applied sciences

Article

Research on Rockburst Risk Level Prediction Method Based on LightGBM−TCN−RF

Li Ma , **Jiajun Cai** *, **Xinguan Dai** and **Ronghao Jia**

College of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710600, China
* Correspondence: caijiajun@stu.xust.edu.cn

Abstract: Rockburst hazards pose a severe threat to mine safety. To accurately predict the risk level of rockburst, a LightGBM−TCN−RF prediction model is proposed in this paper. The correlation coefficient heat map combined with the LightGBM feature selection algorithm is used to screen the rockburst characteristic variables and establish rockburst predicted characteristic variables. Then, the TCN prediction model with a better prediction performance is selected to predict the rockburst characteristic variables at time t + 1. The RF classification model of rockburst risk level with a better classification effect is used to classify the risk level of rockburst characteristic variables at time t + 1. The comparison experiments show that the rockburst characteristic variables after screening allow a more accurate prediction. The overall RMSE and MAE of the TCN prediction model are 0.124 and 0.079, which are better than those of RNN, LSTM, and GRU by about 0.1–2.5%. The accuracy of the RF classification model for the rockburst risk level is 96.17%, which is about 20% higher than that of KNN and SVM, and the model accuracy is improved by 1.62% after parameter tuning by the PSO algorithm. The experimental results show that the LightGBM−TCN−RF model can better classify and predict rockburst risk levels at future moments, which has a certain reference value for rockburst monitoring and early warning.

Keywords: rockburst; risk level; monitoring and warning; LightGBM−TCN−RF model; timing model

Citation: 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. https://doi.org/10.3390/app12168226

Academic Editors: Yougang Sun, Xinwei Yao and Xiaogang Jin

Received: 17 June 2022
Accepted: 15 August 2022
Published: 17 August 2022

1. Introduction

Rockburst is one of the typical dynamic hazards in coal mining. It is a phenomenon of rock bursting and ejection caused by a sudden and violent release of the potential energy accumulated in the rock body under certain conditions, seriously affecting the safe production of the coal industry [1]. Many factors affect rockburst occurrence, and each factor is characterized by complexity and random uncertainty. Scholars worldwide have proposed relevant theories, such as stiffness theory, strength theory, energy theory, and three−factor theory [2,3], which have formed the theoretical basis of coal mine rockburst research, and have laid a research foundation for the monitoring and early warning of rockburst. The widely used monitoring and early warning techniques of rockburst [4–7] based on the physical and mechanical properties of the coal rock body mainly include the borehole coal dust method and the mine pressure observation method. From the geophysical perspective, there are such methods as the microseismic monitoring method, the electromagnetic monitoring method, and the ground sound monitoring method, but due to the difficulty in fully reflecting the complex phenomena in the breeding process of rockburst by single−parameter monitoring and early warning, there is still no unified standard for rockburst monitoring and early warning, making it more difficult to predict the risk level of rockburst accurately. On this basis, scholars have comprehensively analyzed various monitoring data from experience−driven and mechanism−driven perspectives. They have proposed a variety of integrated multiparameter safety evaluation and prediction methods for the burst risk level, in which the integrated index method, the possibility index method, the fuzzy C−mean clustering algorithm, the entropy−based hierarchical analysis [8–11], and other mathematical methods are used. However, the index weight determination and calculation depend on the evaluator's expertise and the authenticity and integrity of the information in the evaluation area. Additionally, with difficulty in making a practical joint analysis of monitoring results, the desired monitoring effect is not achieved. Many experts and scholars have tried to use a data−driven approach to predict rockburst in recent years. They have achieved some results using machine learning techniques. Wen [12] used an improved fruit fly optimization algorithm to select the input layer weights and the implied layer thresholds of the extreme learning machine (ELM) to construct a rockburst hazard prediction model. Lu [13] optimized the kernel parameter σ and penalty factor f of the least−squares supported vector machine (LSSVM) by the particle swarm algorithm to construct a rockburst hazard classification prediction model. He used it to achieve good results in practice. Zhang [14] predicted rockburst classification using a combination of numerical simulation and neural network methods, which is essential for the roadway's early and later construction safety. Chen et al. [15] divided rockburst prediction and warning methods into long-term prediction and short−term warning by analyzing the relevant literature at home and abroad. They outlined the progress and problems in applying the two methods. Zhang [16] used a DNN neural network combined with an SVM classifier to achieve a long-term rockburst prediction. Liu [17] established a CNN−LSTM model to predict the future state of the rockburst characteristic variables and used the PSO−GRNN model according to the future state to predict the t + 1 rockburst level.

Most of the above-mentioned prediction models and methods of the rockburst risk level are for the hazard warning of the current moment. The universal applicability of long-term prediction methods, however, requires further study. Based on the above research, this paper takes a data-driven perspective. The rockburst prediction characteristic variables are established through the feature selection algorithm. The prediction model with the best performance is selected by comparative analysis to predict the rockburst characteristic variables at time t + 1. Through the classification algorithm of supervised learning, the rock burst features at time t + 1 are classified in terms of the burst risk level, and, finally, the long-term prediction of rock burst is realized. The comprehensive analysis of various monitoring system indicators, the selection of important rockburst characteristic variables, and the classification of the risk level by predicting the stated value of the characteristic variables in the future provide a reference for the multi-parameter joint analysis of rockburst and a judgment basis for the safe production of coal mines in the future, and have great significance for promoting rockburst monitoring and early warning.

2. Rockburst Monitoring Data Processing

2.1. Rockburst Monitoring Data Acquisition

Based on the mine rockburst warning project, the data come from Yuanzigou Coal Mine in Baoji, Shaanxi Province, equipped with microseismic monitoring, stress monitoring, and mine pressure monitoring systems. Multiple rockburst events were monitored from January 2020 to 31 March 2021. Due to a large number of monitoring data indicators, the characteristic variables that can characterize the risk of rockburst are initially selected according to practical applications [18], that is, frequency (X1), diurnal frequency variation (X2), energy sum (X3), energy max (X4), energy anomaly coefficient k (X5), and tons of coal release energy (X6) in the microseismic monitoring system, shallow hole stress max (X7), deep hole stress max (X8), and stress increase max (X9) in the stress monitoring system, bolt stress max (X10) and anchor cable prestress max (X11) in the mine pressure monitoring system, AVG daily footage (X12), daily output (X13), and periodic weighting (X14) in progress monitoring, and expert evaluation of the impact hazard level of the working face. Part of the data is shown in Table 1.

Table 1. Rockburst monitoring raw data.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	Risk Level
1	33	0	16,800	2990	1.039477928	1.531868332	7.6	6.7	0	68	138	0	2.4	10,967	0
2	46	13	21,600	2990	1.039477928	1.207041073	8	10	4.3	68	138	0	3.2	17,895	0
3	55	9	22,600	4360	1.039477928	1.329646408	8.2	10.4	0.5	68	138	0	4	16,997	0
	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...															
453	52	−6	21,400	2890	1.477798495	2.127025147	7.6	7.7	0.5	80	103	0	2.4	10,061	1
454	40	−12	13,700	3010	0.89478153	7.413419913	7.6	7.7	0.1	80	103	0	0	1848	1
455	29	−11	4790	829	0.310615395	6.490514905	7.4	7.5	0.1	80	103	0	0	738	1
456	19	−10	9900	2600	0.675306958	2.265446224	7.4	7.5	0.4	80	103	0	1.6	4370	1

Quantitative description of whether there is periodic pressure and rockburst hazard level; 1 for yes, 0 for no. The rockburst risk level is mostly none, weak, medium, and strong, represented by 0, 1, 2, and 3, respectively.

2.2. Rockburst Data Processing

Rockburst is a complex dynamic behavior with many uncertainties, showing diversity, complexity, susceptibility to interference, and other characteristics. In rockburst monitoring and early warning, multiple monitoring system data and other factors may induce rockburst. As a result, a wide range of complex rockburst characteristics are established. Different system indicators have different dimensions and dimension units. To ensure the prediction effect of the model, we need to preprocess the rockburst data before the model is built, and screen the critical characteristic variables of rockburst.

2.2.1. Rockburst Data Preprocessing

During coal mine production, there are problems of monitoring data missing and abnormal monitoring data due to suspension of construction, failure to monitor, or abnormal monitoring equipment. In the data preprocessing process, the small amount of missing monitoring data is compensated with the before−and−after mean method $x_i = (x_{i-1} + x_{i+1})/2$, and abnormal data are replaced using the same method.

Different rockburst monitoring systems cause different magnitudes and units of the data. Suppose the analysis is performed directly with multiple original indicators, among which one feature is more significant in magnitudes, such as energy sum, energy max, and daily output. In contrast, other features are smaller, such as tons of coal release energy and stress increase max. In that case, the final results of model prediction will be dominated by the more prominent feature in magnitude, and the influence of the other features will be weakened. To ensure the role of each feature, the maximum value normalization is performed with all data mapped between 0 and 1. The datasets X1, X2, . . . , and Xn of each feature are transformed by Equation (1).

$$x_{scale} = \frac{x - x_{\max}}{x_{\max} - x_{\min}} \tag{1}$$

In the equation: x represents the data value of the indicator before nondimensionalization, and $x_{\max}$ and $x_{\min}$ represent the maximum and minimum values of the indicator in these sample data. In the normalization process, the differences between the factor feature magnitudes and their effects on the model are reduced, and a better quality rockburst feature magnitude dataset is obtained.

2.2.2. Rockburst Characteristic Variables Screening

Rockburst is triggered by a combination of factors. The Pearson correlation coefficient Equation (2) calculates and determines the interrelationship between the characteristic variables (X1–X14) of rockburst after normalization. The characteristic variables are judged to be independent of each other only when correlation coefficients are zero.

$$\rho(x_1, x_2) = \frac{\mathrm{cov}(x_1, x_2)}{\sqrt{DX_1 DX_2}} \tag{2}$$

As shown in Figure 1, there are strong and weak influence relationships between characteristic variables, and the shades of color represent the strength of the interrelationship.

In other words, there is a connection between the correlations of characteristic variables of rockburst, as shown in Table 2, showing the complexity and diversity of rockburst occurrence. Further feature selection is needed.

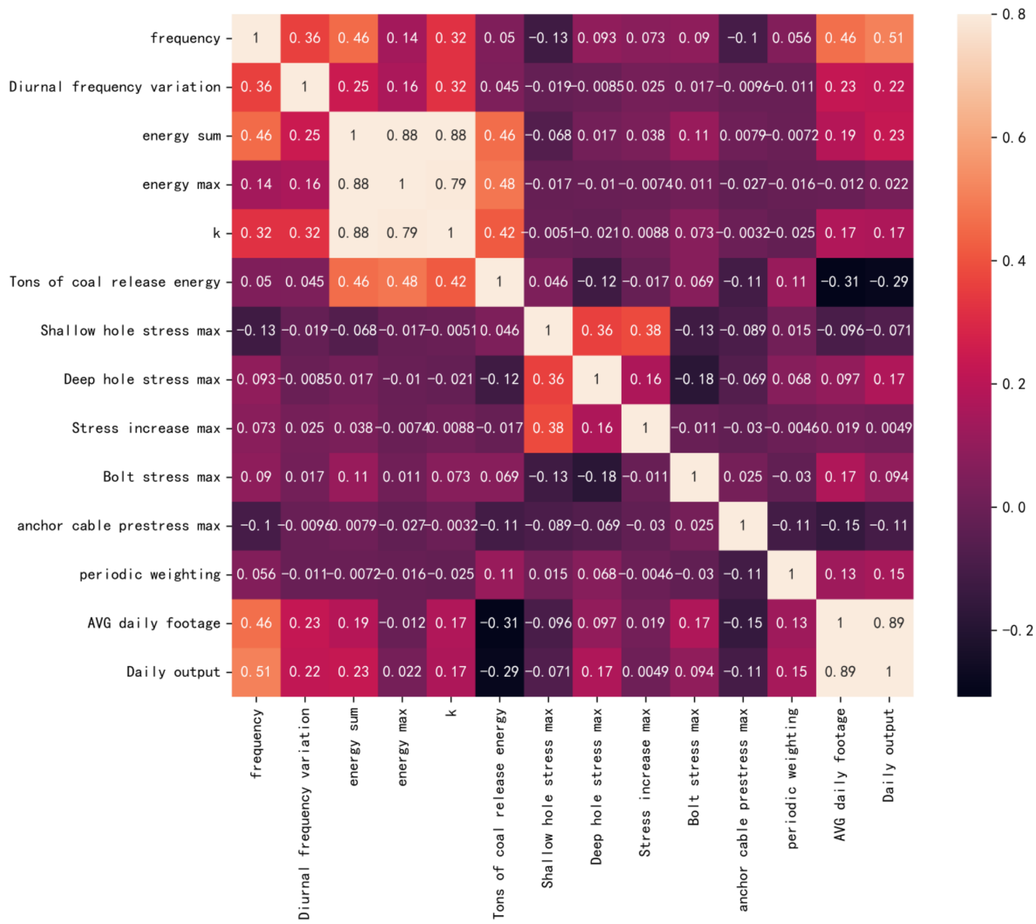

Figure 1. Heat map of the correlation of characteristic variables of rockburst.

Table 2. Correlation coefficient of characteristic variables of rockburst.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
X1	1.000	0.363	0.455	0.138	0.320	0.050	−0.126	0.093	0.073	0.090	−0.101	0.463	0.508	0.056
X2	0.363	1.000	0.246	0.162	0.323	0.045	−0.019	−0.008	0.025	0.017	−0.010	0.225	0.218	−0.011
X3	0.455	0.246	1.000	0.885	0.880	0.462	−0.068	0.017	0.038	0.114	0.008	0.186	0.230	−0.007
X4	0.138	0.162	0.885	1.000	0.793	0.482	−0.017	−0.010	−0.007	0.011	−0.027	−0.012	0.022	−0.016
X5	0.320	0.323	0.880	0.793	1.000	0.417	−0.005	−0.021	0.009	0.073	−0.003	0.174	0.166	−0.025
X6	0.050	0.045	0.462	0.482	0.417	1.000	0.046	−0.122	−0.017	0.069	−0.112	−0.308	−0.290	0.113
X7	−0.126	−0.019	−0.068	−0.017	−0.005	0.046	1.000	0.361	0.381	−0.126	−0.089	−0.096	−0.071	0.015
X8	0.093	−0.008	0.017	−0.010	−0.021	−0.122	0.361	1.000	0.164	−0.183	−0.069	0.097	0.172	0.068
X9	0.073	0.025	0.038	−0.007	0.009	−0.017	0.381	0.164	1.000	−0.011	−0.030	0.019	0.005	−0.005
X10	0.090	0.017	0.114	0.011	0.073	0.069	−0.126	−0.183	−0.011	1.000	0.025	0.170	0.094	−0.030
X11	−0.101	−0.010	0.008	−0.027	−0.003	−0.112	−0.089	−0.069	−0.030	0.025	1.000	−0.152	−0.107	−0.109
X12	0.463	0.225	0.186	−0.012	0.174	−0.308	−0.096	0.097	0.019	0.170	−0.152	1.000	0.888	0.129
X13	0.508	0.218	0.230	0.022	0.166	−0.290	−0.071	0.172	0.005	0.094	−0.107	0.888	1.000	0.149
X14	0.056	−0.011	−0.007	−0.016	−0.025	0.113	0.015	0.068	−0.005	−0.030	−0.109	0.129	0.149	1.000

After determining that there are no unrelated variables among the rockburst characteristic variables, it is necessary to judge the correlation degree of characteristic variables relative to the impact hazard and eliminate the characteristic variable with a lower correlation degree to complete the secondary screening. LightGBM [19,20] is a framework for implementing the gradient boosting decision tree (GBDT) algorithm, which is designed to solve the problems encountered by GBDT on massive data. Its principle is generally similar to that of GBDT. That is, the negative gradient of the loss function is determined as the residual approximation of the current weak classifier decision tree to fit a new decision tree. In other words, a new loss function is added to the original iterative model to keep the predicted value close to the current actual value while ensuring that the state of the original model remains unchanged at each iteration, as shown in Equation (3).

$$\hat{\mathrm{y}}_{\mathrm{i}} = \sum_{q=1}^{Q} f_q(x_i), f_q \in \chi \tag{3}$$

χ is the function space of the iterative tree, and $f_q(x_i)$ denotes the predicted value of the ith sample in the qth tree. For further feature selection, the LightGBM model is used to calculate the importance of each feature relative to the rockburst hazard level, and the heuristic information of the iterative tree is used as an important measure of the feature. The quality of the subset of candidate features was directly affected by the metric of the tree structure. Based on the total number of times each feature is segmented in the iterative tree (T_Split), the LightGBM model measures the feature importance by calculating the sum of the gains (T_Gain) that result from the features being used for segmentation in all decision trees, as shown in Equations (4) and (5).

$$T\_Split = \sum_{t=1}^{k} Split_t \tag{4}$$

$$T\_Gain = \sum_{t=1}^{k} Gain_t \tag{5}$$

X1 to X14 are used as 14 characteristic variables, and the rockburst risk level is used as the label. The LightGBM model is established by dividing the dataset and is then trained. The importance of each feature for the rockburst risk level is calculated, as shown in Figure 2. Based on the actual situation of coal mine production, the less critical characteristic variables, such as diurnal frequency variation (X2), energy anomaly coefficient k (X5), AVG daily footage (X12), and aperiodic weighting (X14), are discarded to establish the new rockburst characteristic variables after screening.

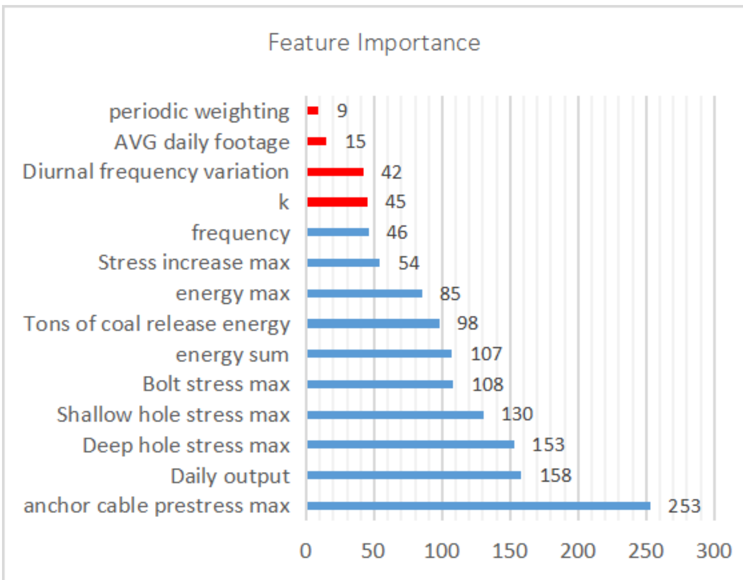

Figure 2. The importance of each characteristic variable of rockburst.

3. Rockburst Characteristic Variable Prediction Model

3.1. Building Process of Rockburst Characteristic Variable Prediction Model

The building process of the rockburst characteristic variable prediction model is shown in Figure 3.

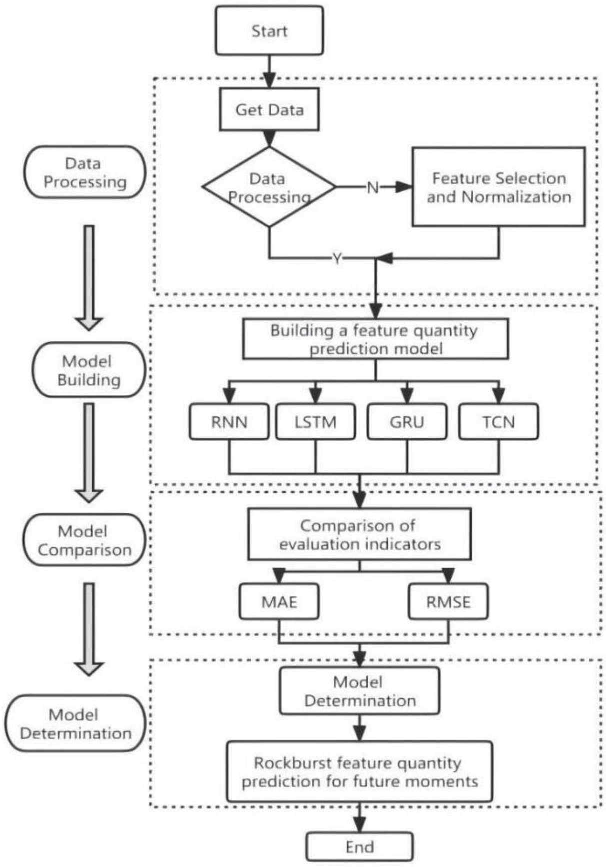

Figure 3. Building process of rockburst characteristic variables prediction model.

① Sample data processing. The missing values and outliers of the dataset are filled and replaced, and the minimum and maximum value normalization of the dataset is carried out, so as to unify the dimensions of the characteristic variable data and establish the ideal dataset.

② Feature variable selection. By calculating the correlation coefficients between characteristic variables of rockburst, and the importance of each characteristic variable relative to the impact hazard, the predicted characteristic variable is selected. The secondary screening is completed to establish the characteristic variables for rockburst prediction.

③ Experimental comparison of rockburst characteristic variable prediction models. RNN, LSTM, GRU, and TCN models are developed for rockburst characteristic variable prediction. The datasets are divided for model training and learning. The mean absolute error (MAE) and root mean square error (RMSE) of each model are compared to analyze the prediction performance of different algorithmic models for rockburst characteristic variables.

④ Model determination. The model with the most desirable evaluation index is selected by comparison. The state value at future t + 1 moment of the rockburst characteristic variables is predicted.

3.2. Rockburst Characteristic Variable Prediction Modeling

Since the rockburst monitoring data are temporal data, four temporal models, recurrent neural network (RNN), long short−term memory network (LSTM), gated recurrent unit

(GRU), and temporal convolutional neural network (TCN), are established to learn, train, and predict the rockburst characteristic variables.

The recurrent neural network (RNN) [21] is used to process sequential data. There are connections between the nodes in the hidden layers of its network structure. The input of the hidden layer contains the input of the input layer and the output of the hidden layer at the previous moment. This enables RNN to memorize the previous information and directly apply it to the analysis and calculation of the current output, solving the situation that traditional neural networks cannot handle time−series changes. The RNN training parameters, however, are used to calculate the local minima of the curve using the gradient descent method. When the minimum point is slightly crossed, a large gradient in the opposite direction of the gradient will make a reverse direction away from the required local minimum and result in the process of repeatedly seeking the minimum. As the number of calculations increases, the gradient decreases rapidly, and it takes a long time to approach this minimum value again, and problems such as gradient disappearance occur.

Long short−term memory (LSTM) is a variant of RNN, a special temporal recurrent neural network that is proposed to solve the problem of gradient disappearance and gradient explosion during the training of long sequences of RNN [17,22]. LSTM has a more complex structure compared with the ordinary RNN. The key is the information state of the neural unit. Three logic gates control the information flow of the neural unit (forgetting gate, input gate, and output gate) as shown in Figure 4. Tanh denotes the tanh function as the activation function of the candidate state, and σ denotes the sigmoid function, whose expressions are shown in Equations (6) and (7), respectively.

$$\tan h(t) = \frac{e^t - e^{t-1}}{e^t + e^{t-1}} \tag{6}$$

$$\sigma(t) = \frac{1}{1 + e^{t+1}} \tag{7}$$

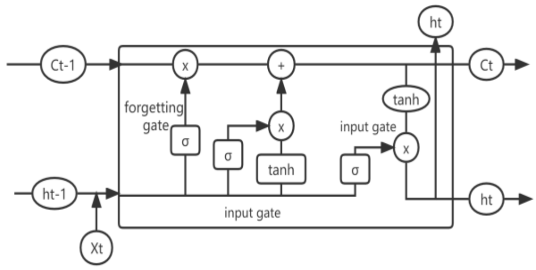

Figure 4. LSTM network structure.

The forgetting gate determines whether the information is discarded or retained. When the previous information in the hidden state and the current input information are entered into the sigmoid function at the same time, the output value is between 0 and 1. The closer it is to 0, the more it should be forgotten, and the closer it is to 1, the more it should be retained. The input gate is used to update the cell state. Additionally, the information is input into the sigmoid function, and the output value is adjusted between 0 and 1 to decide which information to update, with 0 indicating unimportant and 1 indicating important. The output gate can determine the value of the next hidden state, which contains relevant information from previous inputs, so the hidden state can also be used to predict the next moment state.

The gate recurrent unit (GRU) [23] is also a well−known variant of the RNN structure, which is designed for the same purpose as LSTM. The GRU is a variant that can achieve better processing results based on LSTM networks. It combines the forgetting gate and input gate of LSTM into a single update gate, removes the cell states, and uses hidden states for information transfer. Therefore, GRU does not erase the previous useful information over time. It can preserve the information in long-term sequences and avoids the gradient disappearance by using all the temporal information. The network structure of GRU is shown in Figure 5. The update gate is mainly used to control the influence of the information h_{t-1} of the previous state on the current state h_t. The larger value of the update gate means the greater influence of the state information of the previous moment on the current state, i.e., the larger amount of information from the previous moment. The reset gate is used to control the influence of the previous moment's state information h_{t-1} on the candidate state x. The smaller value of the reset gate means that the previous moment's state information has less influence on the candidate information and the amount of incoming information is less. It is simpler than the LSTM network, improving the control effect significantly. GRU has a small number of parameters, which reduces the risk of overfitting and solves the complicated calculation of LSTM.

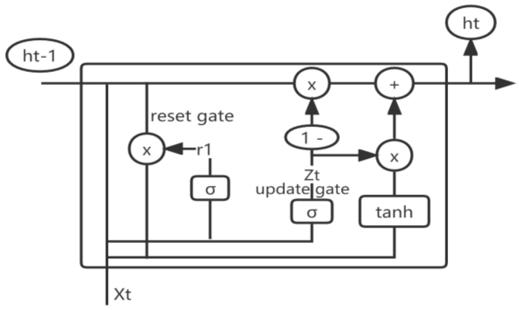

Figure 5. GRU network structure.

Traditional convolutional neural networks are generally considered less suitable for time sequence modeling. In contrast, a temporal convolutional network (TCN) [24,25], a new type of neural network based on causal convolution, dilated convolution, and residual connections, has model predictions that are only influenced by historical data and not permeated by future data. Causal convolution changes p weights at moment t for each layer in the convolutional network, only from the p weight changes of the nodes at moment t and previous moments in its previous layer, as in Equation (8). TCN adopts hole convolution by sampling the input data exponentially layer by layer, thus achieving a larger receptive field with fewer network layers, as in Equation (9), where F is the convolution operation on the sequence element s, f is a filter, k is the size of the filter, d is the expansion factor, and $s - d \cdot i$ illustrates the past direction. In this way, TCN solves the problem that the traditional neural network cannot be applied to the modeling of temporal data. As the depth of the network continues to increase, the amount of information acquired decreases layer by layer. TCN prevents information loss by constructing residual links to transfer information across layers in the network structure. It gives the TCN network the advantages of parallel operation, the stable gradient of general convolutional neural networks, and the ability of long sequence information learning. In addition, TCN keeps more extended memory than LSTM, reducing the complexity of the network while obtaining long-term historically valid information. It can effectively handle sequence modeling tasks even better than other models.

$$p(x) = \prod_{t=1}^{T} p(x_t | x_1, \ldots, x_{t-1}) \tag{8}$$

$$F(s) = (x *_{d*} f)(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-d \cdot i} \tag{9}$$

3.3. Comparative Analysis of the Performance of Rockburst Characteristic Variable Prediction Models

Four time−series models are built to perform characteristic time−series forecasts using the processed rockburst characteristic dataset. To compare the prediction effects of the four models, the dataset is divided into the training set and the test set, and the training data of the ten rockburst characteristic variables are used to train the four models to extract the temporal features of data evolution. During the training process, the activation function is Relu, Adam is selected as the optimizer of the model, the parameter epoch is set to 100, and the number of neurons is 64.

The test set is input to the four trained models to obtain the final predictions of the ten rockburst characteristic variables. The evaluation index root mean square error (*RMSE*), and mean absolute error (*MAE*) are used as in Equations (10) and (11) to compare and analyze the prediction effect of the model test set as in Figure 6. The smaller the index is, the better the model prediction effect.

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(y_{true} - y_{pre}\right)^2} \tag{10}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\left(y_{ture} - y_{pre}\right)\right| \tag{11}$$

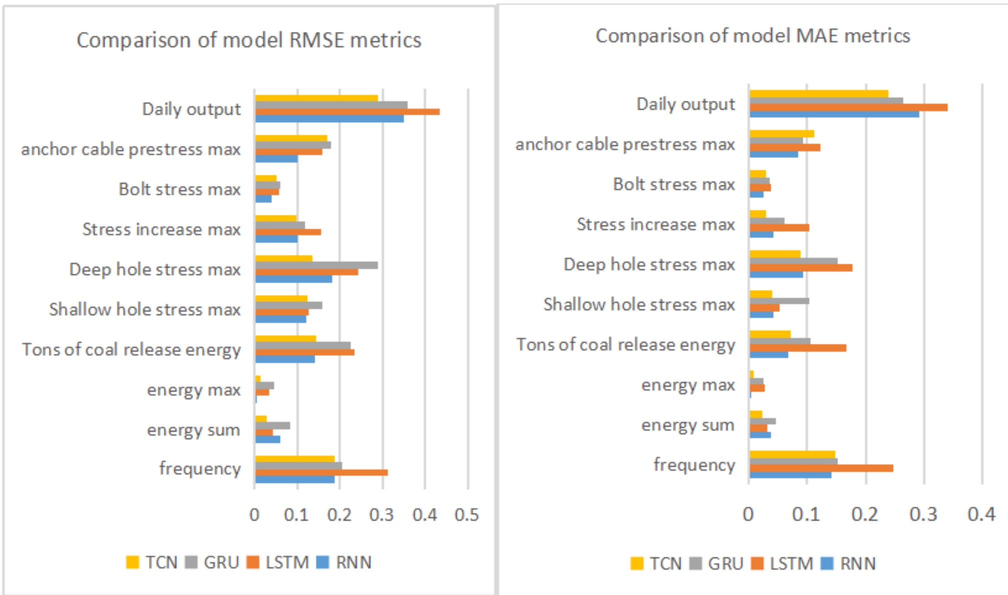

Figure 6. Comparison of the four models' performance indicators.

By comparing the prediction of rockburst characteristic variables by four time−series models, it can be seen that the models' performances in prediction vary for different characteristic variables. As shown in Table 3, the TCN model has minimal overall prediction error with better stability, i.e., its mean RMSE and mean MAE are the smallest, 0.124 and 0.079, respectively. The overall prediction performance of the TCN model is better compared to the other three models.

Table 3. Comparison of four models' indicators.

Model\Indicators	RMSE Mean Value	MAE Mean Value
RNN	0.129	0.0823
LSTM	0.180	0.131
GRU	0.173	0.104
TCN	0.124	0.079

The TCN model is chosen to predict the ten characteristic variables at time t + 1, as shown in Table 4.

Table 4. Rockburst characteristic variable prediction at time t + 1.

Rockburst Characteristic Variables	X1	X3	X4	X6	X7	X8	X9	X10	X11	X13	Rockburst Risk Level
Forecast at the moment t + 1	63	22,796	4534	2.295048	7.8	7.6	0.1	79	125	10,715	To be tested

4. Rockburst Risk Level Classification Model

4.1. Rockburst Risk Level Classification Model Building Process

Figure 7 shows the building process of the rockburst risk level classification model.

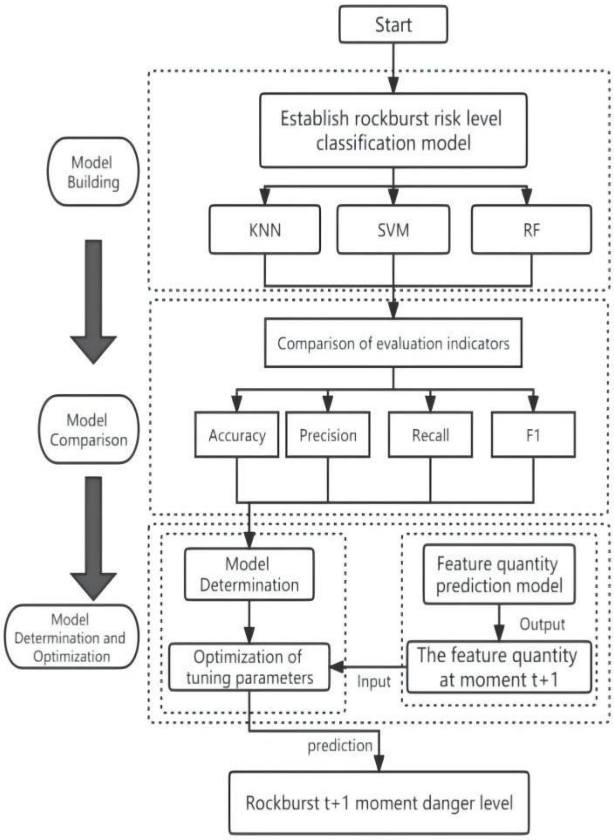

Figure 7. Rockburst risk level classification model building process.

① Classification model establishment. The classification models KNN, SVM, and RF are established for the rockburst risk level, and the dataset is divided for model training and learning.

② Experimental comparison of classification models. Classification evaluation indexes are used to compare the classification effects of the three classification models. The model with the best effect is determined.

③ Model optimization. After the parameters of the determined classification model are optimized, the predicted rockburst characteristic variable is used as the input to achieve the rockburst hazard level prediction.

4.2. Rockburst Risk Level Classification Modeling

The rockburst risk level assessment is an important safeguard to measure production stability and potential threats in coal mines. According to the relevant provisions, the rockburst risk can be divided into four levels: no impact hazard, weak impact hazard, medium impact hazard, and strong impact hazard. With the predicted values of the rockburst characteristic variables at time t + 1, the classification model of the rockburst risk level is established using classification algorithms of machine learning based on supervised learning, such as K−nearest neighbor (KNN), support vector machine (SVM), and random forest (RF), to achieve the prediction of the rockburst risk level in the future.

The basic principle of the K nearest neighbor algorithm (KNN) [26,27] is to determine the class of new values based on the nearest k points (k known samples) and the voting rule that the minority obeys the majority to achieve classification in predicting new values. The model is easy to understand and does not require much tuning to achieve good performance.

Support vector machines (SVMs) [28,29] are supervised learning models with outstanding performance in classification and regression analysis. A machine learning method is adopted for predicting unknown sample classes by learning the features of known samples in different classes. The learning machine's generalization ability is improved by seeking the minimum structural risk, which minimizes the empirical risk and confidence interval, and obtains good statistical prediction laws in the practical situation where the size of overall random statistical samples is small. It can also quickly solve nonlinear pattern recognition and high−dimensional data problems.

Random forest (RF) [30,31] is an integrated learning algorithm consisting of a class of decision tree models as a base learner. When used for classification, the algorithm constructs decision trees by randomly selecting samples and sample features in the dataset, and the process is repeated multiple times. The decision trees are uncorrelated from each other, and the results of all decision trees are counted as the final result. After the classification results of each decision tree in the forest are counted in predicting new samples, the most statistics categories are selected as the new sample prediction results.

4.3. Rockburst Risk Level Classification Model Comparison

Three types of classification and prediction models for rockburst risk are established. The dataset is divided into the training set and test set to select rockburst characteristic variables as features, and the impact hazard level is used as a label. The training set is used to train the three models, and then the test set is input into the trained model for validation. The confusion matrices of the models in Figure 8 are compared, and the accuracy, precision, recall, and F_1 value, as in Equations (12)–(15), are used to evaluate the index comparison, as in Figure 9.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

$$P = \frac{TP}{TP + FP} \tag{13}$$

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

$$F_1 = \frac{2TP}{2TP + FP + FN} \tag{15}$$

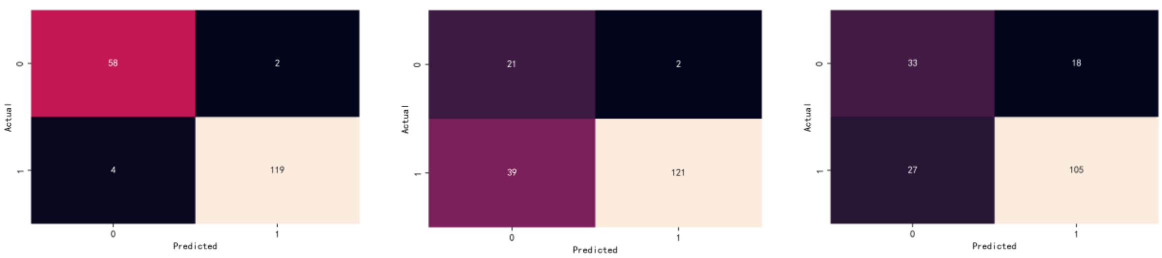

Figure 8. RF, SVM, and KNN confusion matrix comparison. (From left to right, they are RF, SVM, KNN).

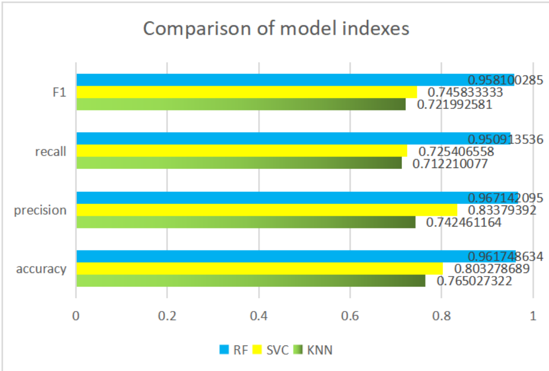

Figure 9. Comparison of three classification model indicators.

In the equation, *FN* is the number of positive examples predicted incorrectly, *TP* is the number of positive examples predicted correctly, *FP* is the number of negative examples predicted incorrectly, and *TN* is the number of negative examples predicted correctly. A comparison of the effects of the rockburst risk level classification prediction models is shown in Figure 9.

It can be seen that the performance of the RF model is significantly better than that of the KNN and SVM models. Its model accuracy is 0.962, and the three indicators of F_1, recall, and precision are better than those of the other models. This shows that the RF model has the most significant recognition ability. That is, the RF model performs better in classification than the other two models, and can be used as the preliminary classification and prediction model for the rockburst level.

4.4. Rockburst Risk Level Classification Model Optimization

The parameters of the primary RF model are optimized and adjusted. Different values of essential parameters of the model affect the results of rockburst risk level prediction. To improve the accuracy of the predictive classification of the model, particle swarm optimization (PSO) [32,33] is used to tune the number of trees (n_estimators) and the maximum depth of the tree (max_depth) among the parameters of the RF model. The PSO algorithm treats each particle in the generated initial population as a feasible solution. The direction and distance of the particle motion during the population motion are determined according to its velocity. Under the condition that the motion of each particle closely follows the current optimal solution, the optimal solution is obtained from the ordered spatially feasible solution by continuously updating its information, as in Figure 10.

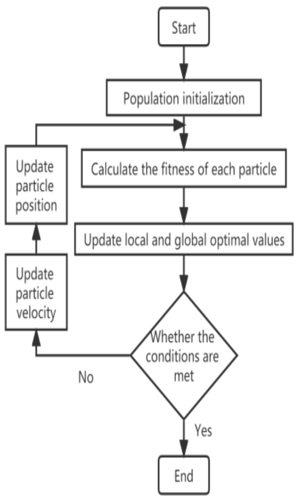

Figure 10. PSO optimization algorithm flow.

The PSO optimization algorithm has the advantages of easy implementation, fast convergence, and high accuracy, so it is attracting increasing attention and has produced significant results in solving practical optimization problems. In the tuning process, the number of particles in the particle swarm is set to 100, the maximum number of iterations is 30, the inertia factor is set to 0.8, and the mean value of the six−fold cross−validation training accuracy is used as the final evaluation metric. The iterative tuning of the RF model parameters is shown in Figure 11.

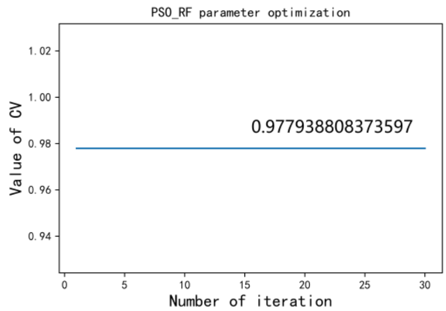

Figure 11. Random forest tuning effect.

After parameter tuning, the accuracy of the RF model for rockburst risk level prediction is increased by 1.62%, and the predicted rockburst risk level at time t + 1 is shown in Table 5. The rockburst risk level at t + 1 is none, which can be used as a basis for judgment and provide a reference for coal mine safety production.

Table 5. Rockburst risk level prediction at time t + 1.

Rockburst Characteristic Variables	X1	X3	X4	X6	X7	X8	X9	X10	X11	X13	Rockburst Risk Level
Forecast at the moment t + 1	63	22,796	4534	2.295048	7.8	7.6	0.1	79	125	10,715	No risk

5. Conclusions

(1) The interrelationship between rockburst characteristic variables was analyzed by using a Pearson correlation coefficient heat map. The importance of characteristic variables for impact hazard was judged through lightGBM model feature selection, and rockburst characteristic variables were established.

(2) Four temporal models, RNN, LSTM, GRU, and TCN, were analyzed. RMSE and MAE were used as evaluation indexes for comparison. The TCN model with the best overall prediction performance was selected. A model was established to predict the status value of the rockburst characteristic variables at t + 1.

(3) By using the classification algorithms, KNN, SVM, and RF, in machine learning, a classification model for rockburst risk level was established. The classification effects of the three models were analyzed and compared. The RF model with the best classification effect was selected, and its parameters were tuned by the PSO algorithm, which improved model prediction accuracy by 1.62%. Based on the predicted status value of the rockburst characteristic variable at future time t + 1, the risk level of the rockburst characteristic variable can be predicted, which provides a reference basis for coal mine safety production and has greater promotion significance for rockburst monitoring and early warning.

Author Contributions: Methodology, L.M.; formal analysis, R.J.; writing—original draft preparation, J.C.; writing—review and editing, L.M.; project administration, X.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (52104215) and the Key Industry Innovation Chain Project of Shaanxi Key Research and Development Plan (2021ZDLGY07−08).

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

References

1. Xie, H.P. Research progress on deep rock mechanics and mining theory. *J. China Coal Soc.* **2019**, *44*, 1283–1305.
2. Jiang, Y.D.; Pan, Y.S.; Jiang, F.X.; Dou, L.M.; Ju, Y. Mechanism and prevention of rock explosion in coal mining in China. *J. China Coal Soc.* **2014**, *39*, 205–213.
3. Xia, Y.X.; Pan, J.F.; Xie, F.; Sun, X.D.; Lu, C.; Zhang, C.Y.; Liu, S.H. Research on the mechanism and application of long-hole fracture pressure relief and anti-scouring in rock-bursting back-hole mining. *J. China Coal Soc.* **2022**, 1–11. [CrossRef]
4. Dou, L.M.; Tian, X.Y.; Cao, A.Y.; Gong, S.Y.; He, H.; He, J.; Cai, W.; Li, X.W. Current situation and problems of rock explosion prevention and control in China's coal mines. *J. China Coal Soc.* **2022**, *47*, 152–171.
5. Zhang, J.W.; Zhong, S. Research progress and technical system analysis of rock explosion disaster in China's coal mines. *Coal Technol.* **2021**, *40*, 105–108.
6. Dou, L.M.; Zhou, K.Y.; Song, S.K.; Cao, A.Y.; Cui, H.; Gong, S.Y.; Ma, X.T. Research on the mechanism, monitoring and warning, and prevention and control technology of coal mine impact pressure. *J. Eng. Geol.* **2021**, *29*, 917–932.
7. Yan, Y.D.; Pan, J.F.; Xi, G.J.; Jiao, B.; Shi, X. Analysis and prevention of impact hazards in the squared-off tectonic zone of consolidated mining. *Ind. Mine Autom.* **2021**, *47*, 7–13.
8. Zhang, W.J. Prediction and Prevention Technology of Rockburst in Deep Buried Extra-Thick Coal Seam Roadway in Mengcun Mine. Master's Thesis, Xi'an University of Science and Technology, Xi'an, China, 2019.
9. Zuo, L.Y. Rockburst hazard prediction of the isolated working face based on multi-factor analysis. *Coal Eng.* **2019**, *51*, 96–100.
10. Wu, J.F.; Cheng, Y.F.; Liang, Y.H.; Zhang, F.; Guo, X.F. Rockburst hazard level prediction by fuzzy C-mean clustering algorithm. *J. Heilongjiang Univ. Sci. Technol.* **2021**, *31*, 350–353+365.
11. Zhu, F.; Zhang, H.W. CW-TOPSIS rockburst evaluation model based on the "AHP + entropy weight method". *China Saf. Sci. J.* **2017**, *27*, 128–133.
12. Wen, T.X.; Li, Y.Z. Preprocessing-based AFOA-ELM rockburst hazard prediction model. *China Saf. Sci. J.* **2019**, *29*, 29–34.
13. Lv, P.F.; Qiu, L. Research on rockburst classification prediction in mines based on PSO-LSSVM. *Min. Saf. Environ. Prot.* **2021**, *48*, 120–125.
14. Zhang, J.; Wang, M.X.; Xi, C.H.; Feng, G. Prediction and Evaluation of Rockburst Based on Depth Neural Network. *Adv. Civ. Eng.* **2021**, *2021*, 8248443. [CrossRef]
15. Chen, J.; Gao, J.K.; Pu, Y.Y.; Jiang, D.Y.; Qi, Q.X.; Wen, Z.J.; Sun, Q.L.; Chen, L.L. A machine learning approach for rockburst prediction and early warning. *J. Min. Strat. Control Eng.* **2021**, *3*, 57–68.
16. Zhang, K. Research on Multiparameter Integrated Early Warning of Coal Column Type Rockburst Based on Neural Network Analysis. Master's Thesis, Shandong University of Science and Technology, Qingdao, China, 2018.
17. Liu, H.M.; Xu, F.Y.; Liu, B.J.; Deng, M. A CNN-LSTM-based method for temporal prediction of rockburst hazard level. *J. Cent. South Univ. (Sci. Technol.)* **2021**, *52*, 659–670.
18. Zhang, X.F.; Qu, X.C.; Wei, Q.D. Development and application of a multidimensional multiparameter monitoring and warning platform for rock blast. *J. Min. Strat. Control Eng.* **2021**, *3*, 69–78.
19. Ding, J.J.; Sun, Y. LightGBM-based multi-category prediction of flight delays. *J. Nanjing Univ. Aeronaut. Astronaut.* **2021**, *53*, 847–854.
20. Li, Z.S.; Yao, X.; Liu, Z.G.; Zhang, J.C. Feature selection algorithm based on LightGBM. *J. Northeast. Univ. (Nat. Sci.)* **2021**, *42*, 1688–1695.
21. Liu, F. Research on a Comprehensive Prediction Method of Coal Mine Impact Hazard Based on Machine Learning. Doctoral Dissertation, Shandong University of Science and Technology, Qingdao, China, 2020.
22. Zhao, H.R.; Zhao, Y.H.; Guo, S. Short-term electric load forecasting is based on the empirical modal decomposition of complementary sets and long- and short-term memory neural networks. *Electr. Power* **2020**, *53*, 48–55.
23. Peng, Y.; Meng, Z.Q.; Yang, L.N. A speech enhancement method based on GRU neural network. *J. Guangxi Univ.* **2021**, *46*, 1533–1548.
24. Fu, Y.; Ren, Z.X.; Wei, S.R.; Wang, Y.; Huang, L.L.; Jia, F. Ultra short term power prediction of offshore wind power based on improved lstm-tcn model. *Proc. CSEE* **2022**, *42*, 4292–4303. [CrossRef]
25. Nie, Q.Q.; Wan, D.S.; Zhu, Y.L.; Li, Z.J.; Yao, C. Hydrological model based on time-domain convolutional neural network. *J. Comput. Appl.* **2022**, *042*, 1756–1761.
26. Bian, Z.K.; Zhang, J.; Wang, S.T. KNN classification algorithm based on class label perception. *J. Pattern Recognit. Artif. Intell.* **2021**, *34*, 873–884. CrossRef]
27. Zhang, Y.; Ren, H.X.; Wang, D.L. The intelligent evaluation system of ship maneuvering is based on the KNN algorithm. *J. Shanghai Marit. Univ.* **2021**, *42*, 33–38.

28. Yang, W.B.; Wang, Z.X.; Tian, H.S.; Wu, F.Y.; Yang, Z.C. Large deformation prediction method for laminated soft rock tunnel based on PSO-SVM algorithm. *Hazard Control Tunn. Undergr. Eng.* **2022**, 1–10. CrossRef]
29. Zhang, M.; Chen, J.H.; Zhou, Z.H. SVM-based rockburst classification prediction model and R language implementation. *Chin. J. Geol. Hazard Control* **2018**, *29*, 64–69.
30. Wan, Q.Z.; Wu, K.C. Islanding detection of DC microgrid based on random forest classification. *J. Acta Energ. Sol. Sin.* **2022**, *43*, 269–276. [CrossRef]
31. Wang, Z.Y.; Zhang, M.Y.; Yu, Y.R.; Ni, P. A random forest sea ice fine classification method integrated texture features and NDVI. *Acta Oceanol. Sin.* **2021**, *43*, 149–156.
32. Gu, Y.F.; Zhang, D.Y.; Ruan, J.F.; Wang, Q.; Bao, Z.D.; Zhang, H.Z. A new model for permeability prediction in oil and gas reserve evaluation. *Prog. Geophys.* **2022**, *37*, 588–599.
33. Zhang, M.C. Prediction of rockburst hazard based on particle swarm algorithm and neural network. *Neural Comput. Appl.* **2022**, *34*, 2649–2659. [CrossRef]