

Article

Deep Learning in Rockburst Intensity Level Prediction: Performance Evaluation and Comparison of the NGO-CNN-BiGRU-Attention Model

Hengyu Liu ¹ , Tianxing Ma ^{2,*}, Yun Lin ¹, Kang Peng ¹, Xiangqi Hu ³, Shijie Xie ¹ and Kun Luo ¹

¹ School of Resources and Safety Engineering, Central South University, Changsha 410083, China; 19891549715@163.com (H.L.); yunlin617@csu.edu.cn (Y.L.); pengkang2020@csu.edu.cn (K.P.); xieshijieabc@126.com (S.X.); 215511037@csu.edu.cn (K.L.)

² Ocean College, Zhoushan Campus, Zhejiang University, Hangzhou 316021, China

³ School of Civil Engineering, Wuhan University, Wuhan 430072, China; geoxqhu@gmail.com

* Correspondence: m16673133031@163.com

Abstract: Rockburst is an extremely hazardous geological disaster. In order to accurately predict the hazardous degree of rockbursts, this paper proposes eight new classification models for predicting the intensity level of rockbursts based on intelligent optimisation algorithms and deep learning techniques and collects 287 sets of real rockburst data to form a sample database, in which six quantitative indicators are selected as feature parameters. In order to validate the effectiveness of the constructed eight machine learning prediction models, the study selected Accuracy, Precision, Recall and F1 Score to evaluate the prediction performance of each model. The results show that the NGO-CNN-BiGRU-Attention model has the best prediction performance, with an accuracy of 0.98. Subsequently, engineering validation of the model is carried out using eight sets of real rockburst data from Daxiangling Tunnel, and the results show that the model has a strong generalisation ability and can satisfy the relevant engineering applications. In addition, this paper also uses SHAP technology to quantify the impact of different factors on the rockburst intensity level and found that the elastic strain energy index and stress ratio have the greatest impact on the rockburst intensity level.

Keywords: rockburst prediction; intelligent optimisation; deep learning; NGO-CNN-BiGRU-Attention model



Citation: Liu, H.; Ma, T.; Lin, Y.; Peng, K.; Hu, X.; Xie, S.; Luo, K. Deep Learning in Rockburst Intensity Level Prediction: Performance Evaluation and Comparison of the NGO-CNN-BiGRU-Attention Model. *Appl. Sci.* **2024**, *14*, 5719. <https://doi.org/10.3390/app14135719>

Academic Editor: Giuseppe Lacidogna

Received: 31 May 2024

Revised: 22 June 2024

Accepted: 27 June 2024

Published: 29 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rockburst [1] is a dynamic instability phenomenon in which the surrounding rock is spalled, blocked, ejected or thrown due to the abrupt and severe dissipation of elastic strain energy in the rock, and its essence is the surrounding rock has gained kinetic energy [2]. As a kind of sudden, random and highly hazardous geological disaster [3], the process of rockburst activity usually has three phases [4], which are the inactive phase, the active phase and the destructive phase. During the inactive phase, the elastic stress energy accumulation is small; in the destruction stage, large-scale rockbursts usually occur; and the active stage is between the inactive stage and the destruction stage of the state, usually accompanied by medium and small-scale rockbursts. In recent years, the frequency and severity of rockbursts have increased with the increase in geological activities due to the increase in underground space engineering activities, as well as the deep mining of mines. Such geological hazards not only affect the construction process [5,6] but also cause harm to social production activities, citizens' lives and property safety. In view of the hazards of rock explosions, how to prevent the occurrence of rock explosion disasters is particularly important, and rock explosion prediction and evaluation is an important prerequisite for the development of prevention and control measures. Therefore, the accurate and reliable prediction of high-intensity rock explosion disasters has become a research hotspot in related fields [7,8].

Due to the extreme complexity of the rock explosion prediction problem, there is currently no set of mature theories and methods at home and abroad, but they can be broadly divided into the field measurement method and theoretical analysis method [9]. The field measurement method is usually the use of geophysical-related professional engineering equipment, the focus of the region of the rock body detection and monitoring, so as to determine whether there is the possibility of a rock explosion, but due to the current detection, monitoring equipment is still in the development stage, the reliability of the information collected is very limited. The theoretical analysis method is founded on the physical parameters of the rock mass, stress state and multiple triggering factors, it employs mathematical methods to predict the intensity of rockbursts, and representative methods include cloud model theory [10,11], fuzzy integrated discrimination [12,13], grey systems theory [14], topology theory [15], the ideal point method [16,17], the distance discrimination method [18,19], etc., but the solution methods are usually cumbersome and the generalisation ability of the established models is weak.

For the past few years, in light of the ongoing development of computer intelligence technology, machine-learning-related methods have better applications in the area of prediction, and other scholars have also conducted related research on rockburst prediction using machine learning techniques. Zhou [20] and others used PSO-SVMs to establish a long-term prediction model for rockbursts in underground caverns, and the results showed that the method has strong robustness; Dong [21] and others employed a Random Forest (RF) model of the underground rock project to forecast the likelihood of rockbursts and the associated rockburst grade and compared the outcomes with those obtained using the SVM and ANN methods; the findings indicate that the RF model is a viable approach for forecasting the intensity of rockbursts; Wu Shunchuan et al. [22], based on the principle of PCA-PNN to establish the rockburst intensity grading prediction model, showed that the outcomes of the assessment indicate that the model convergence speed is faster and the misjudgement rate is lower, and the prediction effect is more satisfactory. Pu Yuanyuan et al. [23] used the t-SNE method to downscale 246 real rockburst data set, and employed the clustering method to re-label the data, and then used the SVC to train the model, so the findings of the final prediction results demonstrate that the method has a strong generalisation ability, and the prediction accuracy is higher. Li et al. [24] performed rockburst prediction using seven machine learning methods based on 314 real cases of rockbursts, and the output showed that the integrated tree had superior prediction performance compared to other ML models; Xue [25] and colleagues developed a PSO-ELM rockburst intensity level prediction model, which they validated with 15 examples of rockbursts at the Jiangbian Hydropower Station; the model exhibited superior predictive capabilities, as evidenced by the results. Wang et al. [26] evaluated the learning ability of Boosting and Bagging algorithms based on rockburst instance data and also used cross-validation as a method of adjusting hyper-parameters and after analyses showed that the Bagging method was the most effective for the prediction of the rockburst intensity level. Ma et al. [27] established the Borderline-SMOTE 1-Adaboost model; with this method, to a certain extent, the problem of data imbalance is solved, and the prediction accuracy is higher after comparing with various models, and the model is applied in the forecasting of rockburst intensity in the Qinling water tunnel, which provides a reference for the early warning of rockburst disasters during the construction of deep buried tunnels. Ma et al. [28] proposed the LightGBM-TCN-RF rockburst intensity level prediction model, which makes the model able to predict the danger degree of rockburst characteristic variables, providing a reference for the field of rock explosion prediction and monitoring.

All of these research methods mentioned above have achieved some success in rockburst prediction research and have certain advantages for specific problems while enriching the disaster theory related to rockburst-prediction research. However, considering that each algorithm inevitably has some shortcomings in practical engineering applications, such as SVM, KNN and other models, the application of the model tends to be less robust, and the model's predicted performance in different regions, different geological condi-

tions, different climates and other conditions of the differences in the performance of the rockburst prediction of the differences exist, often with the input of the parameters of the different changes. With the progress of machine learning technology, more and more scholars are trying to integrate multiple models [29,30] to overcome the shortcomings of the poor generalisation ability of a single model and have achieved some results. In the field of rockburst prediction, several scholars have also made attempts to integrate models [31]; however, the generalisability of these models needs further research [32]. In this paper, to address the limitations of the current research methodology and enhance the model's predictive power and generalisability, an NGO-CNN-BiGRU-Attention rockburst prediction model based on intelligent optimisation algorithms and deep learning techniques is proposed, which is trained based on 287 sets of real rockburst data in different regions, under different geological conditions and at different times and is combined with RF, SVM, KNN, ELM, CNN, BiGRU and CNN-BiGRU-Attention models; the findings demonstrate that the NGO-CNN-BiGRU-Attention model exhibits greater prediction accuracy. In the end, the model is validated using eight sets of real rockburst cases in the Daxiangling Tunnel, and the results indicate that the model exhibits robust generalisation capacity and is suitable for engineering applications.

2. Methods

2.1. Random Forest (RF)

Random Forest (RF) is an integrated learning algorithm; this algorithm predicts by constructing multiple decision trees and combining them together (Figure 1) and therefore effectively avoids the shortcomings of a single decision tree that is prone to overfitting [33]. The following is the detailed algorithmic flow of the Random Forest algorithm:

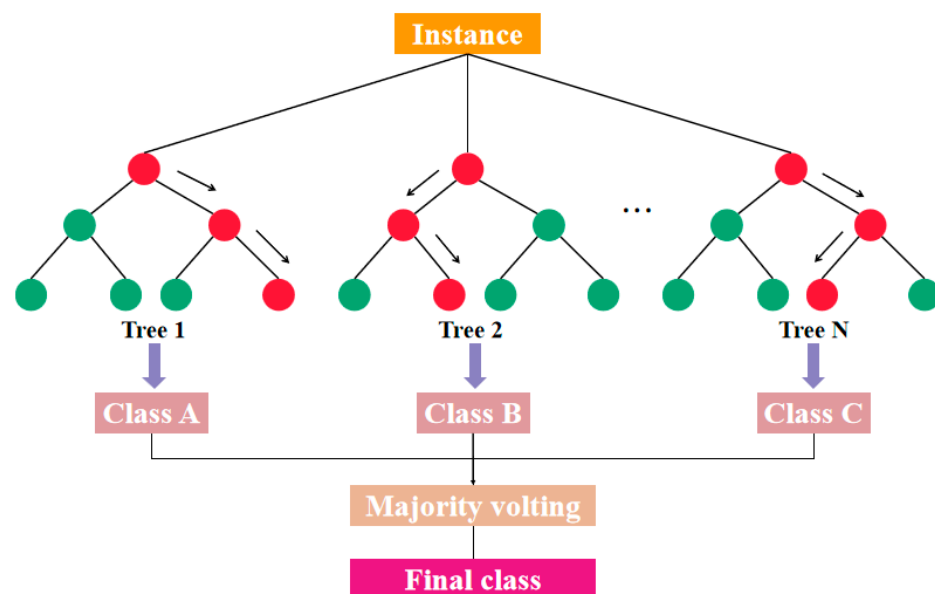


Figure 1. Schematic diagram of RF structure.

Set the number of trees N and the number of randomly selected features m in each tree, and prepare the training set data. Subsequently, the following steps are performed cyclically for each tree n : (1) random sampling; (2) random selection of features; and (3) construction of decision trees.

For the classification prediction problem of the rockburst intensity prediction, the RF model uses voting, where the majority vote determines the final classification result:

$$\hat{y}_{\text{final}} = \operatorname{argmax}_y \sum_{i=1}^N I(\hat{y} = y) \quad (1)$$

where N denotes the number of decision trees; the prediction for each tree is $\hat{y}_i$; and $I(\cdot)$ is an indicator function.

2.2. Support Vector Machine (SVM)

The Support Vector Machine (SVM) model is commonly used as an efficient supervised learning model to deal with linearly differentiable and nonlinear problems [34]. It transforms the initial data to a higher space by means of kernel function mapping, the structure of which is shown in Figure 2. Its kernel function expression is as follows:

$$k(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right) \quad (2)$$

with the introduction of the kernel function, the decision rule used by SVM for classification becomes the following:

$$g(x) = \text{sgn}\left(\sum_{i=1}^n a_i y_i (k(x_i, x) + b)\right) \quad (3)$$

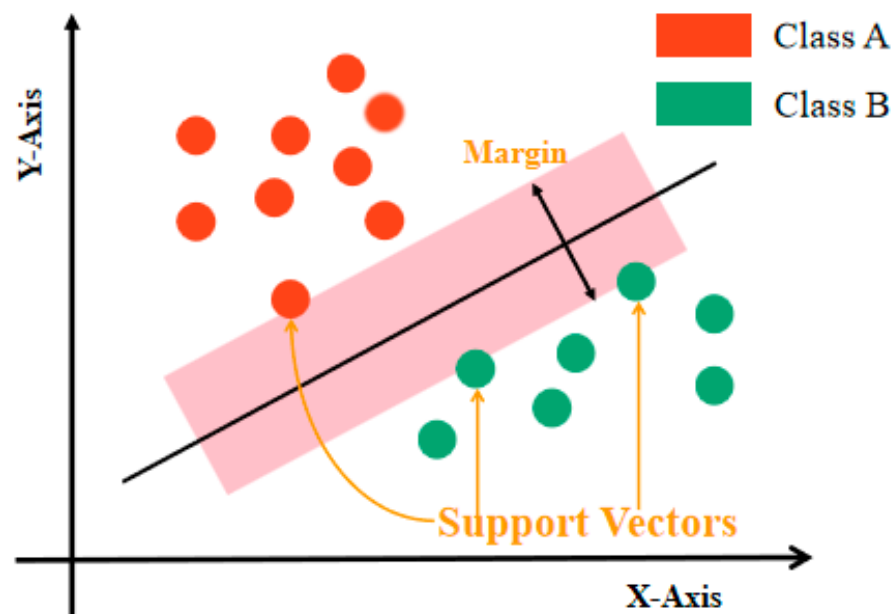


Figure 2. Schematic diagram of SVM structure.

A slack variable ξ_i is introduced to improve the fitness, while a penalty term C is added to constrain the misclassification cases:

$$\begin{cases} \min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ y_i(w^T x_i + b) \geq 1 - \xi_i, \xi_i \geq 0, \quad i = 1, 2, \dots, n \end{cases} \quad (4)$$

This is then converted to a dyadic problem:

$$\begin{cases} w_{\max} = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i=1, j=1}^n a_i a_j y_i y_j k(x_i, x_j), \\ \text{s.t. } \sum_{i=1}^n a_i y_i = 0, 0 \leq a_i \leq C, \quad i = 1, 2, \dots, n. \end{cases} \quad (5)$$

where a_i and a_j denote the Lagrange multipliers; y_i and y_j denote the training sample categories, and $y_i, y_j \in (0, 1)$; the penalty factor, C , is employed to quantify the intricacy inherent to machine learning models. If this parameter is set too high, the model is prone to

overfitting to the training data, which makes the SVM structure complex and reduces the computational efficiency; if the parameter is set too low, it will make the SVM not sufficient to capture the data features, resulting in underfitting.

2.3. KNN Algorithm

The KNN algorithm is a simple and powerful classification algorithm; for the samples to be classified, it searches its pattern space [35] and finds the nearest K training samples as its K-nearest neighbours according to the proximity of the samples and finally predicts the type of the samples to be classified according to the K-nearest neighbours by using some kind of voting strategy, and its structure is shown in Figure 3. The mathematical expressions of its distance metric and classification prediction method are shown in Equations (6) and (7), respectively:

$$\text{dist}(X_1, X_2) = \sqrt{\sum_{i=1}^n (x_{1i} - x_{2i})^2} \quad (6)$$

where X_1 and X_2 are the eigenvectors of the two data points, x_{1i} and x_{2i} are their i th eigenvalue, respectively, and n is the number of features.

$$\hat{y} = \underset{y}{\operatorname{argmax}} \sum_{i=1}^K I(y_i = y) \quad (7)$$

where $\hat{y}$ is the predicted category, y_i denotes the category of the i th nearest neighbour and $I(\cdot)$ represents the indicator function.

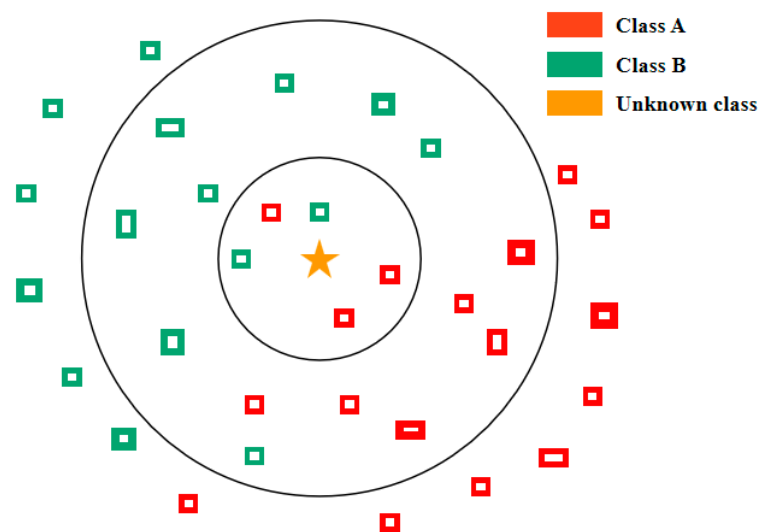


Figure 3. Schematic diagram of KNN structure.

2.4. Extreme Learning Machine (ELM)

The ELM classifier is a machine learning algorithm with better generalisation ability, which is simple to implement, has faster training speed and has better generalisation ability than traditional neural network methods [36], as shown in the mathematical expressions in Equations (8)–(10):

The output H of the implicit layer can be obtained by Equation (8):

$$H = g(XW + b) \quad (8)$$

W_o , the weights of the output layer of the ELM, can be calculated using (9):

$$W_o = (H^T H + \lambda I)^{-1} H^T Y \quad (9)$$

The prediction results can be obtained using Equation (10):

$$\hat{Y} = g(X'W + b)W_o \quad (10)$$

In the above equation, X denotes the input data matrix, W represents the weight from the input layer to the hidden layer, the symbol b represents the bias vector of the hidden layer, $g(\cdot)$ represents the activation function, H represents the output of the hidden layer, H^T represents the transpose of the output matrix H , Y represents the labelling matrix of the training data, λ denotes the regularisation parameter, I denotes the unit matrix, X' denotes the new input data, H' denotes the fixed output of the hidden layer, W_o stands for the output weight obtained by the learning and $\hat{Y}$ means the predicted output result.

2.5. Convolutional Neural Networks (CNNs)

CNN is a class of feedforward networks that include a deep structure and employ convolutional computation [37], which is capable of capturing local patterns and features in the data, and its main components include convolutional, pooling and fully connected layers. Considering the uniqueness of the channel, the input of the l th layer of the CNN is referred to as $x^l \in R^{H^l \times W^l}$, and the output obtained after x^l has gone through the operations of the l th layer is referred to as $x^{l+1} \in R^{H^{l+1} \times W^{l+1}}$.

Convolutional layers extract features from the input using convolutional kernels, convolutional neural networks generally require an increasing number of convolutional kernels as the layers get deeper and deeper and the operation of convolution can be represented by the following equation:

$$x_{i^{l+1},j^{l+1}}^{l+1} = \sum_{i=0}^H \sum_{j=0}^W w_{i,j} \times x_{i^l+j^l}^l \quad (11)$$

where $w_{i,j}$ is the weights learnt by the convolution kernel and (i^{l+1}, j^{l+1}) is the positional coordinate of the convolution result and which needs to satisfy the following expression:

$$\begin{aligned} 0 &\leq i^{l+1} < H^l - H + 1 \\ 0 &\leq j^{l+1} < W^l - W + 1 \end{aligned} \quad (12)$$

where H and W are the dimensions of the convolution kernel.

The role of the pooling layer is mainly feature dimensionality reduction, the prevention of overfitting, feature invariance, etc. The pooling layer includes two kinds of average pooling and maximum pooling. The expression for average pooling is as follows:

$$x_{i^{l+1},j^{l+1}}^{l+1} = \frac{1}{HW} \sum_{i=0}^H \sum_{j=0}^W x_{i^l \times H + i, j^l \times W + j}^l \quad (13)$$

The expression for maximum pooling is as follows:

$$x_{i^{l+1},j^{l+1}}^{l+1} = \max_{0 \leq i < H, 0 \leq j < W} x_{i^l \times H + i, j^l \times W + j}^l \quad (14)$$

The fully connected layer maps feature representations into sample labels. It is a feed-forward neural network:

$$x^{l+1} = f(wx^l + b) \quad (15)$$

In this context, f represents the activation function, w represents the weight matrix and b represents the bias matrix.

2.6. Bidirectional Gated Recirculation Units (BiGRUs)

BiGRU is an improved model based on GRU, which consists of a pair of forward-backward GRUs, where both GRUs simultaneously learn feature mapping from the forward

and reverse directions [38] and extract the sequence features, and then, the two output vectors from the forward and backward GRUs are concatenated to yield the final result. Since the model is able to capture the information from forward and backward directions, it is able to better understand the features and structures in the sequence and thus has higher model efficiency, and the related principle is shown in Figure 4.

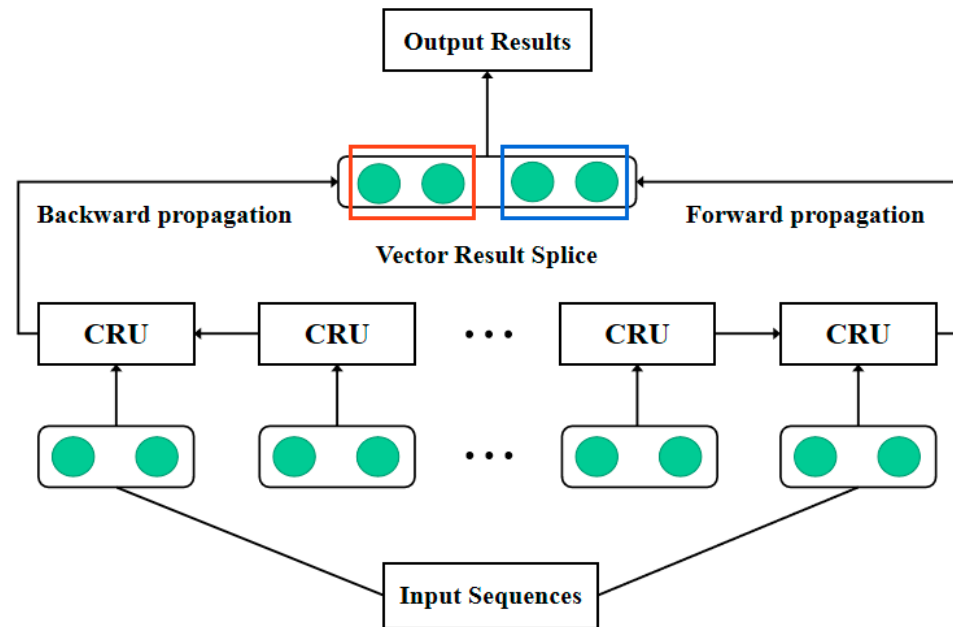


Figure 4. BiGRU Schematic.

When dealing with classification problems, BiGRU is usually used as a feature extractor to convert sequence data into fixed-length feature vectors, which are then fed into a classifier for classification.

- (1) Calculate the hidden state sequence of BiGRU. The hidden state sequences are computed separately for forward GRU and reverse GRU, and then, they are connected to form the hidden state sequences of BiGRU, the mathematical expression of which is shown below:

$$\begin{aligned} z_t &= \sigma(w_z * [h_{t-1}, x_t] + b_z) \\ r_t &= \sigma(w_r * [h_{t-1}, x_t] + b_r) \\ \tilde{h}_t &= \tanh(w_h * [r_t * h_{t-1}, x_t] + b) \\ h_t &= (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \end{aligned} \quad (16)$$

where z_t is an update gate, r_t is a hidden gate, x_t denotes an input at moment t , $\tilde{h}_t$ represents a candidate state at moment t , h_t stands for a hidden state value at moment t , w_z represents the weight matrices of the corresponding update and w_r represents reset gates, respectively, while b_z and b_r are their biases, respectively.

- (2) Feature Extraction Process. Pooling operations, such as average pooling or maximum pooling, are performed on the hidden layer, and the computational formulas are given in (13) and (14), to convert the variable-length hidden state sequence into a fixed-length feature vector.
- (3) Classification of feature vectors. Its mathematical expression is as follows:

$$\text{Output} = \text{softmax}(W_{fc} \cdot \text{Feature} + b_{fc}) \quad (17)$$

In the context of neural networks, the symbol W_{fc} represents the weight matrix of the fully connected layer, while b_{fc} denotes the bias vector, and softmax is the activation function of the classifier, which is used to compute the probability distribution for each category.

2.7. Attention Mechanism

The core principle Attention mechanism is a series of weighted summation operations to re-assign the feature weight coefficients according to the importance of the hidden state [39], which has the ability of data screening and key information extraction, and dynamically adjust the weights and thus has a strong advantage in dealing with nonlinear and complexity problems. In this study, the Attention layer is added after the BiGRU layer, because the Attention mechanism enables the model to identify key information about rockbursts and ignore irrelevant information, improving the prediction accuracy. The formula is as follows:

$$\begin{aligned}\lambda_t &= \tanh(w_t * h_t + b_t), \\ a_t &= \frac{\exp(\lambda_t)}{\sum_{k=1} \exp(\lambda_k)}, \\ v_t &= \sum_t a_t * h_t\end{aligned}\quad (18)$$

The hidden state after the BiGRU layer is designated as h_t , λ_t is the new state, a_t is the attention weights and v_t is the new feature vector finally obtained.

2.8. Northern Goshawk Optimisation Algorithm (NGO)

The Northern Goshawk Algorithm (NGO) is a population-based optimisation algorithm with the advantages of strong global search capability, fast convergence, good robustness and flexibility. The NGO algorithm simulates the behaviour of the Northern Goshawk throughout the process of the hunting process, including prey recognition and attack, chase and escape [40], and establishes mathematical models based on the different phases of the hunt, as shown in Figure 5.

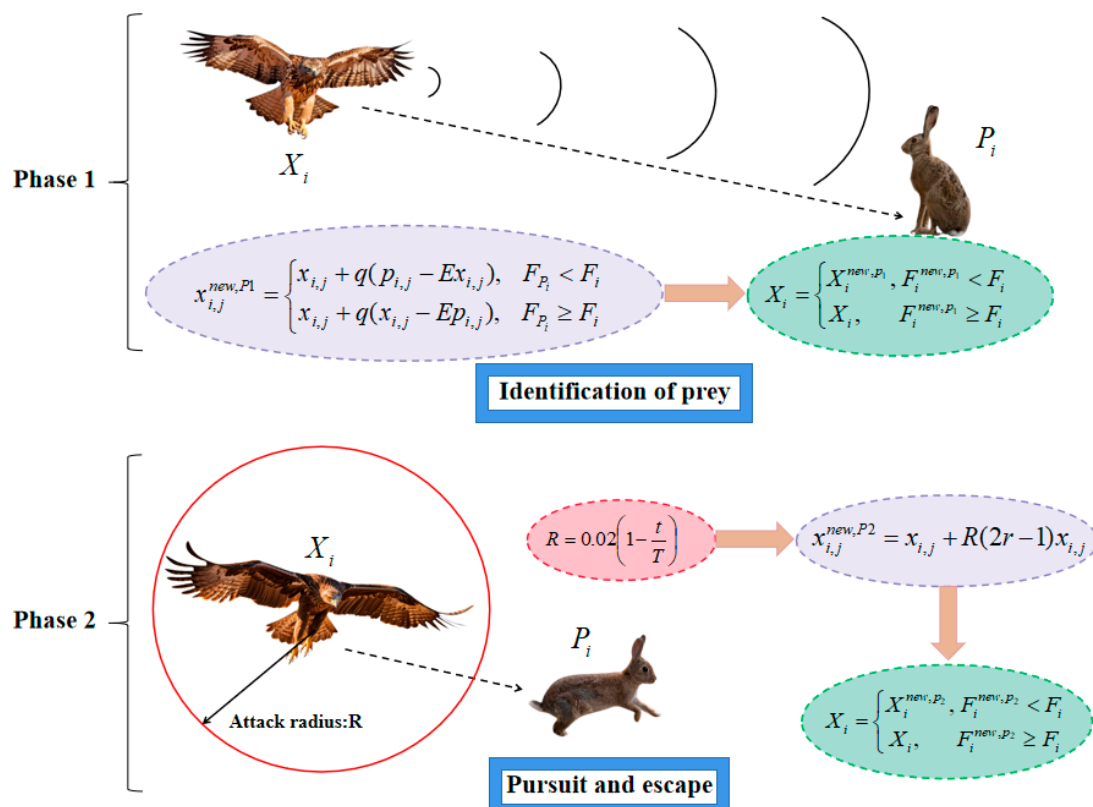


Figure 5. Schematic diagram of the optimisation algorithm for the Northern Eagle.

- (1) Prey identification phase. The northern goshawk will randomly select a prey and then quickly attack it. Its mathematical expression is shown below:

$$P_i = X_k, i = 1, 2, \dots, N, k = 1, 2, \dots, i - 1, i + 1, \dots, N \quad (19)$$

$$x_{i,j}^{new,P1} = \begin{cases} x_{i,j} + q(p_{i,j} - Ex_{i,j}), & F_{P_i} < F_i \\ x_{i,j} + q(x_{i,j} - Ep_{i,j}), & F_{P_i} \geq F_i \end{cases} \quad (20)$$

$$X_i = \begin{cases} X_i^{new,P1}, & F_i^{new,P1} < F_i \\ X_i, & F_i^{new,P1} \geq F_i \end{cases} \quad (21)$$

where P_i represents the prey position corresponding to the i th hawk, N stands for the sample length, k is a random integer that falls within the range $[1, N]$ and not equal to i ; F_{P_i} is the value of the objective function, F_i is the value of the objective function corresponding to the i th solution; X_i is the initial solution of the i th hawk, $x_{i,j}$ is the value of the i th hawk in the j th dimension and $x_{i,j}^{new,P1}$ denotes the value of the i th northern hawk in the j th dimension in the 1st hunting phase; $P1$ is the 1st hunting phase; $F_i^{new,P1}$ is the value of the objective function corresponding to the new solution; and the parameter q and the parameter E are random parameters, where q is a random number belonging to the intervals between the $[0, 1]$, and E has the value 1 or 2.

- (2) Pursuit and escape phase. After a northern goshawk has captured its prey, the prey will often try to escape. During the chase, given the high speeds typically exhibited by northern goshawks, they have the ability to capture prey in a variety of situations. Assuming that the hunt is at an attack position at radius R at this point, the mathematical expression is shown below:

$$x_{i,j}^{new,P2} = x_{i,j} + R(2r - 1)x_{i,j} \quad (22)$$

$$R = 0.02 \left(1 - \frac{t}{T} \right) \quad (23)$$

$$X_i = \begin{cases} X_i^{new,P2}, & F_i^{new,P2} < F_i \\ X_i, & F_i^{new,P2} \geq F_i \end{cases} \quad (24)$$

where $x_{i,j}^{new,P2}$ denotes the value of the j th dimension of the i th northern goshawk in the 2nd hunting stage, and $P2$ represents the 2nd hunting stage; t denotes the current number of iterations, T denotes the maximum number of iterations, respectively; R stands for the radius of the attack; and $F_i^{new,P2}$ denotes the value of the objective function for the stage.

2.9. CNN-BiGRU-Attention Hybrid Neural Network Model Based on NGO Optimisation Algorithm

While CNN, BiGRU, Attention mechanism [41] and NGO algorithms have been utilised in various domains, this study represents the first instance of integrating these methods for predicting rockburst intensity. Furthermore, specific enhancements and optimisations were implemented to address potential limitations.

Given the sensitivity of the original CNN-BiGRU-Attention model to initial parameter settings and its susceptibility to local optimal solutions during training, we propose a hybrid neural network model that combines CNN-BiGRU-Attention with the NGO optimisation algorithm.

This approach leverages CNN's feature extraction capabilities, BiGRU's proficiency in time series modelling and the Attention mechanism's ability to weigh important features with the aim of enhancing accuracy and robustness in rockburst intensity prediction. The Northern Goshawk Optimization Algorithm is employed for optimising hyperparameters (such as the size of CNN's convolutional kernel, number of GRU units and learning rate), as well as weight initialisation. With its strong global search capabilities and adaptability to complex prediction scenarios related to rockbursts, this optimisation algorithm effectively improves the classification prediction results' accuracy and reliability. The model structure is shown in Figure 6:

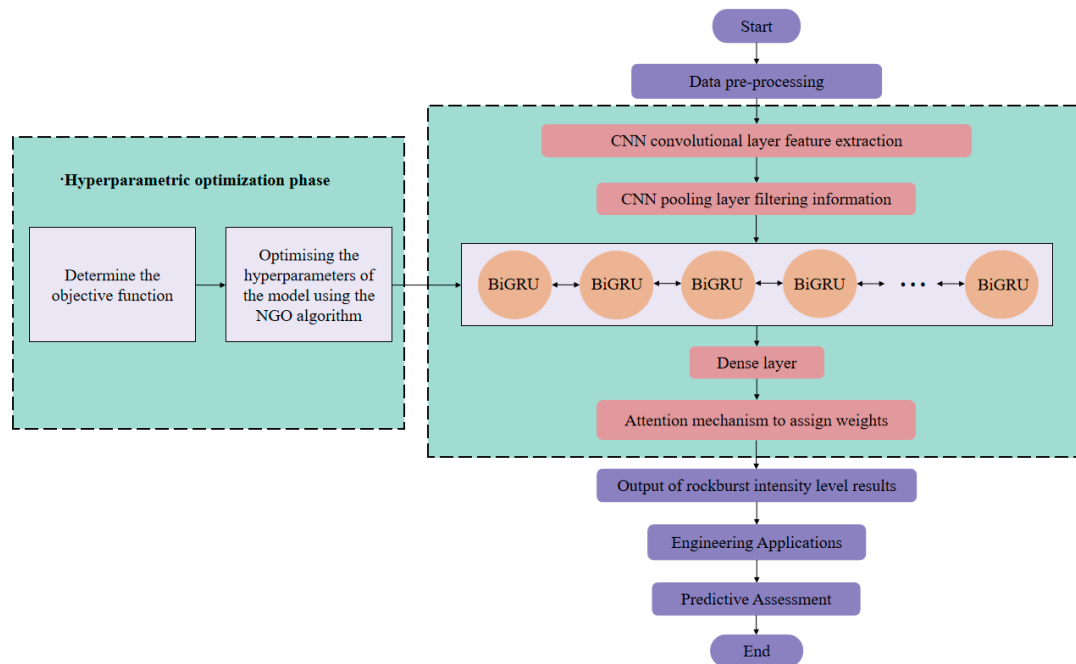


Figure 6. Structure of CNN-BiGRU-Attention model.

Specific operation steps:

- Data preprocessing: firstly, a sample library for data preprocessing and data partitioning must be established; the data set comprises 80% training data and 20% test data.
- Hyper-parameter optimisation: the objective function is defined as the prediction error, and the NGO optimisation algorithm is employed to optimise the hyper-parameters of the model.
- Data prediction: once the training package has been incorporated into the training model, the latter will undergo a hyperparameter search conducted by the NGO. Once this process has reached its conclusion, the model parameters will be established. At this juncture, the data will then be fed into the CNN-BiGRU-Attention for the prediction of rockburst intensity levels.
- Model evaluation: the accuracy, precision, recall and F1 score were employed as indicators to assess the precision of the model's predictions and to facilitate comparisons and analyses with other prediction models.

3. Establishment of the Database

3.1. Data Sources

The study extensively collected data from domestic and international rockburst-related engineering cases and used them to establish a rockburst database (Supplementary Materials). This database consists of 287 groups of data, including 8 groups of data from Li et al. [24], 15 groups of data from Xue et al. [25], 7 groups of data from Wu et al. [42], 12 groups of data from Pu et al. [23], 108 groups of data from Long et al. [43] and 137 groups of data from Afraei et al. [4], which are shown in Table 1.

In general, the selection of predictive indicators should meet the requirements of easy access, strong representativeness and strong characteristics, so this study selected the maximum tangential stress of the surrounding rock σ_θ (MPa), uniaxial compressive strength of rock σ_c (MPa), tensile strength of rock σ_t (MPa), the stress ratio of rock σ_θ/σ_c , rock brittleness ratio of rock σ_c/σ_t and elasticity and strain energy of the rock index W_{et} , six quantitative indexes as a characteristic parameter for analysing the situation of rockbursts. From the viewpoint of the geological structure of rockbursts, the maximum tangential stress σ_θ (MPa) of the enclosing rock body is usually one of the main driving factors for the occurrence of rockbursts. In contrast, although the main maximum (minimum) stress,

equivalent stress and bias stress can also reflect the stress state, the maximum tangential stress of the surrounding rock is more intuitive and direct in characterising the local stress concentration and damage risk of the surrounding rock body. Many existing studies and engineering practices have shown that the maximum tangential stress σ_θ (MPa) plays an important role in rockburst prediction and stability analysis of the surrounding rock [3]. Also, rock uniaxial compressive strength σ_c (MPa) and rock tensile strength σ_t (MPa) can better reflect the mechanical properties of the rock, in the actual situation of rockbursts have been documented and can be used as a prediction and evaluation of rockburst indicators; from the energy point of view, the rockburst is the energy of elastic stress to a certain degree of energy accumulation after the rapid release of the results. Therefore, the rock elastic strain energy index W_{et} was selected as a rockburst prediction index. In addition, the stress ratio σ_θ/σ_c and rock brittleness ratio σ_c/σ_t are also proposed by many researchers to predict the selection of rockburst strength indicators [44–46].

Table 1. Distribution of rockburst labelling in the database [4,23–25,42,43].

References	Number of Cases	Rockburst Label			
		None	Light	Moderate	Strong
Li et al. (2022)	8	0	3	3	2
Xue et al. (2020)	15	3	5	4	3
Wu et al. (2019)	7	0	1	5	1
Pu et al. (2019)	12	0	1	11	0
Long et al. (2024)	108	19	32	43	14
Afraei et al. (2019)	137	25	36	55	21
Total	287	47	78	121	41

In this paper, the rockburst rating criteria selected are a posteriori derived from the rockburst intensity level proposed by Zhou et al. [20] (Table 2), which categorises the intensity level into four distinct categories: None, Light, Moderate and Strong. Table 3 demonstrates the calculation of the statistical analysis metrics for each indicator (i.e., standard deviation, kurtosis, maximum, mean, median and range).

Table 2. Generic classification criteria for rock blast intensity.

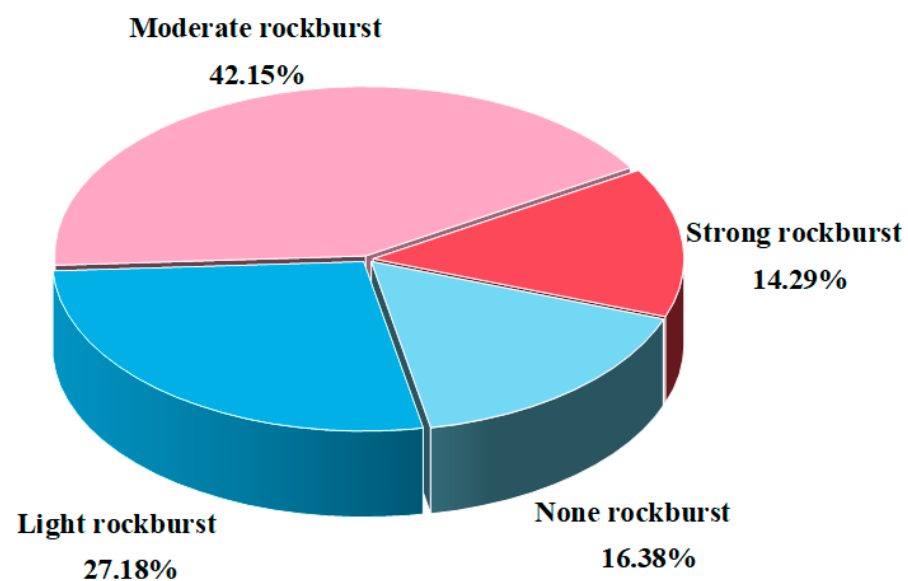
Rockburst Label	Sign	Failure Characteristics
None rockburst	I	There is no sound of rock bursting or rock falling.
Light rockburst	II	The rock surrounding the area displays signs of spalling, cracking or striping. There is no evidence of ejection, and the sound is weak.
Moderate rockburst	III	The rocks surrounding the area exhibit significant deformation and fracturing, resulting in loose rock chips and sudden destruction. This is often accompanied by a crunchy squeaking sound, which is common in local caverns within the surrounding rock.
Strong rockburst	IV	The rocks surrounding the tunnel are severely fractured and suddenly propelled into it, accompanied by strong bursts, roars, air jets and other storm phenomena. This causes rapid expansion into the deep surrounding rocks.

Table 3. Calculation of statistical analysis indicators for each indicator.

Parameters	Data Indicators						
	STD	Kurt	Max	Min	Mean	Median	Range
σ_θ (MPa)	32.65	0.85	167.20	2.60	43.61	49.50	164.60
σ_c (MPa)	53.63	1.08	306.58	15.00	100.54	112.50	291.58
σ_t (MPa)	4.47	1.50	22.60	0.38	5.16	5.20	22.22
σ_θ/σ_c	0.56	43.46	5.26	0.05	0.41	0.44	5.21
σ_c/σ_t	12.12	3.42	80.00	0.15	18.98	21.18	79.85
W_{et}	2.10	−0.62	10.57	0.85	3.93	4.63	9.72

3.2. Data Description and Analysis

The rockburst case datasets used in this study are all from reliable databases and can be verified; the distribution of the rockburst class is shown in Figure 7, in which class I samples are the least, for 47 cases, class II samples are 78 cases; class III samples are the most, for 121 cases; and class IV samples are 41 cases. The fiddle diagram of the initial dataset is shown in Figure 8. The fiddle diagram is a combination of densities, and its shape can characterise the overall distribution of the data; a wider fiddle diagram indicates that the data are more evenly distributed, and a narrower fiddle diagram indicates that the data are more concentrated. The box line part in the middle of the violin diagram indicates the median and interquartile range of the data, and the concentration of the scatter represents the density of a certain value. By observing Figures 7 and 8, it can be found that there are some category imbalances or sampling biases in this dataset, and the appearance of these outliers may be related to the special working condition samples collected [47].

**Figure 7.** Distribution of rockburst grades.

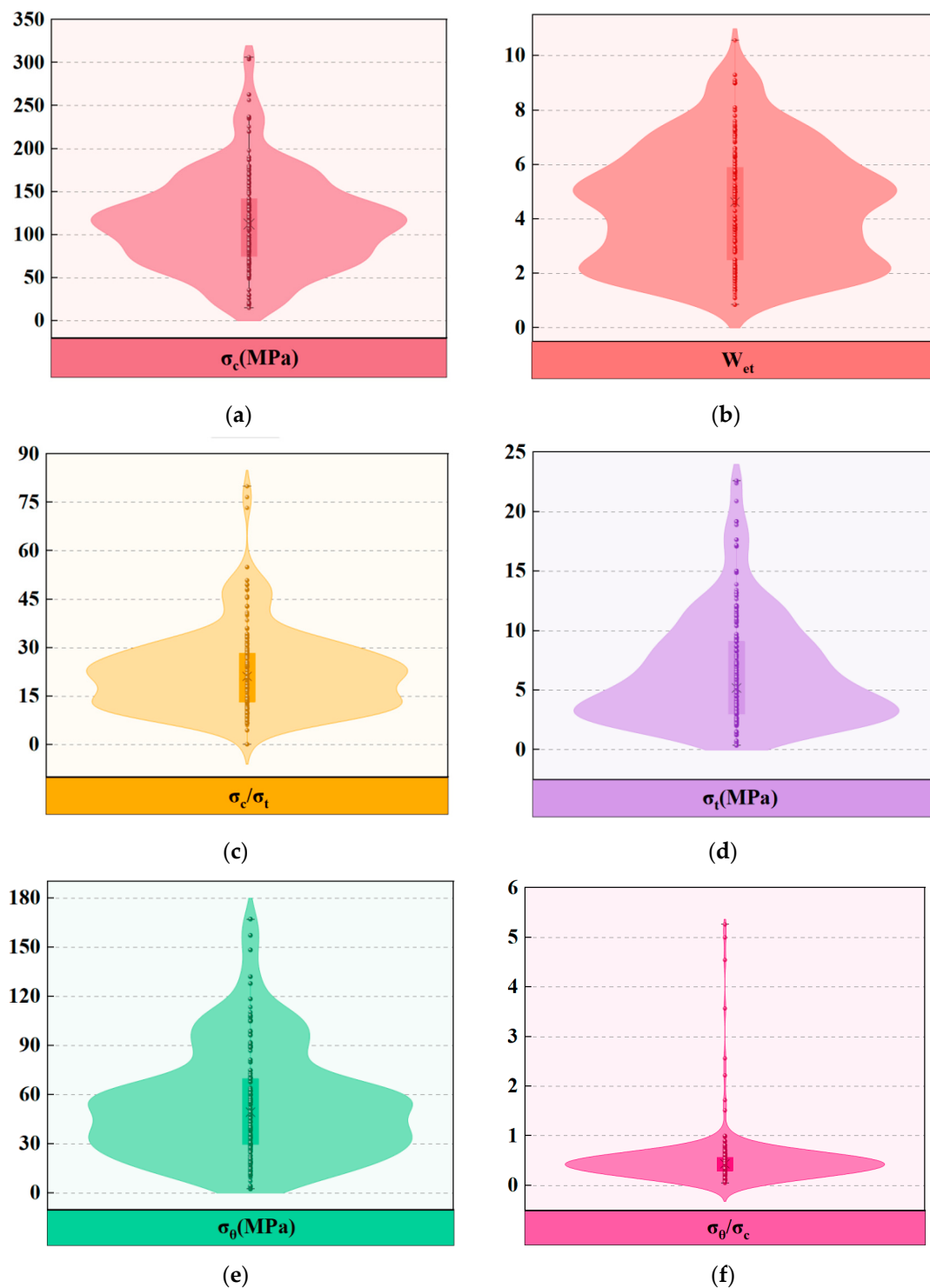


Figure 8. Fiddle map of rockburst data distribution: (a) the maximum tangential stress of the surrounding rock; (b) elasticity and strain energy of the rock index; (c) rock brittleness ratio of rock; (d) tensile strength of rock; (e) the maximum tangential stress; (f) the stress ratio of rock.

3.3. Parameter Correlation Analysis

In order to assess the correlation among the six characteristic parameters of rockbursts, a correlation analysis can be based on the Pearson correlation coefficient method [48] for correlation analysis of the above sample data of rockburst prediction; the formula is shown in Equation (25), and the results are shown in Figure 9. The magnitude of the absolute value of the correlation coefficient can be employed to demonstrate the strength of the correlation.

The shape of the ellipse indicates the strength of the correlation; the more rounded the ellipse, the weaker the correlation, and the flatter the ellipse, the stronger the correlation.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (25)$$

where: $\bar{x}$ and $\bar{y}$ are the mean values of n test values, respectively; $r \in [-1, 1]$.

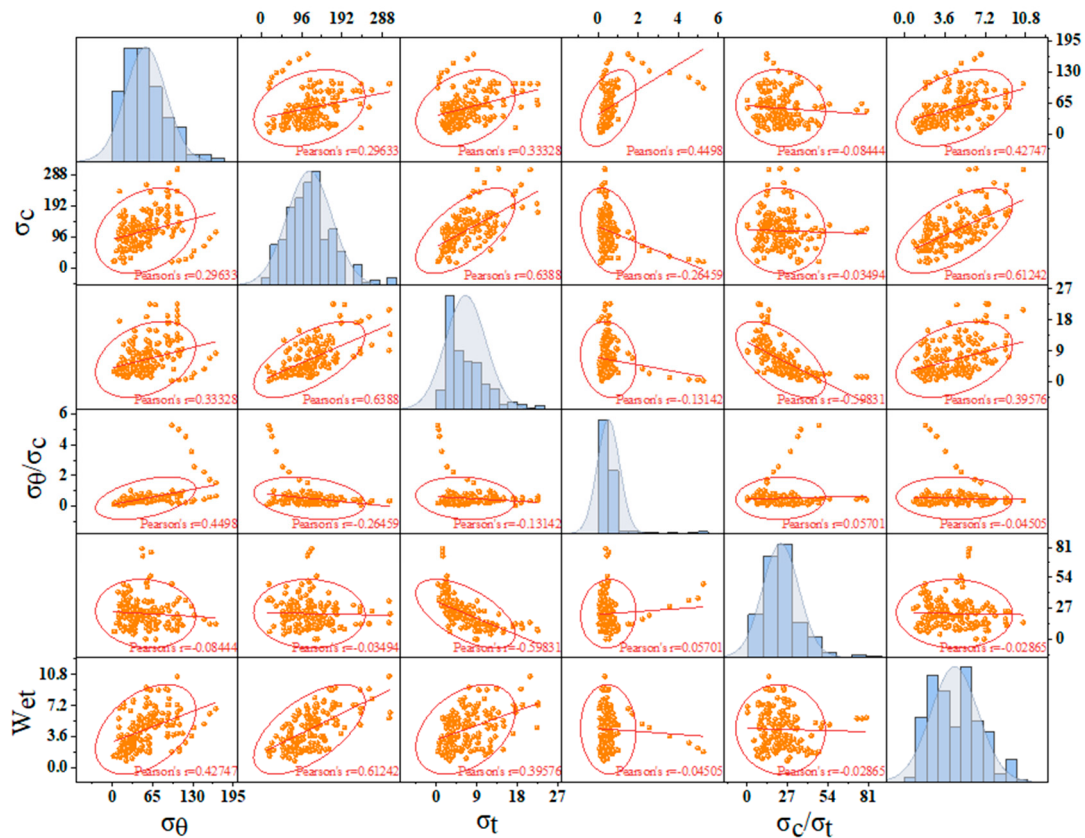


Figure 9. Correlation of rockburst samples.

The diagonal position in Figure 9 is the histogram of the normal distribution of the original dataset; in the normal distribution, the mean is the centre point of the data, and the standard deviation determines the width of the curve [49] and can be a measure of the degree of dispersion between the data points and the mean. A wider histogram indicates that the standard deviation is larger, and otherwise, it is the opposite. As can be seen from the figure, most of the data are more concentrated, and the number of data along both sides of the mean gradually decreases, while the normal histogram of W_{et} is more abnormal, which indicates that there are certain outliers in the sample and also reflects the existence of a certain degree of uncertainty in W_{et} .

A correlation coefficient with an absolute value more than 0.5 implies a high degree of correlation between the two features [50,51]. From Figure 9, we can see that the correlation between σ_c and σ_t reaches 0.6388, which indicates that the correlation between the two is strong, and it has been pointed out in the related literature [52] that there may be some kind of connection between the two. The correlation between σ_c and W_{et} is as high as 0.6124, which explains to a certain extent that when the rock is subjected to high stress, a large amount of elastic strain energy will be accumulated in its interior, and if the rock's compressive strength is lower and it cannot effectively resist the action of external stress,

then the rock will be prone to occur. If the rock's compressive strength is low and can not effectively resist the role of external stress, then the rock is prone to damage [53], resulting in an abrupt release of elastic strain energy, which may cause rockburst. σ_t and σ_c/σ_t have a correlation of -0.5983 , with a strong negative correlation, which indicates that the brittle characteristics of the rock are closely related to rockbursts. The greater the brittleness of the rockbursts, the higher the tendency to be higher [54,55]; in the presence of high stress, the rocks with lower tensile strength and higher brittleness are more likely to be damaged and prone to rockbursts. The correlation between the other features is weak, indicating a strong independence between the features. In summary, the consequences of the data correlation analysis are more consistent with the engineering reality, which further verifies the reliability and authenticity of the dataset.

4. Prediction Results and Analysis

The eight rockburst intensity level prediction models in this study have differences, so machine learning models are employed to predict the rockburst intensity level and to compare the effectiveness and performance of the models.

4.1. Model Prediction Results

By inputting the test set into the rockburst intensity prediction classifier, the prediction consequences of the classifiers of different models based on the test set can be obtained. The confusion matrix, a table employed to assess the efficiency of a classification model, is able to present the model's prediction results in the context of the correspondence between the real categories and the model's predicted categories. Figure 10 illustrates the confusion matrix corresponding to the classification prediction results of each model.

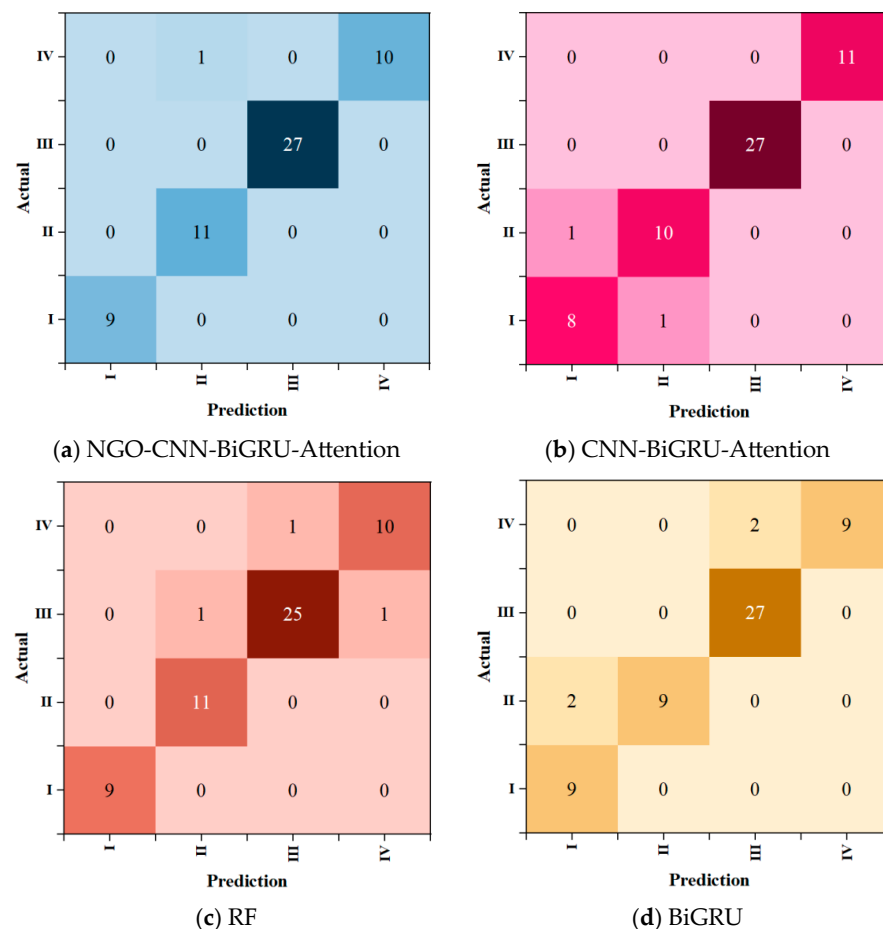


Figure 10. Cont.

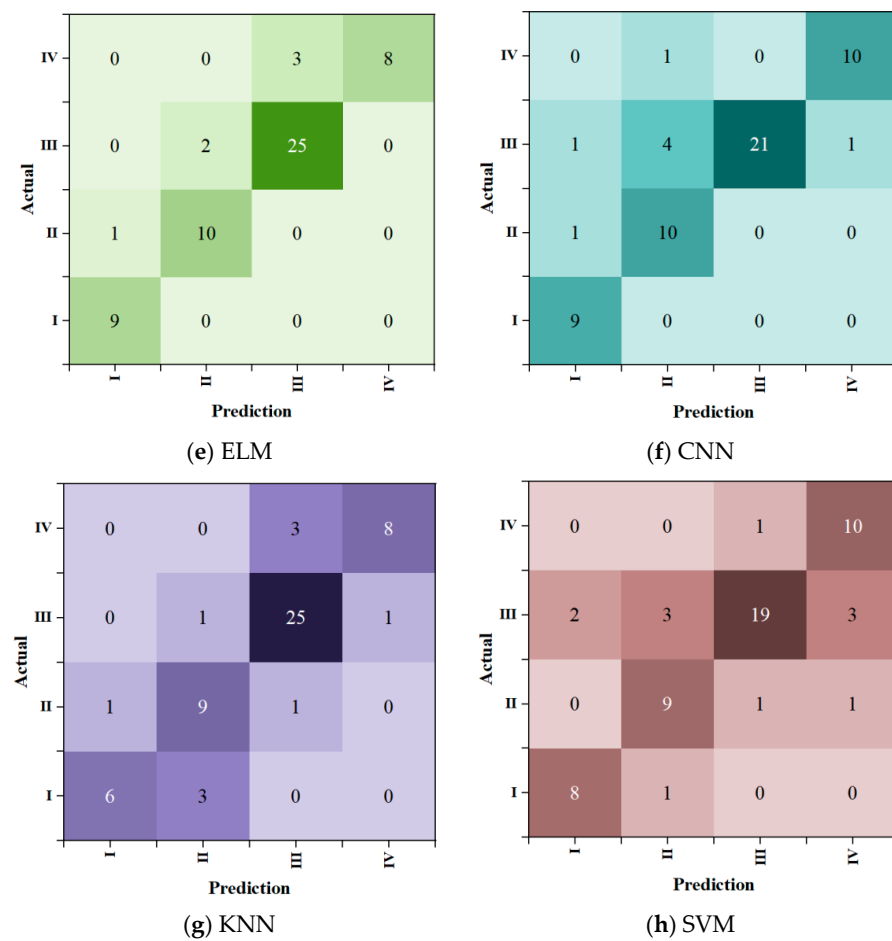


Figure 10. Confusion matrix diagram.

4.2. Evaluation and Comparison of Model Prediction Performance

In order to assess the prediction efficiency and generalisation ability of different rockburst intelligent prediction models, the study selected the Accuracy, Precision, Recall and F1 Score to evaluate the evaluation performance of each method (26)–(29) [28], and the calculated results are shown in Table 4.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (26)$$

$$Precision = \frac{TP}{TP + FP} \quad (27)$$

$$Recall = \frac{TP}{TP + FN} \quad (28)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (29)$$

TP represents true positive, TN true negative, FP false positive and FN false negative.

Table 4 demonstrates the evaluation results of the different models, and it can be seen that the best prediction performance is the NGO-CNN-BiGRU-Attention model with an accuracy of 98.28%, which has a precision of 97.92%, a recall of 97.73% and an F1 score of 97.82%, which are higher than the other models. The order is CNN-BiGRU-Attention, RF, BiGRU, ELM, CNN, KNN and SVM, with accuracies of 96.55%, 94.83%, 93.10%, 89.66%, 86.21%, 82.76% and 79.31%, respectively. It is not difficult to find that there is a difference in the performance of traditional machine learning methods in the rockburst intensity level prediction, which may be due to a single model with poor robustness. Since this study

makes the model better able to capture the complex patterns of the data by combining CNN and BiGRU and introduces the attention mechanism, the model's attention and prediction accuracy can be improved. In addition, the NGO algorithm facilitates the model to better seize the global information. So, although the prediction performance of the CNN-BiGRU-Attention rockburst prediction model shows good performance, with a precision rate of 94.95%, a recall rate of 94.95% and an F1 score of 94.95%, there is still a certain gap compared with the NGO-CNN-BiGRU-Attention model, which also reflects the characteristics of the NGO algorithm's automatic optimisation of the search effect being good and having strong robustness.

Table 4. Evaluation of the performance of the models.

Model	Accuracy	Precision	Recall	F1 Score
NGO-CNN-BiGRU-Attention	98.28%	97.92%	97.73%	97.82%
CNN-BiGRU-Attention	96.55%	94.95%	94.95%	94.95%
RF	94.83%	94.68%	95.88%	95.28%
BiGRU	93.10%	93.73%	90.91%	92.30%
ELM	89.66%	90.65%	89.06%	89.85%
CNN	86.21%	84.85%	89.90%	87.30%
KNN	82.76%	82.51%	78.45%	80.43%
SVM	79.31%	77.78%	83.00%	80.31%

In conclusion, by comparing the predictive capabilities of various models, it is not difficult to find that the NGO-CNN-BiGRU-Attention model, as a superior model, makes the overall model have a stronger characterisation ability by combining a variety of deep learning techniques and is able to forecast the intensity level of rockbursts with a high accuracy rate and provides certain guiding significance for the research of rockburst disaster prevention and control.

4.3. Feature Importance Analysis Based on SHAP Methodology

SHAP is a method employed to elucidate the rationale behind machine learning predictive models [56], which is particularly applicable to black-box models. It determines the extent to which each input variable contributes to the final prediction of the model by using the concept of Shapley values. For the rockburst intensity level prediction problem, the SHAP method can be employed to analyse the relative importance of sample features. The specific results are presented in Figure 11.

The degree of feature importance can be used to demonstrate the extent to which each feature affects the prediction results. As illustrated in Figure 11, each rectangle represents the corresponding feature importance. The more important the feature, the taller the rectangle; the less important the feature, the shorter the rectangle. It becomes apparent that the characteristic importance of the elastic strain energy index W_{et} is the largest, 0.21, indicating that the feature has the greatest impact on the intensity level of the rockburst; the stress ratio σ_θ/σ_c feature importance is as high as 0.201, indicating that the feature has a greater impact on the intensity level of the rockburst, and the feature and the elastic strain energy index W_{et} are close to the characteristic importance of the feature; it can be speculated that there may be a certain degree of correlation between these two features. In addition, the maximum tangential stress σ_θ (MPa), rock uniaxial compressive strength σ_c (MPa), rock tensile strength σ_t (MPa) and rock brittleness ratio σ_c/σ_t characteristic importance are lower than the first two, but their impact on the intensity level of rockbursts cannot be ignored. Therefore, these features should be taken into account in engineering applications and research to make the prediction of rockburst intensity levels more reliable.

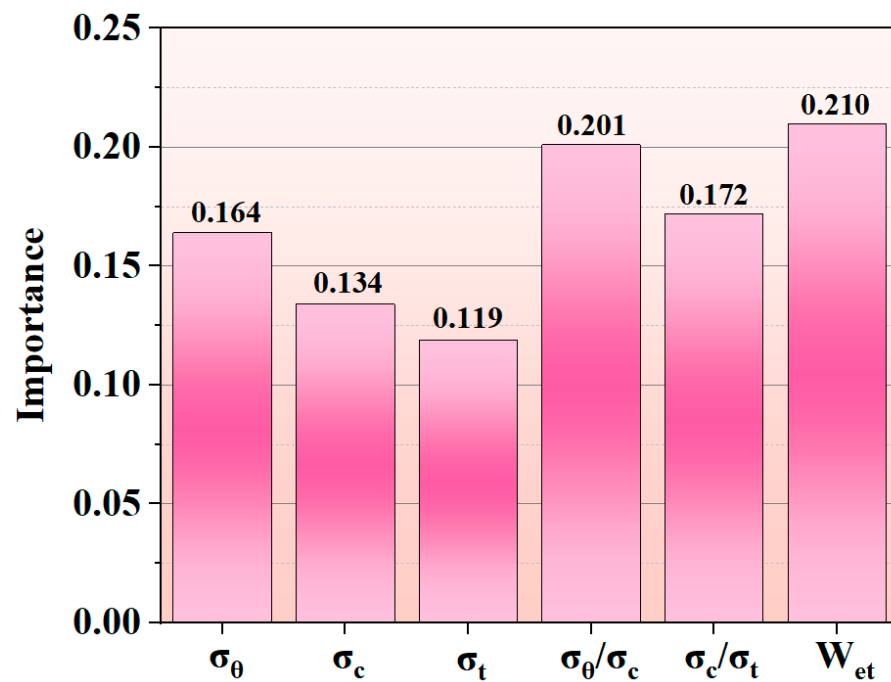


Figure 11. Schematic diagram of feature importance.

5. Engineering Validation

The Daxiangling Tunnel is located in the south of Ya'an City, Sichuan Province, which is the first longest tunnel under construction in Southwest China, with a maximum burial depth of 1648 m. In this study, eight sets of data were selected from the rockburst data of the Daxiangling Tunnel [57] to assess the rockburst hazard and were predicted using the NGO-CNN-BiGRU-Attention rockburst intensity prediction model. Table 5 illustrates the findings of the predictions, and all rockburst cases were predicted correctly, indicating that the NGO-CNN-BiGRU-Attention model shows strong generalisation ability and engineering practicability, which corroborates the previous conclusions.

Table 5. Prediction results of rock explosion cases in Daxiangling Tunnel.

No.	σ_θ	σ_c	σ_t	σ_θ/σ_c	σ_c/σ_t	W_{et}	Actual	Prediction
1	39.40	65.20	2.30	0.60	28.35	3.40	III	III
2	38.20	71.40	3.40	0.54	21.00	3.60	III	III
3	29.71	16.00	2.70	0.26	42.96	3.70	II	II
4	29.10	94.00	2.60	0.31	36.15	3.20	II	II
5	58.20	83.60	2.60	0.70	32.15	5.90	IV	IV
6	57.20	80.60	2.50	0.71	32.24	5.50	IV	IV
7	27.80	90.00	2.10	0.31	42.86	1.80	I	I
8	25.70	59.70	1.30	0.43	45.92	1.70	I	I

In order to further prove the effectiveness of the NGO-CNN-BiGRU-Attention model, the prediction results of this model were compared with those of other models (see Table 6), and the results showed that the accuracy of the NGO-CNN-BiGRU-Attention model was the highest, thus indicating that the model is an effective rockburst prediction method.

Table 6. Comparison of the accuracy of the models.

Model	Accuracy
NGO-CNN-BiGRU-Attention	1
CNN-BiGRU-Attention	1
RF	0.875
BiGRU	0.875
ELM	0.875
CNN	0.75
KNN	0.75
SVM	0.625

The above results show that the NGO-CNN-BiGRU-Attention model has strong prediction performance and generalisation ability in the area of rockburst intensity prediction, but to some extent, the model still has certain deficiencies. Due to the different selections of the sample library, the generalisation ability of the trained model varies; the contribution of different predictors to the rockburst intensity level is different, and the performance of the model will be influenced to some degree by the different predictors selected, so it is necessary to further explore this effect in the research of rockburst prediction.

6. Conclusions

In the present study, for the rockburst intensity level prediction problem, eight machine learning classification prediction models were established to predict the rockburst intensity level, based on six evaluation indexes, and compared and analysed with other machine learning prediction models. Furthermore, the study also utilised the SHAP method to conduct a thorough analysis of the models to explain the affect or importance of each feature on the prediction of the rockburst intensity level, and the final conclusions were obtained as follows.

The study collected 287 sets of real rockburst data from different regions, geological conditions and times to constitute a sample library for the machine learning prediction model, which included six characteristic parameters, such as the maximum tangential stress σ_θ (MPa), uniaxial compressive strength of rock σ_c (MPa), the tensile strength of rock σ_t (MPa), stress ratio σ_θ/σ_c , brittleness ratio of rock σ_c/σ_t and elastic strain energy index W_{et} , and it analysed the correlation between the different evaluation indexes using the Pearson correlation coefficient method.

The study constructed eight machine learning classification prediction models based on machine learning techniques and evaluated the prediction performance of each model using calculations and found that the NGO-CNN-BiGRU-Attention model achieved an accuracy of 0.98, which is the highest among all models tested, and the other models' accuracies were, in order, 0.96 for CNN-BiGRU-Attention, 0.94 for RF, 0.93 for BiGRU, 0.89 for ELM, 0.86 for CNN, 0.82 for KNN, and 0.79 for SVM. Since the NGO algorithm enables the model to better capture the global information, although the CNN-BiGRU-Attention rockburst prediction model shows good prediction performance, there is a certain level of accuracy compared with the NGO-CNN-BiGRU-Attention model; this reflects the good automatic optimisation and robustness of the NGO algorithm and also shows that the NGO-CNN-BiGRU-Attention model has good performance in the area of rockburst intensity level prediction, which can provide some guidance for related research.

For the prediction results of different models, the SHAP method was employed to quantify the impact of different evaluation indexes on the intensity level of rockbursts, and the results show that the elastic strain energy index W_{et} and stress ratio σ_θ/σ_c have the greatest influence on the intensity level of rockbursts, and their characteristic importance is as high as 0.21 and 0.201, respectively. For the other characteristics, although their importance is not as high, they have a non-negligible influence on the intensity level of rockbursts. The other features are not as important, but their influence on the intensity level of the rockburst is not negligible. Therefore, these features should be taken into account in

engineering applications and research to make the prediction of rockburst intensity levels more reliable.

The NGO-CNN-BiGRU-Attention model is applied to eight sets of real rockburst cases in the Daxiangling Tunnel, and the findings indicate that the model has excellent prediction performance, strong generalisation ability and good engineering practicability, which can be applied to the prediction of the rockburst intensity level in real engineering.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app14135719/s1>, Database of rockburst samples.

Author Contributions: Conceptualisation, H.L. and T.M.; methodology, K.P.; software, H.L.; validation, T.M., K.L. and H.L.; formal analysis, T.M.; investigation, X.H.; resources, H.L.; data curation, T.M.; writing—original draft preparation, H.L.; writing—review and editing, S.X.; visualisation, T.M.; supervision, Y.L.; project administration, K.P.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key Research and Development Program of China—2023 Key Special Project (No. 2023YFC2907400), the National Natural Science Foundation of China (No.52104109) and the Natural Science Foundation of Hunan Province, China (No.2022JJ40602).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

RF	Random Forest
SVM	Support Vector Machine
KNN	K-Nearest Neighbors algorithm
ELM	Extreme Learning Machine
CNN	Convolutional Neural Networks
BiGR	Bidirectional gated recirculation units
NGO	Northern Goshawk Optimisation Algorithm
σ_θ	the maximum tangential stress of the surrounding rock
σ_c	uniaxial compressive strength of rock
σ_t	tensile strength of rock
σ_θ / σ_c	the stress ratio of rock
σ_c / σ_t	rock brittleness ratio of rock
W_{et}	elasticity and strain energy of the rock index

References

1. Wang, C.; Xu, J.; Li, Y.; Wang, T.; Wang, Q. Optimization of BP Neural Network Model for Rockburst Prediction under Multiple Influence Factors. *Appl. Sci.* **2023**, *13*, 2741. [\[CrossRef\]](#)
2. Qian, Q. Definition, mechanism, classification and quantitative forecast model for rockburst and pressure bump. *Rock Soil Mech.* **2014**, *35*, 1–6.
3. Xu, G.; Li, K.; Li, M.; Qin, Q.; Yue, R. Rockburst Intensity Level Prediction Method Based on FA-SSA-PNN Model. *Energies* **2022**, *15*, 5016. [\[CrossRef\]](#)
4. Afraei, S.; Shahriar, K.; Madani, H.S. Developing intelligent classification models for rock burst prediction after recognizing significant predictor variables, Section 1: Literature review and data preprocessing procedure. *Tunn. Undergr. Space Technol. Inc. Trenchless Technol. Res.* **2019**, *83*, 324–353. [\[CrossRef\]](#)
5. Li, N.; Feng, X.; Jimenez, R. Predicting rock burst hazard with incomplete data using Bayesian networks. *Tunn. Undergr. Space Technol. Inc. Trenchless Technol. Res.* **2017**, *61*, 61–70. [\[CrossRef\]](#)
6. Li, M.L.; Li, K.G.; Qin, Q.C.; Wu, S.; Liu, Y.; Liu, B. Discussion and selection of machine learning algorithm model for rockburst intensity grade prediction. *Chin. J. Rock Mech. Eng.* **2021**, *40*, 2806–2816.

7. Qiao, L.; Dong, J.; Liu, J.; Chen, L. Review on the Study Progress of Rockburst Mechanism and Prediction in Underground Metal Mines in China. *Met. Mine* **2023**, 14–28.
8. Li, X.B.; Gong, F.Q.; Wang, S.F.; Li, D.Y.; Tao, M.; Zhou, J.; Huang, L.Q.; Ma, C.D.; Du, K.; Feng, F. Coupled static-dynamic loading mechanical mechanism and dynamic criterion of rockburst in deep hard rock mines. *Chin. J. Rock Mech. Eng.* **2019**, *38*, 708–723.
9. Liu, W.; Li, J.; Li, L. Review of research status on rockburst. *Gold* **2010**, *31*, 26–28.
10. Wang, Y.; Jing, H.; Zhang, Q.; Wei, L.; Xu, Z. A normal cloud model-based study of grading prediction of rockburst intensity in deep underground engineering. *Rock Soil Mech.* **2015**, *36*, 1189–1194.
11. Yang, J.; Wang, G.Y.; Liu, Q.; Guo, Y.; Liu, Y.; Gan, W.; Liu, Y. Retrospect and Prospect of Research of Normal Cloud Model. *Chin. J. Comput.* **2018**, *41*, 724–744.
12. Wu, L. The Study on Fuzzy Comprehensive Evaluation and Its Application. Master's Thesis, Taiyuan University of Technology, Taiyuan, China, 2006.
13. Yang, J.; Li, X.; Zhou, Z.; Lin, Y. A Fuzzy Assessment Method of Rock-burst Prediction Based on Rough Set Theory. *Met. Mine* **2010**, 26–29.
14. Liu, S. Emergence and Development of Grey System Theory and Its Forward Trends. *J. Nanjing Univ. Aeronaut. Astronaut.* **2004**, *36*, 267–272.
15. Cai, W.; Yang, C. Basic theory and methodology on Extenics. *Chin. Sci. Bull.* **2013**, *58*, 1190–1199.
16. Wang, Y.C.; Shang, Y.Q.; Sun, H.Y.; Yan, X.S. Research and application of rockburst intensity prediction model based on entropy coefficient and ideal point method. *J. China Coal Soc.* **2010**, *35*, 218–221.
17. Jia, Y.P.; Lv, Q.; Shang, Y.Q.; Du, L.L.; Zhi, M.M. Rockburst prediction based on rough set and ideal point method. *J. Zhejiang Univ. Eng. Sci.* **2014**, *48*, 498–503.
18. Luo, L.; Cao, P. Model of weighted distance discriminant analysis and application for deep roadway. *J. Cent. South Univ. Sci. Technol.* **2012**, *43*, 3971–3975.
19. Wang, J.; Li, X.; Yang, J. A Weighted Mahalanobis Distance Discriminant Analysis for Predicting Rock-Burst in Deep Hard Rocks. *J. Min. Saf. Eng.* **2011**, *28*, 395–400.
20. Zhou, J.; Li, X.; Shi, X. Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. *Saf. Sci.* **2012**, *50*, 629–644. [[CrossRef](#)]
21. Dong, L.; Li, X.; Peng, K. Prediction of rockburst classification using Random Forest. *Trans. Nonferrous Met. Soc. China* **2013**, *23*, 472–477. [[CrossRef](#)]
22. Wu, S.; Zhang, C.; Cheng, Z. Prediction of intensity classification of rockburst based on PCA-PNN principle. *J. China Coal Soc.* **2019**, *44*, 2767–2776.
23. Pu, Y.; Apel, B.D.; Xu, H. Rockburst prediction in kimberlite with unsupervised learning method and support vector classifier. *Tunn. Undergr. Space Technol. Inc. Trenchless Technol. Res.* **2019**, *90*, 12–18. [[CrossRef](#)]
24. Li, D.; Liu, Z.; Armaghani, D.J.; Xiao, P.; Zhou, J. Novel ensemble intelligence methodologies for rockburst assessment in complex and variable environments. *Sci. Rep.* **2022**, *12*, 1844. [[CrossRef](#)] [[PubMed](#)]
25. Xue, Y.; Bai, C.; Qiu, D.; Kong, F.; Li, Z. Predicting rockburst with database using particle swarm optimization and extreme learning machine. *Tunn. Undergr. Space Technol. Inc. Trenchless Technol. Res.* **2020**, *98*, 103287. [[CrossRef](#)]
26. Wang, S.M.; Zhou, J.; Li, C.Q.; Armaghani, D.J.; Li, X.B.; Mitri, H.S. Rockburst prediction in hard rock mines developing bagging and boosting tree-based ensemble techniques. *J. Cent. South Univ.* **2021**, *28*, 527–542. [[CrossRef](#)]
27. Ma, K.; Shen, Q.Q.; Sun, X.Y.; Ma, T.H.; Hu, J.; Tang, C.A. Rockburst prediction model using machine learning based on microseismic parameters of Qinling water conveyance tunnel. *J. Cent. South Univ.* **2023**, *30*, 289–305. [[CrossRef](#)]
28. Ma, L.; Cai, J.; Dai, X.; Jia, R. Research on Rockburst Risk Level Prediction Method Based on LightGBM–TCN–RF. *Appl. Sci.* **2022**, *12*, 8226. [[CrossRef](#)]
29. Xu, J.; Yang, Y. A survey of ensemble learning approaches. *J. Yunnan Univ. Nat. Sci. Ed.* **2018**, *40*, 1082–1092.
30. Kong, Y.; Jing, M. Research of the Classification Method Based on Confusion Matrixes and Ensemble Learning. *Comput. Eng. Sci.* **2012**, *34*, 111–117.
31. Tan, W.; Hu, N.; Ye, Y.; Wu, M.; Huang, Z.; Wang, X. Rockburst intensity classification prediction based on four ensemble learning. *Chin. J. Rock Mech. Eng.* **2022**, *41*, 3250–3259.
32. Dychkovskiy, R.; Falshtynskiy, V.; Ruskykh, V.; Cabana, E.; Kosobokov, O. A modern vision of simulation modelling in mining and near mining activity. *E3S Web Conf.* **2018**, *60*, 00014. [[CrossRef](#)]
33. Zhou, Y.S.; Cui, J.L.; Zhou, L.Y.; Sun, H.X.; Liu, S.Q. Study on the Evaluation of Personal Credit Risk Based on the Improved Random Forest Model. *Credit Ref.* **2020**, *38*, 28–32.
34. Gao, J.; Li, Y.; Guo, Z.; Tong, W. A Fault Diagnosis Method for Fire Control Systems Based on KPCA-ISSA-SVM. *J. Ordnance Equip. Eng.* **2024**, 1–8.
35. Zhang, Z.; Huang, Y.; Wang, H. A New KNN Classification Approach. *Comput. Sci.* **2008**, 170–172.
36. Xu, R.; Liang, X.; Qi, J.S.; Li, Z.Y.; Zhang, S.S. Advances and Trends in Extreme Learning Machine. *Chin. J. Comput.* **2019**, *42*, 1640–1670.
37. Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; Zhang, L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Trans. Image Process. A Publ. IEEE Signal Process. Soc.* **2017**, *26*, 3142–3155. [[CrossRef](#)] [[PubMed](#)]

38. Cai, T.; Zeng, X. Short Term Load Forecasting Method Based on Multi-feature Extracted Attention-BiGRU. *Hebei Electr. Power* **2023**, *42*, 1–7.
39. Xu, W.; Chen, X.; Zu, T. Research on RMB Exchange Rate Prediction Based on CNN BiGRU Attention Fusion Model. *J. Anqing Norm. Univ. Nat. Sci. Ed.* **2023**, *29*, 35–41.
40. Li, B.; Guo, Z.; Gao, P. Application of improved northern goshawk optimization algorithm in photovoltaic array. *J. Electron. Meas. Instrum.* **2023**, *37*, 131–139.
41. Zhang, J. A sensitive entity identification method for the Internet based on CNN and BiGRU-attention. *Netw. Secur. Technol. Appl.* **2020**, *04*, 61–65.
42. Wu, S.; Wu, Z.; Zhang, C. Rock burst prediction probability model based on case analysis. *Tunn. Undergr. Space Technol. Inc. Trenchless Technol. Res.* **2019**, *93*, 103069. [\[CrossRef\]](#)
43. Long, G.; Wang, H.; Hu, K.; Zhao, Q.; Zhou, H.; Shao, P.; Liao, J.; Gan, F.; He, Y. Probability prediction method for rockburst intensity based on rough set and multidimensional cloud model uncertainty reasoning. *Environ. Earth Sci.* **2024**, *83*, 84. [\[CrossRef\]](#)
44. Adoko, A.C.; Gokceoglu, C.; Wu, L.; Zuo, Q.J. Knowledge-based and data-driven fuzzy modeling for rockburst prediction. *Int. J. Rock Mech. Min. Sci.* **2013**, *61*, 86–95. [\[CrossRef\]](#)
45. Xie, S.; Lin, H.; Chen, Y.; Ma, T. Modified Mohr–Coulomb criterion for nonlinear strength characteristics of rocks. *Fatigue Fract. Eng. Mater. Struct.* **2024**, *47*, 2228–2242. [\[CrossRef\]](#)
46. Xie, S.; Lin, H.; Duan, H. A novel criterion for yield shear displacement of rock discontinuities based on renormalization group theory. *Eng. Geol.* **2023**, *314*, 107008. [\[CrossRef\]](#)
47. Ma, T.; Lin, Y.; Zhou, X.; Wei, P.; Li, R.; Su, J. Entropy weight-normal cloud model for predicting the risk of water breakout in coal bed floor. *Chin. J. Saf. Sci.* **2022**, *32*, 171–177.
48. Chen, Y. Methods for Calculating the Correlation Coefficient. *J. China Exam.* **2011**, 15–19.
49. Ma, T.; Lin, Y.; Zhou, X.; Zhang, M. Grading Evaluation of Goaf Stability Based on Entropy and Normal Cloud Model. *Adv. Civ. Eng.* **2022**, *2022*, 9600909. [\[CrossRef\]](#)
50. Wang, T.; Luo, R.; Ma, T.; Chen, H.; Zhang, K.; Wang, X.; Chu, Z.; Sun, H. Study and verification on an improved comprehensive prediction model of landslide displacement. *Bull. Eng. Geol. Environ.* **2024**, *83*, 90. [\[CrossRef\]](#)
51. Xu, Y.P.; Lin, Z.H.; Ma, T.X.; She, C.; Xing, S.M.; Qi, L.Y.; Farkoush, S.G.; Pan, J. Optimization of a biomass-driven Rankine cycle integrated with multi-effect desalination, and solid oxide electrolyzer for power, hydrogen, and freshwater production. *Desalination* **2022**, *525*, 115486. [\[CrossRef\]](#)
52. Sun, L. Discusses the connection between tensile strength and uniaxial compressive strength of rock. *Agric. Sci. -Technol. Inf.* **2012**, 48–49.
53. Zhou, H.; Xie, H.; Zuo, J. Developments in researches on mechanical behaviors of under the condition of high ground pressure in the depths. *Adv. Mech.* **2005**, *35*, 91–99.
54. Zhou, Q.; Li, H.; Yang, C. Review of evaluation of rockburst and hazard in underground engineerings. *Rock Soil Mech.* **2003**, 669–673.
55. Zhang, J.; Ai, C.; Li, Y.W.; Zeng, J.; Qiu, D.Z. Brittleness evaluation index based on energy variation in the whole process of rock failure. *Chin. J. Rock Mech. Eng.* **2017**, *36*, 1326–1340.
56. Qi, W.; Sun, R.; Zheng, T.; Qi, J. Prediction and analysis model for ground peak acceleration based on XGBoost and SHAP. *Chin. J. Geotech. Eng.* **2023**, *45*, 1934–1943.
57. Xiong, Y.; Chen, G.; Liu, X. Rockburst Prediction Based on Limit Tree Machine Learning Algorithm. *Chin. J. Undergr. Space Eng.* **2023**, *19*, 908–919.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.