Advances in Computer, Signals and Systems (2024)
Clausius Scientific Press, Canada
DOI: 10.23977/acss.2024.080610
ISSN 2371-8838 Vol. 8 Num. 6

Research on the prediction of impact ground pressure hazard in deep coal mining based on moving average method

Bingbing Hou

School of Mathematics and Statistics, Guangxi Normal University, Guilin, China

Keywords: Moving Average Method, Decision Tree, Support Vector Machine, Impact Ground pressure

Abstract: As the mining depth of underground increases, the ground stress increases, which inevitably leads to an increase in the probability of impact ground pressure. The hidden danger of impact ground pressure seriously affects the safe and efficient mining of coal mines, so the early warning of impact ground pressure has an important role. In this paper, the identification and prediction of precursor characteristic signals of impact ground pressure are realized by moving average method, decision tree and support vector machine. The data are preprocessed by removing noise signals and normalization, extracting the "Class C" and "non-Class C" features of the preprocessed data, and adjusting the parameters to establish and optimize the interference signal recognition model based on the classification of the feature tree, and applying the model to identify the interference signal and determine the interference signal. The model is used to identify the interfering signals and determine the time interval of the interfering signals. Based on the feature tree classification algorithm of particle swarm optimization, the precursor feature signal identification model is established and applied to identify the precursor feature signals and determine their time intervals, and finally the feature tree algorithm is used to predict and analyze the probability of the appearance of precursor features.

1. Introduction

Coal, as one of the major sources of energy, supports China's industrial production and economic development. Abundant coal resources can meet domestic demand and reduce energy import dependence to a certain extent. The coal industry also creates a large number of jobs and promotes local economic development[1].However, with the increase of mining depth, the ground stress increases, and the risk of underground coal-rock power disaster is getting bigger and bigger, which seriously affects the safe and efficient mining of coal mines, and the coal safety problem cannot be ignored. Among them, the problem of impact ground pressure is gradually getting people's attention, and the monitoring and early warning and effective prevention and control of impact ground pressure is an important part of coal mine safety to be overcome at this stage[2].Research on the risk prediction of impact ground pressure in deep coal mining is being actively carried out. Through detailed investigation and analysis of the geological structure of the mine, the nature of the rock strata, and the parameters of the ground stress, it is possible to preliminarily determine the potential hazardous area of the ground pressure and predict the trend of the change of the ground pressure. At present, the research on hazard prediction of impact ground pressure in deep coal mining involves many fields and various methods, and it is constantly advancing the level of technology and prediction accuracy to ensure the safety of miners and the stable operation of mines.

2. Extract overall discrete data features

Initially, it is judged that there are also large differences in the mean and variance of continuous type numerical features. In turn, it is judged that there is also a difference between the features after the possible wavelet transform and the Fourier transform, as shown in Figure 1.

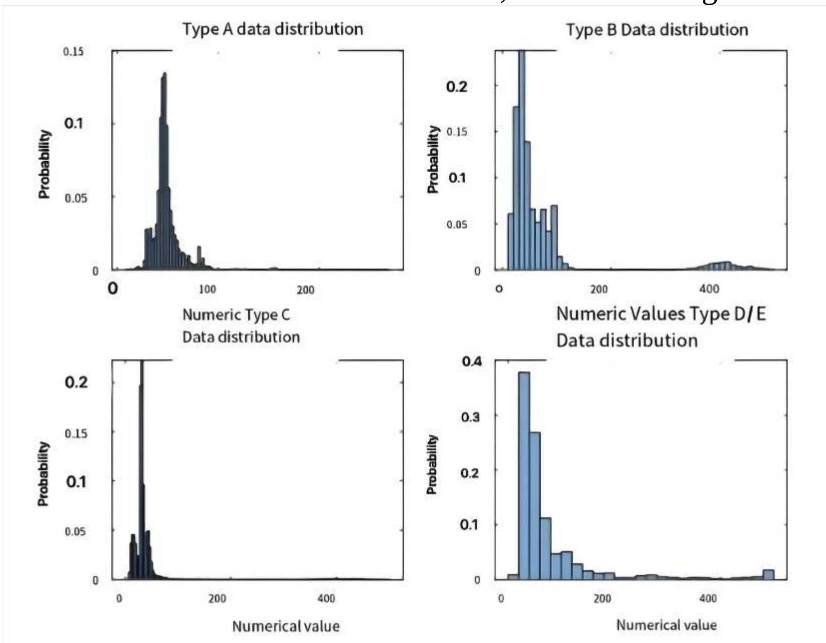

Figure 1: Distribution of type data

2.1 Extracting features of continuous data under a sliding window

Considering that the duration of anomalous conditions is not very short, there will be some continuity in the features. Therefore, we use the moving average method,[3] to set the window size to 20 and the step size to 1 for continuous data feature extraction. Considering that there is a large time jump between the C class data. In this paper, more than 20 consecutive neighboring class C data are merged into multiple groups. The AE and EMR intensities vary with the time signal. But there may exist multiple different signals at the same time. It is conjectured that there are multiple sensors on the site. Assume that there is a lag in the chip or timer of the sensors that does not strictly require an interval of 30 seconds. Therefore the relationship between time points and signal strength can be converted into the relationship between sequence points and signal strength. In this paper, we can ignore the time data and utilize the sequence points instead of the time point operation, as shown in Figure 2.

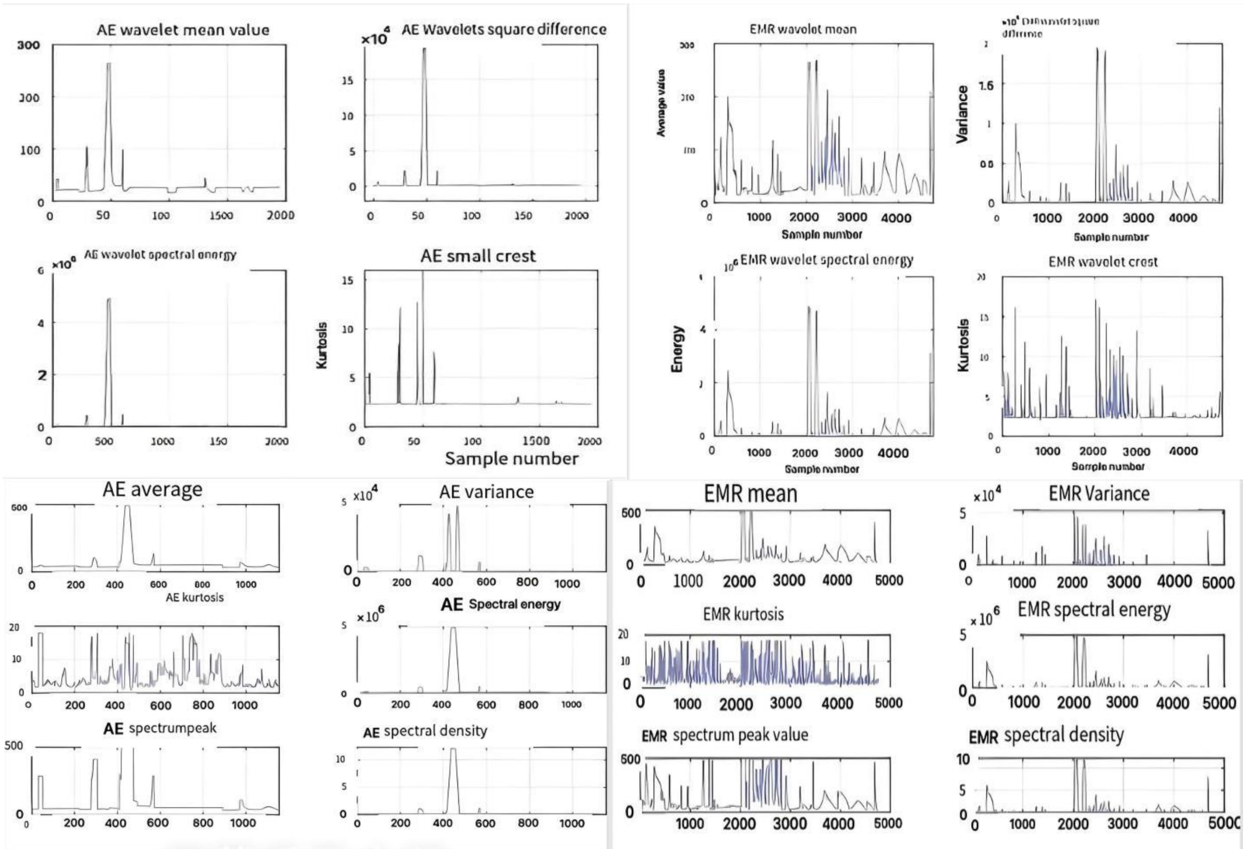

Figure 2: Visualization of the features of AE and EMR

(math.) a Fourier transform

$$F(w) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \tag{1}$$

wavelet transform (math.)

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt \tag{2}$$

Capture local features of the signal such as breakpoints, singularities, etc.
kurtosis

$$Kurtosis = \frac{N(N+1)}{(N-1)(N-2)(N-3)} \sum_{i=1}^{N} \left(\frac{x_i - Mean}{Standard\ Deviation}\right)^4 - \frac{3(N-1)^2}{(N-2)(N-3)} \tag{3}$$

A high kurtosis indicates that the data has heavy tails or many outliers, and a low kurtosis indicates that the data distribution is flatter than a normal distribution.

Spectral energy

$$Spectral\ Energy = \sum |FFT(x_i)|^2 \tag{4}$$

It can be used to evaluate the energy distribution of a signal at various frequencies and is commonly used for energy analysis and feature extraction of signals.

In this paper, the "non-C class" data are extracted using the moving average method for all the windows, the window size is 20, and the step size is 1. The results include 10 features, including wavelet mean, wavelet variance, wavelet energy, wavelet kurtosis, variance, etc., of the "non-C class" data, and a feature matrix is obtained. The result includes 10 features such as wavelet mean, wavelet variance, wavelet energy, wavelet kurtosis, variance, etc. The feature matrix is obtained. And add 1 column of number 5 as the label of "non-C class" data.

2.2 Applying the category weighting method for preprocessing

Suppose there are Ncategories, the number of samples in each category is respectively n 1, n 2,... , nN, then the weight of the category w_i is

$$w_i = \frac{N \times n_{total}}{n_i \times \sum_{j=1}^{N} n_j} \tag{5}$$

Where N is the total number of categories, andn_{total} is the total number of samples from all categories, and n_i is the number of samples in the i number of samples in the first category.

The weights calculated in this way ensure that categories with fewer samples receive higher weights in model training, thus balancing the unbalanced dataset.

2.3 Modeling based on decision tree classification algorithm

Information entropy is a measure of the uncertainty of a random variable in a data set. In decision trees, entropy is used to measure the impurity of the data at a given node[5]. The mathematical formula is

$$H(S) = -\sum_{x \in X} p(x) log_2 p(x) \tag{6}$$

Information gain refers to the decrease in entropy or increase in information obtained after selecting a feature for data segmentation. The formula is calculated as

$$Gain(S, A) = H(S) - \sum_{t \in T} \frac{|S_t|}{|S|} H(S_t) \tag{7}$$

Where S is the current dataset, A is the attribute, and T is the subset after partitioning A by the attribute S.

First data preparation is done by collecting datasets containing features and labels (categories). Then missing values, outliers and duplicates are processed to ensure data quality. Finally feature engineering is performed to try to introduce new features or perform feature transformation to improve the model performance.

3. Model output results

3.1 Model parameters

The model parameters are shown in Table 1.

Table 1: Model parameters

The function calculates the total number of times	Total time/s	Total objective function calculation time/s	Estimated best practicable point	Observed best feasible point
30	11.2487	1.3524	952	1068

3.2 Decision Tree Structure

Decision tree structure obtained after training with clear branching. It can assist in determining the type of data. As shown in Fig. 3.

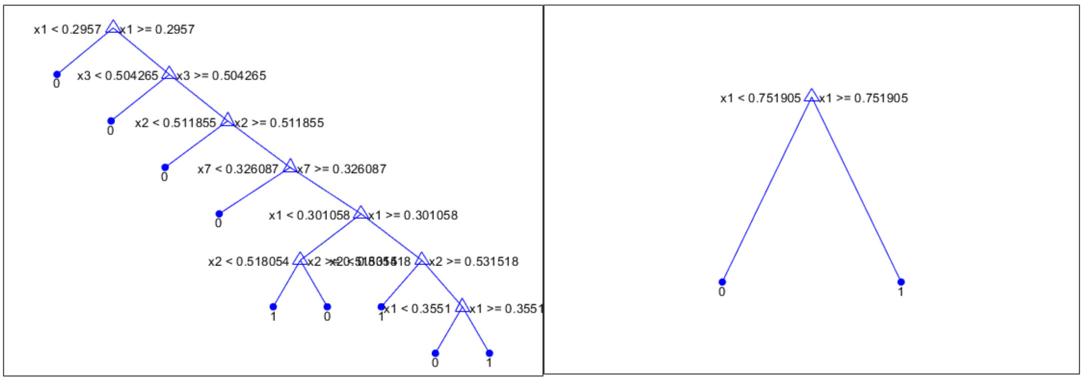

Figure 3: Decision tree structure for EMR and AE

3.3 Results of model evaluation

The results of the model evaluation are shown in Table 2.

Table 2: Results of model evaluation

Accuracy	Precision	F1 Score	AUC
99.98%	1.00	1.00	1.00

3.4 ROC curve evaluation model accuracy

The Receiver Operating Characteristic curve is a graph used to evaluate the performance of a classification model. It is plotted with False Positive Rate, FPR, on the horizontal axis and True Positive Rate, TPR, on the vertical axis, FPR being the proportion of samples that are actually negative but incorrectly predicted to be positive. TPR, also known as recall or sensitivity, refers to the proportion of samples that are actually positive but are correctly predicted to be positive. As can be seen from the ROC curve below, the predictions are relatively accurate, as shown in Figure 4.

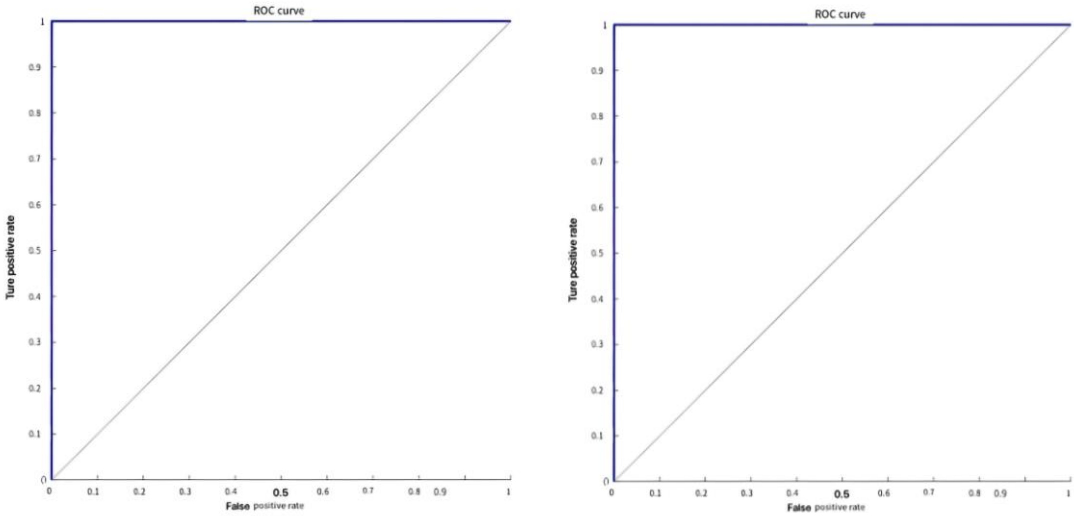

Figure 4: ROC curves for EMR and AE

3.5 Feature tree objective function modeling

A target feature model is a model used for supervised learning tasks, which is mainly used to predict one or more target features, also known as dependent variables or labels, based on the input features or independent variables, to get the classification results of labeled classes. In this paper, the feature values are divided as input features and trained to get the cloth labeling function model. The model prediction curve fits well with the real curve. It shows that the model has high performance. The model can be used to classify the results for prediction, as shown in Fig. 5.

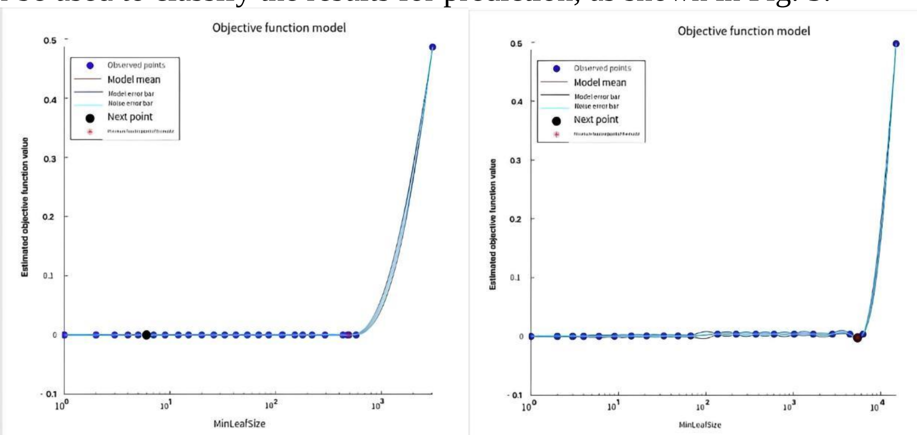

Figure 5: Objective function model of EMR and AE

4. Modeling of decision tree classification algorithm based on particle swarm optimization algorithm

4.1 Extraction of continuous data features (moving average method)

Numerical analysis of the grouped table is performed using the moving average method, setting the window size bit 20 and the step size to 1. The idea is the same as 5.1.2, obtain 10 features. Construct the feature matrix, and add the label column data 2 to the last column of the matrix to indicate the label class B data, as shown in Figure 6.

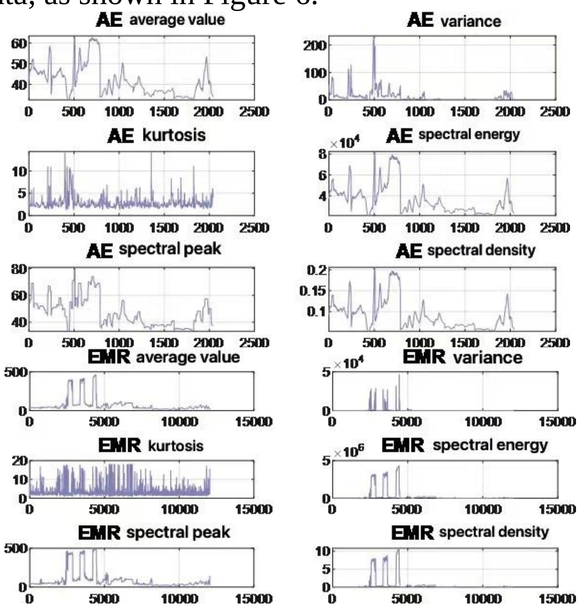

Figure 6: Visualization of the features of AE and EMR

4.2 Modeling of decision tree classification algorithm based on particle swarm optimization algorithm

The decision tree classification model of Particle Swarm Optimization is a model that combines the classification ability of decision trees with the optimization mechanism of PSO. This model is particularly suitable for finding the optimal or near-optimal decision tree structure on complex datasets to improve the classification performance. The following is the mathematical modeling process and related mathematical formulas for the establishment of this model[6].

4.2.1 Overview of Particle Swarm Optimization (PSO) Algorithm

Particle swarm optimization is an evolutionary computing technique, based on population intelligence, to solve optimization problems by simulating the social behavior of bird flocks[7].Suppose there are N categories and the number of samples in each category is n1,n2,... ,nN then the weight of the category is

$$\mathrm{w_i} = \frac{\mathrm{N} \times \mathrm{n_{total}}}{\mathrm{n_i} \times \sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{n_j}} \tag{8}$$

Where, N is the total number of categories. $\mathrm{n_{total}}$ is the total number of samples from all categories. n_i is the number of samples from the i-th category[8].

The weights thus computed ensure that categories with fewer samples are given higher weights in model training, thus balancing the unbalanced dataset. Each solution (particle) moves through the solution space, updating its speed and position by tracking the individual historical best position (individual optimum, pbest) and the group historical best position (global optimum, gbest).

4.3 Output results

4.3.1 Model parameters

The model parameters are shown in Table 3

Table 3: Model parameters

particle count	algebraic	Total time/s	Total objective function calculation time/s	Estimated best practicable point	Observed best feasible point
5	30	125.3279	5.3524	904	997

4.3.2 Decision tree structure

The following is the structure of the decision tree obtained after training[9].The nodes are clear and the branches are well defined. It can assist in determining the type of data. As shown in Fig. 7.

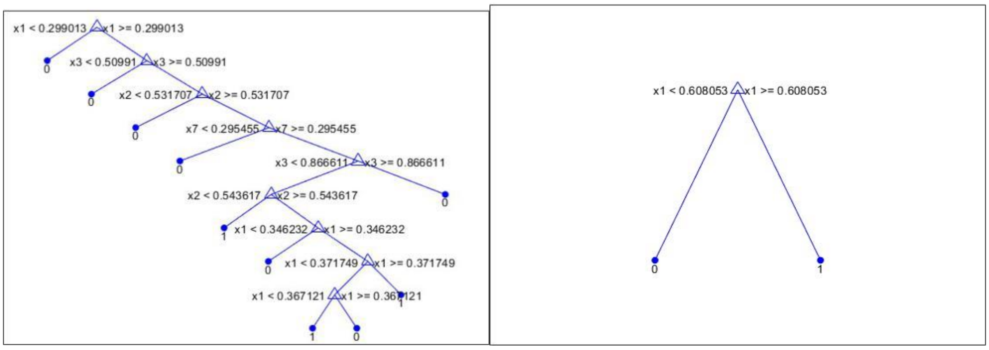

Figure 7: Decision tree structure for EMR and AE

4.4 Feature Tree Objective Function Modeling under Particle Swarm Optimization

The paper divide the eigenvalues selected by particle swarm optimization as input features and train them to get the cloth-scaled function model. The model prediction curve fits well with the real curve. It shows that the model has high performance. The model can be used to classify the results for prediction, as shown in Fig. 8[10].

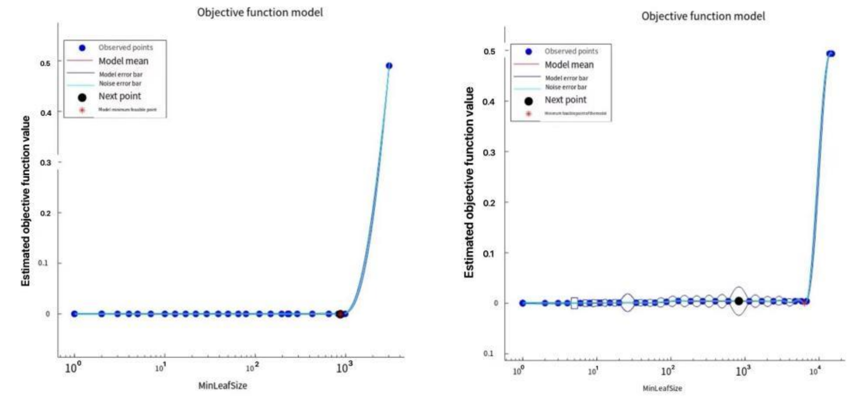

Figure 8: Objective function model for EMR and AE

5. Conclusions

In this paper, continuous similar samples are extracted and grouped to reduce the influence of eigenvalues analyzed from discontinuous similar samples on the model. The model uses wavelet transform on the imported data to minimize the interference of noise on the data analysis. During the iteration process, the decision tree may find that a feature value is a better segmentation attribute than another feature, and therefore, the position of the particles is updated to a new decision tree structure with the better feature value as the root node. Through continuous iteration, the model will try to find the most optimized decision tree structure, which will improve the accuracy of classification.

The moving average method can effectively reduce noise and identify potential trends by smoothing the data, which is very helpful for predicting the risk of rock burst in deep mining of coal mines. At the same time, the moving average method is relatively simple, easy to implement and understand. By calculating the average value of coal mining data in different time periods, the moving average method can identify the trend and periodic changes in the mining process, which is helpful to judge the potential danger of ground pressure.

References

[1] Bouaoud A , Lebey M , Wang H ,et al.Forecasting Flashover in a Compartment Fire by Using a Real-time Moving Average from Temperature Recordings[J].Journal of Civil Engineering and Construction, 2023.DOI:10.32732/ jcec. 2023.12.1.55.

[2] Kaharuddin S , Ain M F , Mamat M N ,et al. Evaluating the Effectiveness of Exponentially Weighted Moving Average Filter in Enhancing Landslide Detection from Accelerometer Data[C]//Symposium on Intelligent Manufacturing and Mechatronics.Springer, Singapore, 2024.DOI:10.1007/978-981-97-0169-8_59.

[3] Wu J, Grande G, Triolo F. Sjoberg L. Rizzuto D.Pyko A.Ljungman P.Bellander T.Eneroth K.Air pollution, social engagement, and depression in older adults: Results from a Swedish population-based cohort study [J].Environmental Pollution, 2023, 336(Nov.):1.1-1.10.

[4] Meng Fanxiao, Zhang Yuran, Quan Ranran, Zhao Nan. Application of weighted moving average and curve fitting model in settlement monitoring [A]. Mapping and spatial geographic information, 2023, 8(46):52-54.

[5] Xiao L, Zeng Y. Turbofan engine RUL prediction based on decision tree feature extraction with RSM___LightGBM [A]. Engineering Science and Technology II Series; Agricultural Science and Technology, V235.13. Agricultural Equipment and Vehicle Engineering, 2024, 4(62): 133-135.

[6] He Xiucheng, Qin Dongdong, Zhao Junjie, Tian Keqin, He Youjian. Design of yaw control system for wind turbine based on decision tree algorithm [A]. Machinery and electronics, 2024, 6(42):78-80.

[7] Yang Fan, Wu Jingxiu, Fan Ziwu, Li Zixiang, Zhu Shentao. Fast integrated learning particle swarm optimization algorithm [A]. Water conservancy and hydropower technology (Chinese and English), 2024, 6 (28):3-15.

[8] G. P. Xu, J. Yu, B.Y. Wu, et al. Wind turbine gearbox fault feature extraction and diagnosis based on Zoom-FFT-CEEMD and wavelet packet noise reduction [A]. Power Industry. tm315. Microtronics, 2024, 4(52):29-32.

[9] Guo Zizhe, Xu Shuangfei, Zhou Hui et al. Cost-benefit analysis of 13-valent pneumococcal polysaccharide conjugate vaccine for inclusion in the immunization program of Qinghai Province based on decision tree-Markov model [A]. Medicine and Health Science and Technology, R186. China Vaccine and Immunization, 2024, 2(30):199-203.

[10] Men Hong, Zhao Huaquan, Dou Guidong, Xie Fei. Evaluation and prevention of impact ground pressure hazard based on seismic wave CT detection [B].Shanxi coal, 2024, 6(43):83-86.