Signal recognition and prediction system based on random forest model

Junhao Hu

China University of Petroleum-Beijing at Karamay, Karamay City, Xinjiang Uygur Autonomous Region 834000

3034308054@qq.com

Abstract. To accurately predict the occurrence of rock bursts during deep coal mining and ensure the safe and efficient operation of coal mines, a signal recognition and prediction system based on the Random Forest model is proposed. This system utilizes electromagnetic radiation (EMR) and acoustic emission (AE) signal data, employing feature extraction and point biserial correlation analysis to screen out the most relevant feature parameters. A Random Forest binary classification model is constructed to identify interference signals. Subsequently, by introducing new features such as moving average slope and exponentially weighted moving average (EWMA), a time series analysis of precursor feature signals is conducted. A real-time warning model based on the Random Forest algorithm is developed, dynamically calculating the probability of precursor feature signal occurrence by integrating historical data and real-time data changes. This approach improves the accuracy and recall rate of signal recognition and prediction, providing reliable data support for mine safety management.

Keywords: Feature extraction, Point biserial correlation analysis, Random Forest model, Subsampling

1. Introduction

During the deep mining process of coal mines, rock burst disasters have become one of the significant threats to mine safety. As mining depth and intensity increase, rock bursts occur more frequently, severely affecting the safety and economic efficiency of coal production. Rock burst refers to the phenomenon where, when the stress in coal and rock masses accumulates to a certain level, a large amount of energy is suddenly released, causing the coal and rock masses to fracture and produce strong vibrations. The suddenness and destructiveness of this disaster make it a major challenge in ensuring the safe production of coal mines.

To effectively predict the occurrence of rock bursts, researchers have proposed various prediction methods [1][14]-[15], including traditional mathematical models and modern intelligent algorithms. Early prediction methods were mainly based on analyzing the stress state of coal and rock masses, such as strength theory, stiffness theory, and energy theory. Although these methods could explain the mechanism of rock bursts to some extent, their models were too simplified to accurately predict rock bursts in the complex and variable coal mine environment [2].

In recent years, with the development of artificial intelligence and big data technology, machine learning-based prediction methods have gradually gained attention. These methods utilize large amounts

of historical data, and by employing complex algorithms for modeling and analysis, they can predict the occurrence of rock burst more accurately. For example, common rock burst prediction methods include: the support vector machine (SVM) method, which is suitable for handling small samples and high-dimensional data; convolutional neural networks (CNN) and long short-term memory networks (LSTM)[3], which have significant advantages in processing time series and image data, can automatically extract features and perform complex pattern recognition, and are suitable for real-time monitoring and prediction of rock burst signals; and Bayesian methods, which combine prior knowledge and data for prediction and are suitable for environments with high uncertainty.

Among all machine learning methods, the Random Forest algorithm, as an ensemble learning method, improves prediction accuracy and stability by constructing multiple decision trees and combining their results. It performs excellently in handling high-dimensional data and nonlinear relationships, making it suitable for rock burst prediction in complex environments.

This paper proposes a signal recognition and prediction system based on the Random Forest model, aiming to achieve accurate prediction of rock bursts through the analysis of electromagnetic radiation (EMR) and acoustic emission (AE) signals. The system first uses feature extraction and point biserial correlation analysis to select the feature parameters most strongly correlated with rock bursts, constructing a Random Forest binary classification model to identify interference signals. Subsequently, using new features such as moving average slope and exponentially weighted moving average (EWMA), it conducts time series analysis on precursor feature signals to predict the possible occurrence time period of rock bursts, providing reliable data support for mine safety management.

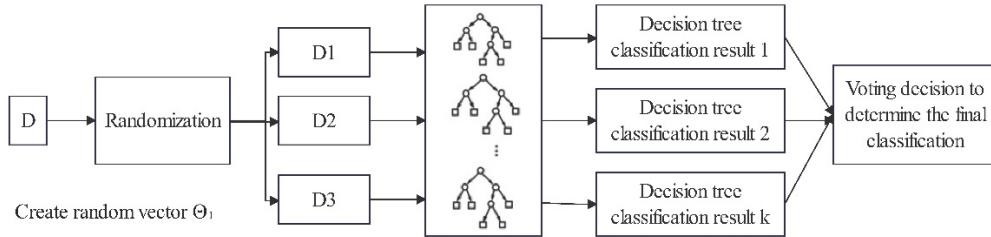## 2. Analysis of Rock Burst Prediction Problems

Studies have shown [5] that before the occurrence of a rock burst, the coal and rock masses exhibit specific precursor features in the form of electromagnetic radiation (EMR) and acoustic emission (AE) signals. These signals typically show a significant cyclic increasing trend within approximately seven days before a rock burst occurs. To achieve precise prediction of rock bursts, it is essential to focus on identifying and predicting these precursor feature signals. This article focuses on analyzing these precursor feature signals, aiming to establish a mathematical model to predict the possible time period of rock bursts, thereby ensuring the safety of coal mine workers.

3. Signal Recognition and Prediction Model Based on Random Forest

3.1. Random Forest Algorithm

Random Forest (RF) is a statistical learning theory that uses the bootstrap resampling method to extract multiple samples from the original sample. Each bootstrap sample is used to build a decision tree model, and then multiple decision trees are combined to make predictions. The final prediction result is obtained through voting [4].

Random Forest Classification (RFC) is an ensemble classification model composed of many decision tree classification models $\{h(X, \ominus_k), k = 1, 2, \ldots\}$, where the parameter sets $\{\ominus k\}$ are independent and identically distributed random vectors. Given a set of independent variables X, each decision tree classification model has one vote to select the optimal classification result. The basic idea of RFC is as follows: first, use bootstrap sampling to extract k samples from the original training set, with each sample having the same sample size as the original training set; second, establish k decision tree models for the k samples to obtain k classification results; finally, vote on the k classification results for each record to determine its final classification, as shown in Fig. 1.

Figure 1: RF Diagram

Random Forest increases the diversity among classification models by constructing different training sets, thereby enhancing the extrapolation and predictive ability of the ensemble classification model. Through k rounds of training, a sequence of classification models $\{h_1(x), h_2(x), \ldots, h_k(x)\}$ is obtained, which together form a multi-classification model system. The final classification result of this system is determined by a simple majority voting method. The final classification decision is given by:

$$H(x) = \underset{y}{\text{argmax}} \sum_{i=1}^{k} I\left(h_i(x) = y\right) \tag{1}$$

where *H(x)* represents the combined classification model;
h_i denotes a single decision tree classification model;
y represents the output variable (or target variable);
$I(h_i(x))$ is the indicator function.
Equation (1) illustrates the use of majority voting to determine the final classification.

3.2. Feature Extraction

Traditional signal features [11]-[13] commonly include time-domain and frequency-domain features. Time-domain features encompass mean value, variance, standard deviation, and peak value, while frequency-domain features include spectral energy, spectral entropy, and frequency. By reviewing extensive literature [6][16] and analyzing the trend changes in precursor feature signals, this paper introduces four new features: moving average slope, EWMA, EWMA rate of change, and percentage change. To smooth data while preserving necessary details, a sliding window technique is employed with window size of 5. For each set of acoustic emission and electromagnetic radiation signal data, the sliding maximum, sliding minimum, sliding mean, sliding standard deviation, sliding spectral energy, sliding spectral entropy, moving average slope, EWMA, EWMA rate of change, and percentage change are calculated.
The calculation formulas for each parameter are as follows:
Sliding Maximum:

$$Max = max(x_1, x_2, \ldots, x_n) \tag{2}$$

Sliding Minimum:

$$Min = min(x_1, x_2, \ldots, x_n) \tag{3}$$

Sliding Mean:

$$Mean = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4}$$

Sliding Standard Deviation:

$$SD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - Mean)^2} \tag{5}$$

where $x_1, x_2, \ldots, x_n$ represents the values within the sliding window, and n denotes the window size, which is 5.

Sliding Spectral Energy:

Calculated using Fourier Transform (FFT):

$$E = \sum_{k=1}^{N} |X(k)|^2 \tag{6}$$

where $X(k)$ is the spectral amplitude at frequency k.

Sliding Spectral Entropy:

Calculated using Fourier Transform:

$$P_i = \frac{|X(i)|^2}{\sum_{k=1}^{N} X(k)|^2} \tag{7}$$

where $X(i)$ is the normalized spectral amplitude at frequency i.

Moving Average Slope:

Indicates the rate of change of the average value within a given window, helping to capture short-term signal trends:

$$S_t = \frac{\sum_{i=0}^{n-1} i(y_{t-i} - \bar{y}_t)}{\sum_{i=0}^{n-1} i^2} \tag{8}$$

where y_t is a time series, n is the window size, and $\bar{y}_t$ is the average value within the window.

Exponentially Weighted Moving Average (EWMA):

Emphasizes recent observations for data smoothing and trend highlighting:

$$EWMA_t = \alpha \cdot y_t + (1 - \alpha) \cdot EWMA_{t-1} \tag{9}$$

where α is the smoothing constant, $0 \leq \alpha \leq 1$, and y_t is the observation at time t.

EWMA Rate of Change ($\Delta Rate_{EWMA,t}$):

Measures the rate of change of EWMA, reflecting the acceleration of signal changes:

$$\Delta Rate(EWMA, t) = \frac{EWMA_t - EWMA_{t-1}}{EWMA_{t-1}} \tag{10}$$

Percentage Change ($\Delta Rate_{EWMA,t}$):

Describes the percentage change between data points in the time series, aiding in trend intensity identification:

$$\Delta Rate_{per,t} = \frac{y_t - y_{t-1}}{y_{t-1}} \times 100\% \tag{11}$$

3.3. Feature Determination

Point biserial correlation analysis is a statistical method that can be used to study the relationship between two or more variables. It is primarily used to understand the strength and direction of the correlation or relationship between variables. After extracting the features of the signals, point biserial correlation analysis is employed to analyze the precursor feature signals, aiming to identify the time intervals of precursor features in electromagnetic radiation and acoustic emission signals.

The process for calculating the correlation between each feature parameter and the class is as follows:

Step 1: Calculate the means of the quantitative parameter $\bar{X}$ and the dichotomous parameter $\bar{Y}$.

Step 2: Compute the deviation product for each data point between the quantitative parameter and the dichotomous parameter, and sum all the deviation products to get the total sum.

Step 3: Calculate the standard deviations of the quantitative parameter and the dichotomous parameter.

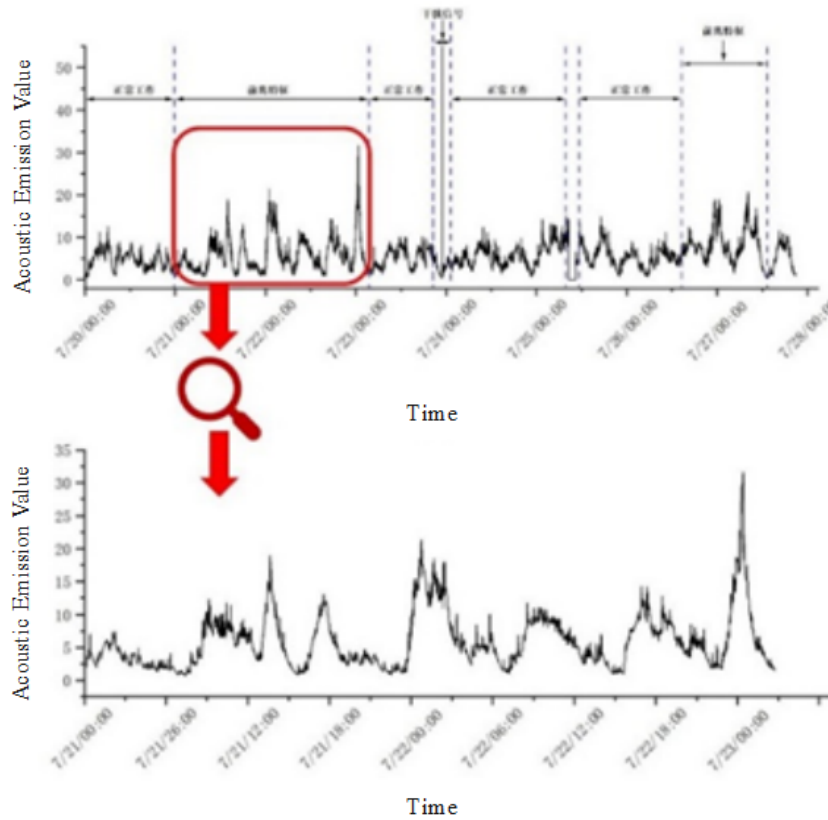**Step 4:** Calculate the point biserial correlation coefficient.

The formula for calculating the point biserial correlation coefficient is as follows:

$$r_{pd} = \frac{\sum(X_i - \bar{X}) \times (Y_i - \bar{Y})}{S_X \times S_Y} \tag{12}$$

where X_i and Y_i represent the values of the quantitative parameter and the dichotomous parameter for the i data point, respectively, and $\bar{X}$ and $\bar{Y}$ are the means of the two variables.

4. Case Analysis

4.1. Data Description and Data Processing

Figure 2: Amplitude Distribution Chart at the Time of Launch

Based on the data detected during actual production at a certain mining site, a dataset of acoustic emission signals and electromagnetic radiation signals was compiled. This data includes the statistical values of acoustic emission signal quantities and electromagnetic radiation signal quantities at different time periods during the production process, as well as the statistical classification of signal categories detected at different time periods. After feature extraction and processing, the following data was obtained.

The following table shows a portion of the AE data after preliminary organization and feature processing:

Table 1. Partially Processed AE Data Table after Feature Processing

Acoustic Emission Intensity (AE)	199.75	197.99	178.599	180.623	200.02	203.4	178.074
Time	2021-11-1 0:04	2021-11-1 0:06	2021-11-1 0:06	2021-11-1 0:08	2021-11-1 0:08	2021-11-1 0:10	2021-11-1 0:10
Sliding Mean	193.6004	192.9124	188.3902	187.9244	191.3964	192.1264	188.1432

Table 1. (continued).

Sliding Standard Deviation	9.875563316	9.296588557	9.740097854	10.11371041	10.81046097	11.60930171	12.47860444
Sliding Minimum	182.66	182.66	178.599	178.599	178.599	178.599	178.074
Sliding Maximum	201.43	201.21	199.75	199.75	200.02	203.4	203.4
Sliding Window Spectral Energy	938003.1395	931244.1174	888220.3815	883912.3743	916983.209	924161.5983	886503.7483
Sliding Window Spectral Entropy	0.187843186	0.191330625	0.200692129	0.198975824	0.203849353	0.198585345	0.232281857
Moving Average Slope	0.81184	-0.1376	-0.90444	-0.09316	0.6944	0.146	-0.79664
EWMA	192.1613143	194.2249917	188.804842	186.0047239	190.7589277	195.0219028	189.3287226
EWMA Rate of Change	0.873957644	0.412735492	-1.084029943	-0.560023635	0.950840777	0.852595019	-1.138636052
Percentage Change	0.093561809	-0.008811014	-0.09793929	0.01133265	0.107389424	0.01689831	-0.124513274

The partially processed electromagnetic radiation signal data is shown in the following table:

Table 2. Partially Processed Electromagnetic Radiation Signal Data Table after Feature Processing

Electromagnetic Radiation (EMR)	Time	Sliding Mean	Sliding Standard Deviation	Sliding Minimum	Sliding Maximum	Sliding Window Spectral Energy
58.15	2020-4-8 0:16	56.336	1.510142377	54.56	58.15	79366.4277
57.92	2020-4-8 0:17	56.894	1.467985014	54.56	58.15	80944.7307
59.62	2020-4-8 0:19	57.538	1.852922017	54.56	59.62	82799.8693
61.24	2020-4-8 0:21	58.874	1.552829675	57.44	61.24	86677.8097
62.65	2020-4-8 0:22	59.916	2.025963968	57.92	62.65	89789.2217
64.76	2020-4-8 0:24	61.238	2.646945409	57.92	64.76	93822.3793
65.38	2020-4-8 0:25	62.73	2.400104164	59.62	65.38	98433.9275
66.8	2020-4-8 0:27	64.166	2.214967268	61.24	66.8	102980.9497

Electromagnetic Radiation (EMR)	Time	Sliding Window Spectral Entropy	Moving Average Slope	EWMA	EWMA Rate of Change	Percentage Change
58.15	2020-4-8 0:16	0.127273947	0.146	56.81090226	0.154218366	0.012360724
57.92	2020-4-8 0:17	0.11336435	0.1116	57.20358426	0.078536402	-0.003955288
59.62	2020-4-8 0:19	0.146416939	0.1288	58.04176051	0.167635249	0.029350829
61.24	2020-4-8 0:21	0.123861834	0.2672	59.13631214	0.218910326	0.02717209
62.65	2020-4-8 0:22	0.140628734	0.2084	60.32821077	0.238379727	0.023024167
64.76	2020-4-8 0:24	0.18175119	0.2644	61.82275227	0.2989083	0.03367917
65.38	2020-4-8 0:25	0.160346281	0.2984	63.01771148	0.238991842	0.009573811
66.8	2020-4-8 0:27	0.151693398	0.2872	64.28498587	0.253454878	0.02171918

Step 1: Given Dataset

First, a dataset of acoustic emission signals and electromagnetic radiation signals is provided, which includes their feature values and target values. Based on the five features identified earlier, the selected feature values are the sliding mean, sliding standard deviation, sliding minimum, sliding maximum, and sliding spectral energy. The signal category is used as the target value, where interference signals are labeled as 1 and non-interference signals as 0.
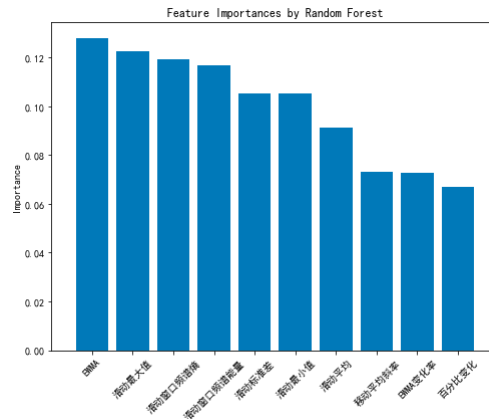
Step 2: Data Preprocessing

Check the dataset for missing and abnormal values. Since the selected features are sliding parameters of the data within their respective windows and the sliding window size is set to 5, the last four rows of the dataset for both acoustic emission and electromagnetic radiation signals contain missing values. These rows should be removed.

Step 3: Splitting Training and Testing Sets
Divide 70% of the data into the training set and 30% into the testing set.

4.2. Feature Selection and Determination

Using point-biserial correlation analysis, it was found that for both electromagnetic radiation and acoustic emission signals, the five metrics—sliding mean, moving average slope, EWMA, EWMA change rate, and percentage change—show a strong correlation with precursor features. Therefore, these five metrics are selected as feature values for both electromagnetic radiation and acoustic emission signals. Consequently, the trend features of the data before the occurrence of danger for both electromagnetic radiation and acoustic emission signals are: sliding mean, moving average slope, EWMA, EWMA change rate, and percentage change.

Figure 3: Feature Importance of Random Forest

4.3. Model Setup

Step 1: Constructing Decision Trees [7-10]:

Random forests are ensemble models consisting of multiple decision trees, so the first step is to construct these trees. A decision tree is a tree-like model used for classification or regression of instances. The construction process of a decision tree is recursive, involving the selection of the best features to split the dataset into different subsets until a stopping condition is met (such as reaching the maximum depth or having fewer samples than a certain threshold in a node).

Step 2: Random Feature Selection:

During the construction of each decision tree, a feature is selected from the set of all features for splitting. To introduce randomness, a subset of features is randomly selected at each node for consideration. This approach ensures that each tree is different, increasing randomness and enhancing the model's generalization ability.

Step 3: Random Sampling of Data:

In building each decision tree, random sampling with replacement is typically performed on the training set to generate different subsets of training data. This method, known as "bootstrap sampling," results in slightly different training datasets for each decision tree, increasing model diversity.

Step 4: Building Multiple Decision Trees:

To create a random forest model, multiple decision trees need to be built and combined to form a robust ensemble model. The number of trees to be constructed can be specified; in this case, we set the number of trees to 100. More trees generally improve the model's stability and accuracy, but this needs to be balanced with computational cost and time.

Step 5: Voting by Decision Trees [17-18]:

When predicting for test samples, each decision tree provides a prediction result. The random forest model aggregates these results through voting to determine the final prediction. The calculation formula is as follows:

$$\hat{y}(x) = argmax_c \sum_{t-1}^{T} P_t(c|x) \tag{13}$$

Among them, $\hat{y}(x)$ represents the predicted category of sample x;

$argmax_c$ is the category with the highest summed probability;

$P_t(c|x)$ denotes the probability that the t decision tree in the random forest predicts sample belongs to category c.

Step 6: Model Tuning

The random forest model has several important hyperparameters that need to be tuned, such as the number of decision trees, the maximum depth of each tree, and the minimum number of samples required at each node. In this step, we use cross-validation to select the optimal hyperparameter combination to enhance the model's performance, ensuring that it maintains good predictive accuracy on unseen data.
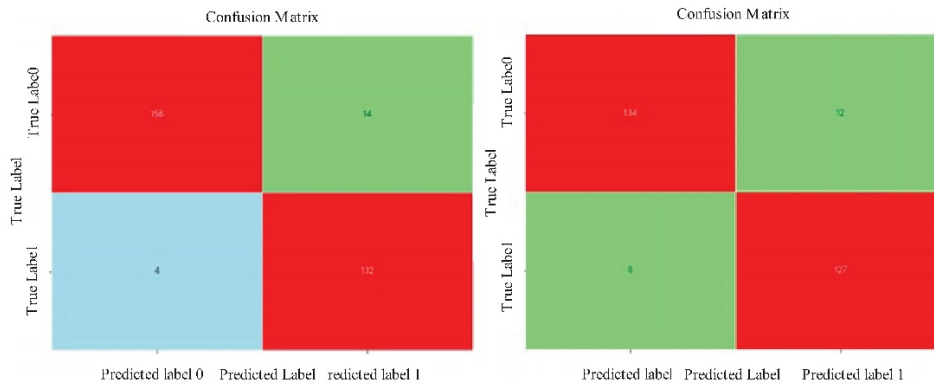
Expressed mathematically:
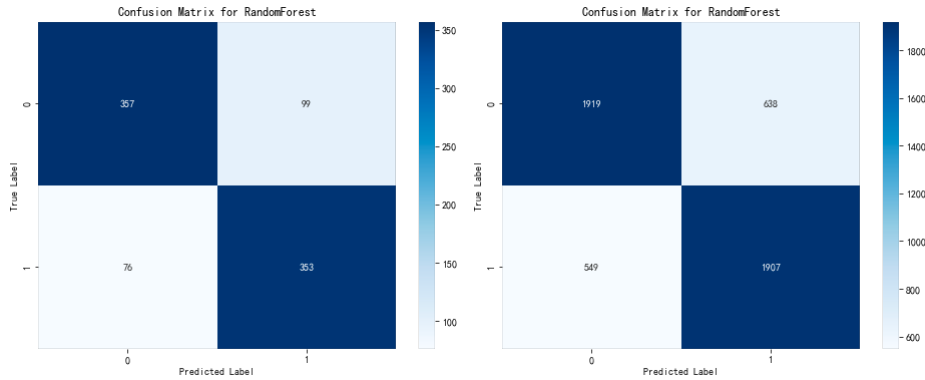
$$RF(x) = mode(h(x_2 k))_{k=1}^{K} \tag{14}$$

where X is the feature space; Y is the target variable (label); $h(X, \Theta_k)$ represents the prediction result of the sample by the k decision tree; Θ_k denotes the parameters of the k decision tree obtained through introducing randomness during the training process; "mode" refers to the majority voting mechanism, meaning the final classification result is determined by selecting the most frequent class label among all trees.

5. Model Solution and Results Explanation

Based on the importance indicators of various features mentioned above, we selected EWMA, sliding maximum, sliding window spectral entropy, sliding window spectral energy, and sliding standard deviation as the feature vectors. The signal class is used as the feature value (where precursor feature signals are labeled as 1 and non-precursor feature signals are labeled as 0). First, we employed down-sampling to balance the sizes of data labeled as 0 and 1. Subsequently, a random forest binary classification model was established, with the dataset split into training and testing sets for model training and evaluation. The resulting confusion matrix is shown below:

Figure 4: Confusion Matrix Chart

To demonstrate that our model has good accuracy and can correctly identify precursor feature signals, we set a continuous window size of 28 when using the model for signal identification. This measure ensures that only when the number of continuously predicted data points exceeds this window length will it be classified as a precursor feature signal. Using this method, we successfully identified the time intervals of the first five precursor feature signals in electromagnetic radiation and acoustic emission signals. The specific results are shown in the tables below:

Table 3. Time Intervals of Electromagnetic Radiation Precursors

No	Start Time Interval	End Time Interval
1	2020-04-08 00:16:48	2020-04-08 00:39:05
2	2020-04-08 00:55:32	2020-04-08 01:11:02
3	2020-04-08 01:19:29	2020-04-08 02:09:50
4	2020-04-08 03:42:48	2020-04-08 04:04:06
5	2020-04-08 04:39:56	2020-04-08 04:54:28

Table 4. Time Intervals of Acoustic Emission Precursors

No	Start Time Interval	End Time Interval
1	2021-11-01 14:40:22	2021-11-01 15:03:55
2	2021-11-26 05:10:49	2021-11-26 05:27:07
3	2021-12-07 03:09:43	2021-12-07 04:02:15
4	2022-01-04 04:08:19	2022-01-04 05:36:07
5	2022-01-04 05:39:45	2022-01-04 06:10:51

Based on the binary random forest model established above, the partial results are calculated as follows:

Table 5. Predicted and Actual Values of Acoustic Intensity Signals

	Moving Average	Moving Std Dev	Moving Min	Moving Max	Moving Window Spectral Energy	Moving Window Spectral Entropy	Moving Average Slope	EWMA	EWMA Rate of Change	Percentage Change	RF_Predictions	True Label
206657	32.9360	0.212673	32.820	33.310	27119.954700	0.039099	0.00520	32.886401	-0.006640	0.000000	0	0
340414	38.9638	0.433776	38.271	39.298	37956.324378	0.061182	-0.03228	38.827723	-0.055672	-0.022627	1	1
340346	40.9210	0.345143	40.603	41.400	41864.397260	0.044562	0.00812	41.022850	0.037715	0.018100	1	1
194080	33.8366	0.503763	33.372	34.698	28625.425257	0.078400	-0.04204	33.731197	-0.035920	-0.010672	0	0
312905	33.4422	0.420769	32.809	33.886	27961.288988	0.066813	-0.02564	33.345023	-0.053602	-0.025224	1	0
537438	20.9664	0.117651	20.832	21.095	10989.886642	0.034817	0.01044	20.975322	0.011968	0.000854	0	0

Table 5. (continued).

340372	39.4242	0.367022	38.993	39.779	38858.035693	0.051942	0.01784	39.457979	0.032102	0.020157	1	1
416059	34.0566	0.106824	33.901	34.187	28996.414202	0.018525	0.00652	34.077195	0.010980	0.004318	0	0
453955	36.6240	0.521181	35.990	37.260	33535.650700	0.056102	-0.01440	36.717905	0.004210	-0.004064	1	1
66507	37.6000	3.286335	33.000	42.000	35452.000000	0.310272	0.04000	37.174835	-0.017484	0.000000	0	0

Table 6. Predicted and Actual Values of Electromagnetic Radiation Signals

	Moving Average	Moving Std Dev	Moving Min	Moving Max	Moving Window Spectral Energy	Moving Window Spectral Entropy	Moving Average Slope	EWMA	EWMA Rate of Change	Percentage Change	RF_ Predictions	True Label
559691	13.2420	0.446621	12.480	13.630	4.385759e+03	0.152540	-0.02280	13.108247	-0.062825	-0.084373	0	0
337753	29.8094	0.389955	29.429	30.373	2.221653e+04	0.063330	-0.02112	29.755067	-0.032607	-0.009091	0	0
503311	19.6580	0.136638	19.450	19.800	9.661111e+03	0.038407	-0.01560	19.743003	-0.001300	-0.003535	1	1
108607	33.6000	2.607681	29.000	35.000	2.829200e+04	0.283935	0.16000	33.282157	0.171784	0.206897	0	0
95518	456.0000	9.746794	445.000	471.000	5.199350e+06	0.106155	-0.76000	456.220167	0.277983	0.015487	1	1
141136	97.5492	0.000447	97.549	97.550	2.378962e+05	0.000058	-0.00008	97.549544	-0.000054	0.000000	1	1
479648	21.0340	0.462255	20.240	21.350	1.106287e+04	0.108054	0.03360	21.009860	0.001014	-0.014071	1	1
73555	53.2000	3.271085	49.000	58.000	7.086300e+04	0.226430	0.12000	53.143129	-0.114313	-0.018868	0	1
595385	10.0000	0.707107	9.000	11.000	2.505000e+03	0.259946	0.00000	9.663877	-0.066388	-0.100000	0	0
105699	30.4000	1.516575	28.000	32.000	2.312700e+04	0.192029	0.00000	30.757716	0.124228	0.032258	0	0

Prediction Results:

```
Results for RandomForest:
Accuracy: 0.8022598870056498
Classification Report:
              precision    recall  f1-score   support

           0       0.82      0.78      0.80       456
           1       0.78      0.82      0.80       429

    accuracy                           0.80       885
   macro avg       0.80      0.80      0.80       885
weighted avg       0.80      0.80      0.80       885
```

Figure 6: Prediction Results Chart

6. Conclusion

In summary, this paper proposes a signal recognition and prediction system based on the random forest model, applied to the hazard prediction of dynamic pressure in deep coal mining. Through the extraction and analysis of features from electromagnetic radiation (EMR) and acoustic emission (AE) signal data, we successfully constructed a random forest binary classification model to identify interference signals and precursor feature signals. To enhance the accuracy and reliability of predictions, we introduced new features such as moving average slope and exponential weighted moving average (EWMA), and employed down-sampling techniques to balance the data categories.

Overall, the proposed method not only effectively improves the accuracy of dynamic pressure prediction but also provides reliable data support for coal mine safety management. Future research

could further optimize model parameters, explore additional factors affecting dynamic pressure, and validate the model's applicability under different mining conditions to enhance the system's generalization ability and practical value.

References

[1] Wu, B. (2024). Study on the prediction model for dynamic pressure hazard based on supervised learning [Doctoral dissertation, Anhui University of Science and Technology].

[2] Lu, C., et al. (2005). Spectrum analysis and signal recognition of rock mass microseismic monitoring. Journal of Geotechnical Engineering, 27(7), 772-775.

[3] Zhou, X., He, X., & Zheng, C. (2019). Radio signal recognition based on deep learning in images. Journal on Communication/Tongxin Xuebao, 40(7).

[4] Mai, Q., & Wu, X. (2024). Current status of coal mine dynamic pressure hazard prediction and monitoring technology. Shaanxi Coal, 43(01), 87-92.

[5] Liang, Y., Shen, F., Xie, Z., et al. (2023). Research on dynamic pressure prediction method based on LSTM model. China Mining, 32(05), 88-95.

[6] Di, Y. (2023). Research on comprehensive early warning for dynamic pressure based on deep learning [Doctoral dissertation, China University of Mining and Technology].

[7] Fang, K., Wu, J., Zhu, J., et al. (2011). A review of random forest methods. Statistical & Information Forum, 26(03), 32-38.

[8] Zhang, Z., et al. (2009). A review of radar radiation source signal recognition. Ship Electronic Engineering, 4, 10-14.

[9] Rigatti, S. J. (2017). Random forest. Journal of Insurance Medicine, 47(1), 31-39.

[10] Speiser, J. L., et al. (2019). A comparison of random forest variable selection methods for classification prediction modeling. Expert Systems with Applications, 134, 93-101.

[11] Zhong, Z., & Li, H. (2020). Recognition and prediction of ground vibration signals based on machine learning algorithms. Neural Computing and Applications, 32(7), 1937-1947.

[12] Zhao, D., & Yan, J. (2011). Performance prediction methodology based on pattern recognition. Signal Processing, 91(9), 2194-2203.

[13] Keenan, R. J., et al. (2001). The signal recognition particle. Annual Review of Biochemistry, 70(1), 755-775.

[14] Yuan, R., Li, H., & Li, H. (2012). Distribution characteristics and precursor information discrimination of coal pillar-type dynamic pressure microseismic signals. Journal of Rock Mechanics and Engineering, 31(01), 80-85.

[15] Zhang, J., et al. (2019). An automatic recognition method of microseismic signals based on EEMD-SVD and ELM. Computers and Geosciences, 133, 104318.

[16] Zhang, J., Jiang, R., Li, B., et al. (2019). An automatic recognition method of microseismic signals based on EEMD-SVD and ELM. Computers and Geosciences, 133, 104318.

[17] Probst, P., & Boulesteix, A. L. (2018). To tune or not to tune the number of trees in random forest. Journal of Machine Learning Research, 18(181), 1-18.

[18] Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(3), e1301.