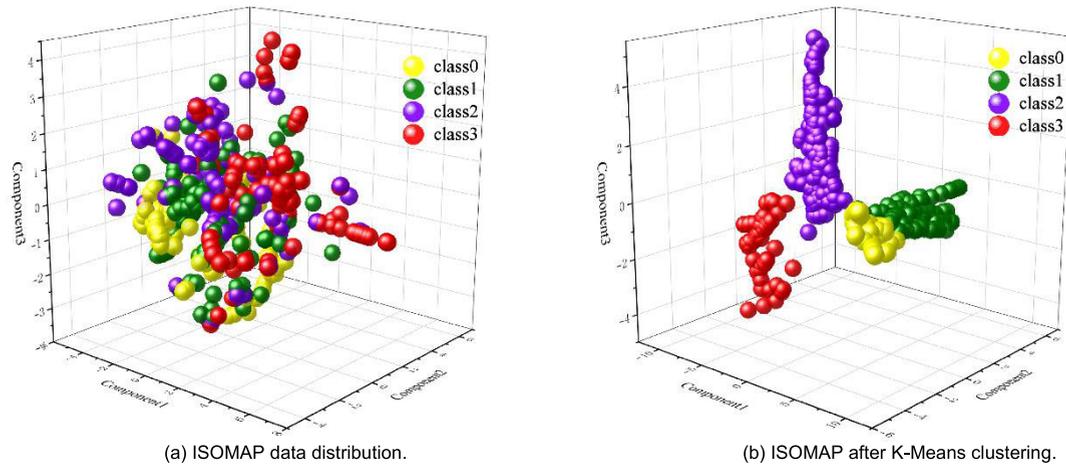
FIGURE 16. ISOMAP algorithm K-Means clustering before and after comparison.

(a) Calinski-Harabasz Index.

(b) Silhouette Coefficient.
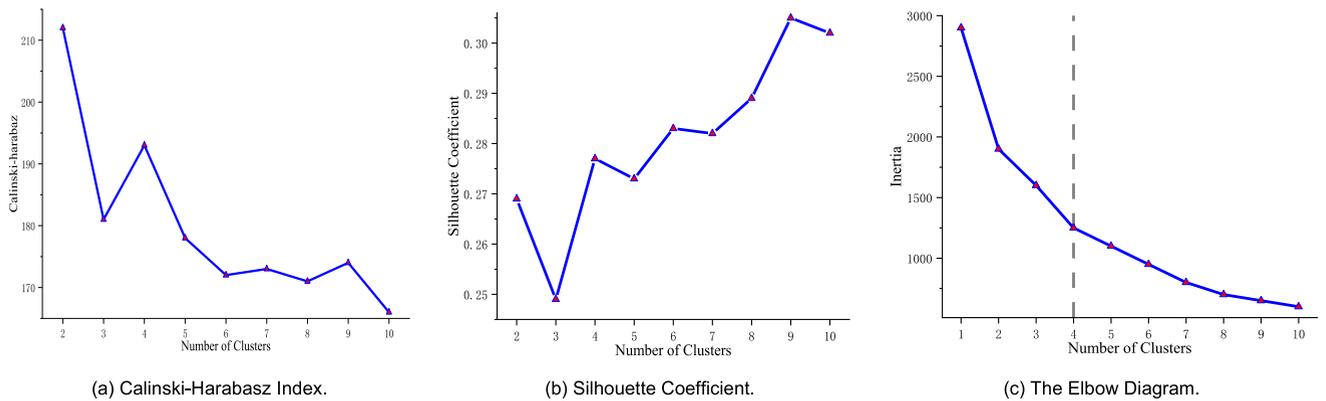
(c) The Elbow Diagram.

FIGURE 17. K-Means clustering model evaluation metrics.

3) COMPARISON OF PREDICTED RESULTS

The prediction results of single-model, integrated stacking, and deep forest models are compared in Table (7) and table (8) for the dataset after ISOMAP reconstruction labeling. The deep forest model achieves the highest prediction accuracy of 98.60%, which is an average improvement of 1.40%

compared to the single-model prediction results and 1.40% compared to stacking, and achieves 100% accuracy for strong rockburst and no rockburst, and reduces the misclassification rate of light rockburst and moderate rockburst by 15.96%.

The integrated model with eight algorithms stacked is analyzed, and Table 7 shows that the accuracy of the

TABLE 7. Single model and Stacking model prediction results.

Model	Partial Parameter Optimization Value	Prediction Accuracy
Ridge Regression	alpha=0.1	0.7413
DT	max depth=8, min samples leaf=2 min_samples_split=4	0.9650
RF	max depth=16 n_estimators=50	0.9231
AdaBoost	n_estimators=100	0.9671
GBR	max depth=2 n_estimators=250	0.9441
XGBoost	gamma=0 n_estimators=500	0.9371
SVR	c=10	0.9510
MLP	hidden_layer_sizes=50 solver='lbfgs'	0.9441
Staking		0.9720
KNN	n_neighbors=2	0.9790
CatBoost	n_estimators=1000, loss_function='RESE' l2_leaf_reg=3	0.9650

TABLE 8. Evaluation indexes of deep forest model.

		Precision	Recall rate	F1	Confusion Matrix	Accuracy
Deep Forest	N	1.0000	0.9697	0.9846	$\begin{bmatrix} 32 & 1 & 0 & 0 \\ 0 & 37 & 0 & 0 \\ 0 & 0 & 46 & 0 \\ 0 & 0 & 1 & 26 \end{bmatrix}$	0.9860
	L	0.9737	1.0000	0.9867		
	M	0.9787	1.0000	0.9892		
	S	1.0000	0.9630	0.9811		

integrated model is 97.2%, which is higher than the accuracy of the integrated DT (Decision Tree), GBR (Gradient Boosting Regression), SVR (Support Vector Regression), and other algorithms. but 0.7% less accurate than the KNN algorithm in this paper.

The integrated model is more influenced by the combination strategy [27] and hyperparameter optimization, reducing the generalization ability [28]. When multiple nonlinear systems determine decision boundaries, the integrated model training process is influenced by trees, making the training results unstable and the training time long, while a deep forest has higher accuracy and a faster iteration time. Therefore, the deep forest model has stronger robustness and better adaptability.

The two-level rock burst prediction model constructed in this article achieved an overall prediction accuracy of 81.25% and a misclassification rate of 6.25% for the first-level model, which utilized Yeo-Johnson transformation and CTGAN enhancement. The second-level prediction model achieved an accuracy of 98.60%, which is a significant improvement compared to the original classification data with a prediction accuracy of 58.89%. This provides valuable reference for engineering applications.

V. CONCLUSION

1) This paper conducted Yeo-Johnson, Box-Cox transformation, and uniform transformation on the original rock burst data in terms of data structure and enhanced data using the CTGAN framework, and established prediction models for six rock burst features. The results showed that the Yeo-Johnson algorithm is suitable for a wider range of data distributions, exhibiting stronger adaptability. The Yeo-Johnson-CTGAN dataset performed the best, achieving a prediction accuracy of 81.25%. Compared to the original model's prediction results, the average accuracy was improved by 0.19 after the first-level model processing. In addition, expanding the data features using polynomial extension is also an effective means of improving prediction results. The prediction results may vary due to the influence of the model structure, but overall, there was a slight improvement with an average prediction accuracy increase of 2.5%.

2) The accuracy of the prediction was further improved through different model selection and improvement strategies. A CTGAN expansion plan that fits the original distribution was selected for data expansion, and its effectiveness was verified through quantitative comparison. To achieve accurate recognition in engineering application scenarios, improvements were made in the complexity of the model by increasing the degree of aggregation of the same type of rock burst data, expanding the inter-class boundary, and visualizing the process. Finally, an ensemble model was used to further improve the prediction accuracy, reaching 98.6%.

3) Model interpretation is an important direction for understanding model decisions. In this paper, we used the SHAP visualization tool to validate the model evaluation results through a Summary plot and a Dependence plot. These plots showed the contribution of each feature to the prediction results and the importance ranking of each feature.

4) Although this paper significantly improved the prediction accuracy under the two-level prediction model, there is still insufficient research on the model decision-making process, limited sample size, and the sensitivity to different algorithms that need to be further studied. The next step should be to expand the research scope, gain a deeper understanding of the inherent decision-making logic within the model, and expand the types of data that the model can handle.

REFERENCES

[1] X. Yin, Q. Liu, Y. Pan, X. Huang, J. Wu, and X. Wang, "Strength of stacking technique of ensemble learning in rockburst prediction with imbalanced data: Comparison of eight single and ensemble models," *Natural Resour. Res.*, vol. 30, no. 2, pp. 1795–1815, Apr. 2021.

[2] Q. Wu, F. Han, S. Liang, F. Sun, D. Wan, H. Su, F. Ma, and Q. Wu, "Development law of mining fracture and disaster control technology under hard and thick magmatic rock," *Sustainability*, vol. 14, no. 18, 2022, Art. no. 11140.

[3] J. Guo, J. Guo, Q. Zhang, and M. Huang, "Research on rockburst classification prediction based on BP-SVM model," *IEEE Access*, vol. 10, pp. 50427–50447, 2022.

[4] A. Iannacchione and J. C. Zelanko, "Occurrence and remediation of coal mine bumps: A historical review," Tech. Rep.

[5] D. Guo, H. Chen, L. Tang, Z. Chen, and P. Samui, "Assessment of rockburst risk using multivariate adaptive regression splines and deep forest model," *Acta Geotech.*, vol. 17, no. 4, pp. 1183–1205, Apr. 2022.

[6] A. Keneti and B.-A. Sainsbury, "Review of published rockburst events and their contributing factors," *Eng. Geol.*, vol. 246, pp. 361–373, Oct. 2018.

[7] X. Chen, W. Li, and X. Yan, "Analysis on rock burst danger when fully-mechanized caving coal face passed fault with deep mining," *Saf. Sci.*, vol. 50, no. 4, pp. 645–648, Apr. 2012.

[8] G. Feng, G. Xia, B. Chen, Y. Xiao, and R. Zhou, "A method for rockburst prediction in the deep tunnels of hydropower stations based on the monitored microseismicity and an optimized probabilistic neural network model," *Sustainability*, vol. 11, no. 11, p. 3212, 2019.

[9] S. Wu, Z. Wu, and C. Zhang, "Rock burst prediction probability model based on case analysis," *Tunnelling Underground Space Technol.*, vol. 93, Nov. 2019, Art. no. 103069.

[10] S. Afraei, K. Shahriar, and S. H. Madani, "Developing intelligent classification models for rock burst prediction after recognizing significant predictor variables, section 1: Literature review and data preprocessing procedure," *Tunnelling Underground Space Technol.*, vol. 83, pp. 324–353, Jan. 2019.

[11] Y. Xue, C. Bai, F. Kong, D. Qiu, L. Li, M. Su, and Y. Zhao, "A two-step comprehensive evaluation model for rockburst prediction based on multiple empirical criteria," *Eng. Geol.*, vol. 268, Apr. 2020, Art. no. 105515.

[12] S.-J. Miao, M.-F. Cai, Q.-F. Guo, and Z.-J. Huang, "Rock burst prediction based on in-situ stress and energy accumulation theory," *Int. J. Rock Mech. Mining Sci.*, vol. 83, pp. 86–94, Mar. 2016.

[13] N. M. Khan, M. Ahmad, K. Cao, I. Ali, W. Liu, H. Rehman, and T. Ahmed, "Developing a new bursting liability index based on energy evolution for coal under different loading rates," *Sustainability*, vol. 14, no. 3, p. 1572, 2022.

[14] Z. H. Ouyang, Q. X. Qi, S. K. Zhao, B. Y. Wu, and N. B. Zhang, "The mechanism and application of deep-hole precracking blasting on rockburst prevention," *Shock Vib.*, vol. 2015, p. 7, May 2015, Art. no. 625691.

[15] C.-L. Wang, C. Cao, C.-F. Li, X.-S. Chuai, G.-M. Zhao, and H. Lu, "Experimental investigation on synergetic prediction of granite rockburst using rock failure time and acoustic emission energy," *J. Central South Univ.*, vol. 29, no. 4, pp. 1262–1273, Apr. 2022.

[16] J. Liu, S. D. Liu, and Y. Cao, "Self-potential characteristics in deep rock mass damage based on point discharge mechanism," *J. Geophys. Eng.*, vol. 16, no. 4, pp. 742–752, 2019.

[17] F. X. Jiang, S. H. Yang, Y. H. Cheng, X. M. Zhang, Z. Y. Mao, and F. J. Xu, "A study on microseismic monitoring of rock burst in coal mine," *J. Geophys.*, vol. 49, no. 5, pp. 1511–1516, Sep. 2006.

[18] W. Cai, L. Dou, M. Zhang, W. Cao, J.-Q. Shi, and L. Feng, "A fuzzy comprehensive evaluation methodology for rock burst forecasting using microseismic monitoring," *Tunnelling Underground Space Technol.*, vol. 80, pp. 232–245, Oct. 2018.

[19] X. Chen, Q. Lian, X. Chen, and J. Shang, "Surface crack detection method for coal rock based on improved YOLOv5," *Appl. Sci.*, vol. 12, no. 19, p. 9695, Sep. 2022.

[20] Y. G. Yang and F. L. Niu, "Using unsupervised machine learning for clustering seismic noise: A case study of a dense seismic array at the Weifang segment of the tanlu fault," *Chin. J. Geophys.*, vol. 65, no. 7, pp. 2573–2594, Jul. 2022.

[21] Y. Y. Di and E. Y. Wang, "Rock burst precursor electromagnetic radiation signal recognition method and early warning application based on recurrent neural networks," *Rock Mech. Rock Eng.*, vol. 54, no. 3, pp. 1449–1461, Mar. 2021.

[22] L. Qiong, C. Zheng, H. Jian-Jun, H. Si-Yu, W. Rui, Y. Hao-Tao, and S. Hua-Jun, "Reconstructing the 3D digital core with a fully convolutional neural network," *Appl. Geophys.*, vol. 17, no. 3, pp. 401–410, Sep. 2020.

[23] S. Zhai, G. Su, S. Yin, B. Zhao, and L. Yan, "Rockburst characteristics of several hard brittle rocks: A true triaxial experimental study," *J. Rock Mech. Geotech. Eng.*, vol. 12, no. 2, pp. 279–296, Apr. 2020.

[24] M. Cheng, L. Zhong, Y. Ma, X. Wang, P. Li, Z. Wang, and Y. Qi, "A new drought monitoring index on the Tibetan plateau based on multisource data and machine learning methods," *Remote Sens.*, vol. 15, no. 2, p. 512, 2023.

[25] J. Zhou, X. Li, and H. S. Mitri, "Evaluation method of rockburst: State-of-the-art literature review," *Tunnelling Underground Space Technol.*, vol. 81, pp. 632–659, Nov. 2018.

[26] 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, Sep. 2016, Art. no. 04016003.

[27] D. Li, Z. Liu, D. J. Armaghani, P. Xiao, and J. Zhou, "Novel ensemble intelligence methodologies for rockburst assessment in complex and variable environments," *Sci. Rep.*, vol. 12, no. 1, p. 23, Feb. 2022.

[28] D. Li, Z. Liu, D. J. Armaghani, P. Xiao, and J. Zhou, "Novel ensemble tree solution for rockburst prediction using deep forest," *Mathematics*, vol. 10, no. 5, p. 787, Mar. 2022.

[29] L. Xu, M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, "Modeling tabular data using conditional GAN," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–11.

[30] W. Yan, J. Chen, Y. Tan, R. He, and S. Yan, "Surface dynamic damage prediction model of horizontal coal seam based on the idea of wave lossless propagation," *Int. J. Environ. Res. Public Health*, vol. 19, no. 11, p. 6862, Jun. 2022.

[31] L. Sun, N. Hu, Y. Ye, W. Tan, M. Wu, X. Wang, and Z. Huang, "Ensemble stacking rockburst prediction model based on Yeo–Johnson, K-means SMOTE, and optimal rockburst feature dimension determination," *Sci. Rep.*, vol. 12, no. 1, p. 15352, 2022.

[32] Y. He and Y. Zheng, "Short-term power load probability density forecasting based on Yeo–Johnson transformation quantile regression and Gaussian kernel function," *Energy*, vol. 154, pp. 143–156, Jul. 2018.

[33] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[34] Z. A. Lin, A. Khetan, G. Fanti, and S. Oh, "PacGAN: The power of two samples in generative adversarial networks," in *Proc. 32nd Conf. Neural Inf. Process. Syst.*, vol. 31, Montreal, QC, Canada, 2018, pp. 1–10.

[35] Y. Liu, J. Chen, C. Tan, J. Zhan, S. Song, W. Xu, J. Yan, Y. Zhang, M. Zhao, and Q. Wang, "Intelligent scanning for optimal rock discontinuity sets considering multiple parameters based on manifold learning combined with UAV photogrammetry," *Eng. Geol.*, vol. 309, p. 22, Nov. 2022, Art. no. 106851.

[36] S. E. Sesnie, C. I. Espinosa, A. K. Jara-Guerrero, and M. F. Tapia-Armijos, "Ensemble machine learning for mapping tree species alpha-diversity using multi-source satellite data in an Ecuadorian seasonally dry forest," *Remote Sens.*, vol. 15, no. 3, p. 583, 2023.

[37] Z. Xu, W.-C. Ni, and Y.-H. Ji, "Rotation forest based on multimodal genetic algorithm," *J. Central South Univ.*, vol. 28, no. 6, pp. 1747–1764, Jun. 2021.

[38] Z.-H. Zhou and J. Feng, "Deep forest," *Nat. Sci. Rev.*, vol. 6, no. 1, pp. 74–86, 2019.

[39] J. Tian, Y. Wang, and S. Gao, "Analysis of mining-related injuries in Chinese coal mines and related risk factors: A statistical research study based on a meta-analysis," *Int. J. Environ. Res. Public Health*, vol. 19, no. 23, p. 16249, Dec. 2022.

[40] L.-J. Dong, X.-B. Li, and K. Peng, "Prediction of rockburst classification using random forest," *Trans. Nonferrous Met. Soc. China*, vol. 23, no. 2, pp. 472–477, Feb. 2013.

[41] J. Zhang, Y. Wang, Y. Sun, and G. Li, "Strength of ensemble learning in multiclass classification of rockburst intensity," *Int. J. Numer. Anal. Methods Geomech.*, vol. 44, no. 13, pp. 1833–1853, Sep. 2020.

[42] Y. Xue, Z. Li, S. Li, D. Qiu, Y. Tao, L. Wang, W. Yang, and K. Zhang, "Prediction of rock burst in underground caverns based on rough set and extensible comprehensive evaluation," *Bull. Eng. Geol. Environ.*, vol. 78, no. 17, pp. 417–429, Feb. 2019.

[43] Y. Y. Pu, D. B. Apel, and H. W. Xu, "Rockburst prediction in kimberlite with unsupervised learning method and support vector classifier," *Tunnelling Underground Space Technol.*, vol. 90, pp. 12–18, Aug. 2019.

[44] R. Liu, Y. C. Ye, N. Y. Hu, H. Chen, and X. H. Wang, "Classified prediction model of rockburst using rough sets-normal cloud," *Neural Comput. Appl.*, vol. 31, no. 12, pp. 8185–8193, Dec. 2019.

[45] Q. Jia, L. Wu, B. Li, C. Chen, and Y. Peng, "The comprehensive prediction model of rockburst tendency in tunnel based on optimized unascertained measure theory," *Geotech. Geol. Eng.*, vol. 37, no. 4, pp. 3399–3411, Aug. 2019.

[46] Y. G. Xue, C. H. Bai, D. H. Qiu, F. M. Kong, and Z. Q. Li, "Predicting rockburst with database using particle swarm optimization and extreme learning machine," *Tunnelling Underground Space Technol.*, vol. 98, p. 12, Apr. 2020, Art. no. 103287.

[47] Y. H. Wang, W. D. Li, Q. G. Li, and G. H. Tan, "Method of fuzzy comprehensive evaluations for rockburst prediction," *Chin. J. Rock Mech. Eng. Geol.*, vol. 5, pp. 15–23, Feb. 1998.

[48] K. Jiang, J. Lu, and K. Xia, "A novel algorithm for imbalance data classification based on genetic algorithm improved SMOTE," *Arabian J. Sci. Eng.*, vol. 41, no. 8, pp. 3255–3266, Aug. 2016.

[49] L. Ma, J. Cai, X. Dai, and R. Jia, "Research on rockburst risk level prediction method based on LightGBM-TCN-RF," *Appl. Sci.*, vol. 12, no. 16, p. 8226, Aug. 2022.

[50] B. Ibrahim, I. Ahenkorah, and A. Ewusi, "Explainable risk assessment of rockbolts' failure in underground coal mines based on categorical gradient boosting and Shapley additive explanations (SHAP)," *Sustainability*, vol. 14, no. 19, p. 16, Oct. 2022, Art. no. 11843.

[51] H. L. Lee, J. S. Kim, C. H. Hong, and D. K. Cho, "Ensemble learning approach for the prediction of quantitative rock damage using various acoustic emission parameters," *Appl. Sci.*, vol. 11, no. 9, p. 16, May 2021, Art. no. 4008.

[52] S. Xu, X. Yang, F. Zhang, and J. Liu, "Application of a microseismic method of rock burst risk assessment under blasting mining in ashele copper mine," *Shock Vib.*, vol. 2022, pp. 1–13, Feb. 2022.

QINGWANG LIAN is currently a Professor with the Key Laboratory of In-Situ Property-Improving Mining of the Ministry of Education, Taiyuan University of Technology. He has been engaged in science and technology management and technology development for a long time and has undertaken teaching and research work in the fields of mining engineering and mine safety economy. His research interests include mine pressure and rock control, with a special focus on the theory and technology of coal mine impact ground pressure. He has completed four national research projects, including the National Science and Technology Support Program "Key Technology Research and Demonstration of Production System for Comprehensive Mining Face of 10 Million Tons per Year" and "Integrated Demonstration of Key Technology for High Recovery Rate Mining of Difficult Resources in Large Coal Base." He has won four national research projects and many provincial science and technology projects. He has been awarded the first, second, and third prizes for scientific and technological progress in Shanxi Province and the first prize for scientific and technological progress in the China Coal Industry Association. He has published many academic papers, patents, and books.

XINLIN CHEN was born in Shanxi, China, in 1996. She is currently pursuing the master's degree with the Key Laboratory of In-Situ Property-Improving Mining of the Ministry of Education, Taiyuan University of Technology. Her research interests include computer vision, image processing, and deep learning applications in smart mines. She has published two articles in international journals on these topics.

JIN SHANG was born in Shanxi, China, in 1996. He is currently pursuing the master's degree with the Key Laboratory of In-Situ Property-Improving Mining of the Ministry of Education, Taiyuan University of Technology. His research interests include data modeling, the application of machine learning in geotechnical engineering, and the risk analysis of rockburst impacts. He has published one academic article on these topics.

HAORU YANG was born in Shanxi, China, in 1996. He is currently pursuing the master's degree with the Key Laboratory of In-Situ Property-Improving Mining of the Ministry of Education, Taiyuan University of Technology. His research interests include numerical modeling and the influence of the microstructural characteristics of coal on the impact damage of coal bodies. He has published an article in an international journal on these topics.

· · ·