Article

Research on Coal and Gas Outburst Risk Warning Based on Multiple Algorithm Fusion

Yanlei Guo [1,2], Haibin Liu [1,*], Xu Zhou [2] , Jian Chen [2] and Liwen Guo [2]

1 School of Management, China University of Mining and Technology-Beijing, Beijing 100083, China; bqt1900502012@student.cumtb.edu.cn
2 School of Emergency Management and Safety Engineering, North China University of Science and Technology, Tangshan 063000, China; sxzhouxu@126.com (X.Z.); chenjian@ncst.edu.cn (J.C.); guoliwen64@163.com (L.G.)
* Correspondence: hbliu@cumtb.edu.cn

Featured Application: The XGBoost–GR–stacking gas outburst early warning model established in this article demonstrates high accuracy and practical performance, making it suitable for gas outburst risk warning in mining safety.

Abstract: To improve the accuracy of gas outburst early warning, this paper proposes a gas outburst risk warning model based on XGBoost–GR–stacking. The statistic is based on gas outburst data from 26 mines and establishes a data generation model based on XGBoost. The obtained virtual datasets are analyzed through visualization analysis and ROC curve analysis with respect to the original data. If the augmented data has an ROC area under the curve of 1, it indicates good predictive performance of the augmented data. Grey correlation analysis is used to calculate the grey correlation degrees between each indicator and the "gas emission". The indicator groups with correlation degrees greater than 0.670 are selected as the main control factor groups based on the sorting of correlation degrees. In this study, SVM, RF, XGBoost, and GBDT are selected as the original models for stacking. The original data and virtual data with correlation degrees greater than 0.670 are used as inputs for SVM, RF, XGBoost, GBDT, and stacking fusion models. The results show that the stacking fusion model has an MAE, MSE, and R2 of 0.031, 0.031, and 0.981. Comparing the actual and predicted values for each model, the stacking fusion model achieves the highest accuracy in gas outburst prediction and the best model fitting effect.

Keywords: directional splitting; damage region; coalbed methane mining; coalbed permeability enhancement

1. Introduction

Coal and gas outbursts are a frequent and dangerous occurrence in the coal mining industry. These accidents pose a significant threat to both the equipment used in coal mining and the safety of miners. As mining operations continue to deepen and intensify, the frequency of coal and gas outburst accidents has been on the rise, greatly impacting the safety of coal mine production. In fact, China has experienced a significant number of these accidents, accounting for over 40% of the global total. As of 2022, approximately 33.6% [1] of Chinese coal mines are classified as high-risk mines prone to coal and gas outbursts. To address these challenges and support the goals outlined in the national "14th Five-Year Plan", experts and scholars have been conducting research and analysis to understand the mechanisms and risks involved in coal and gas outbursts. Their aim is to predict and prevent such accidents in order to ensure the safe production of coal mines and contribute to the sustainable development of China's energy sector.

Coal and gas outburst accidents are the result of complex phenomena, driven by the dynamic interaction of coal and gas within the mine. This involves the uncontrolled

evolution of various nonlinear factors, making it a highly destructive gas dynamic phenomenon [2]. In studying the mechanisms behind these outbursts, scholars have examined multiple factors such as ground stress, gas properties, and coal mechanical properties. They have constructed theoretical models based on traditional algorithms, although traditional algorithms often struggle to analyze the nonlinear relationship between these factors. With advancements in intelligent algorithms and machine learning, more and more experts and scholars are utilizing these technologies to analyze coal and gas outburst risks, yielding promising results. The challenge lies in uncovering the nonlinear relationships between various influencing factors, accurately predicting the risk and severity of coal seam gas outbursts, and implementing early warning systems to prevent or mitigate the disasters caused by these outbursts. This has become a critical task in ensuring the safe production of mines [3].

Over the past century, experts and scholars from both domestic and international institutions have conducted extensive research on the mechanism of coal and gas outbursts and have proposed numerous hypotheses. However, there is still no unified theory that can completely reveal the development mechanism of coal and gas outbursts. The theory proposed by Pingping Ye [4] analyzed the mechanism of pore pressure on coal and conducted deformation tests of coal rock under cyclic loading and unloading of pore pressure. Norbert [5] derived a relationship model between coal porosity and mechanical strength and used porosity and gas pressure to classify and predict the level of outburst hazards. Zhao [6] through theoretical analysis, pointed out that structurally weak coal has a lower bearing capacity, and the fine coal particles formed after fragmentation have an extremely fast gas desorption rate. This can sustain the development of outbursts, making structurally weak coal seams prone to coal and gas outbursts. Wold [7] suggested that coal and gas outbursts are influenced by factors such as gas pressure, gas composition, coal permeability, and adsorption desorption characteristics, and analyzed the relationship between these influencing factors and the control of outbursts. Dazhao Li [8] proposed a coal and gas outburst support model, and analyzed the mechanism of non synchronous deformation induced coal and gas outburst in soft and hard layer. Chaolin Zhang [9] systematically summarized the research progress on coal and gas outburst mechanisms in China from three aspects: theoretical hypothesis, physical simulation, and numerical simulation. Hu Qianting [10] described in detail the entire process of occurrence and development of outbursts based on numerical simulations and theoretical analysis. The process was ultimately divided into four stages: preparation, initiation, development, and termination. Lijun Zhao [11] summarized the research progress on the mechanism of coal and gas outburst, and analyzed the shortcomings of existing theoretical models. Liangcheng Wang [12] and Shoujian Peng [13], based on theoretical exploration combined with a large number of numerical simulation analyses, discussed the evolution process of coal and gas outbursts. Guo Pinkun [14], combining experimental research, established a model for the development of layer fractures during the outburst process and explored the mechanism of layer fracture development. Xu Mangui [15] and others constructed a microelement model of coal-rock mass and believed that the destruction of coal and gas microelements is the primary cause of outbursts.

With the increasing popularity of data analysis and data science theory, numerous experts and scholars have extensively analyzed various factors related to coal and gas outbursts using traditional algorithm and mathematical analysis methods. Dan Dakuo [16] combined mathematical and statistical analysis methods to determine the prediction indicators and critical values for coal and gas outbursts, with gas as the dominant factor. This achieved the prediction of the risk level of coal and gas outbursts. Si Hu [17] extracted 27 factors that influence the occurrence of coal and gas outburst accidents. By using association analysis and cross-coupling analysis, they conducted statistical analysis and in-depth exploration of coal and gas outburst accidents of average and above average severity that had occurred in the last 15 years. Wen Changping [18] constructed attribute measurement functions to calculate single-index attribute measurements and comprehensive sample

attribute measurements. By applying confidence criteria, they conducted attribute recognition of gas outbursts in tunnel samples and established an attribute recognition model for gas outburst evaluation in the tunnel survey and design stage. Wang Gang [19] analyzed the energy relationship in the process of coal and gas outbursts using the energy method. They obtained the relationship between the energy conditions of coal and gas outbursts, coal seam geostress, cohesion coefficient, coal seam thickness, and the risk of gas outburst accidents. Cao Shugang [20] conducted experiments on the adsorption desorption deformation process of outburst-prone coal under different gas pressure conditions. They found a good power-function relationship and quadratic function relationship between the desorption shrinkage deformation of coal samples and the original gas pressure. Dingding Y [21] studied the influence of temperature on the "energy-mass" characteristics of gas and discovered the function relationship between the initial gas expansion energy released and temperature under different conditions, improving the prediction indicators for outburst hazards. Li Yunbo [22] studied the initial gas desorption velocity and amount of gas-prone coal and structurally weak coal using a self-made gas desorption experimental apparatus. They analyzed and established mathematical models for the influencing factors during the initial gas desorption period of structurally weak coal. They concluded that the initial desorption velocity of gas exhibits a power-law relationship with adsorption equilibrium pressure, and that the initial desorption curve of structurally weak coal conforms to the Vent formula.

Based on mathematical theory and machine learning, predictive methods have shown a high degree of adaptability to the complex problem of coal and gas outbursts, which involve non-linear relationships among various factors. An increasing number of experts and scholars are adopting intelligent algorithms to predict coal and gas outbursts and they have achieved a certain amount of success. Xiang Zeng Du [23] used a grey comprehensive correlation analysis model to quantitatively analyze six predictive indicators of coal and gas outbursts and determine the optimal prediction indicators. This provides a quantitative basis for the selection of prediction indicators for coal and gas outbursts. Zhou Xihua [24] predicted the intensity of coal and gas outbursts using an RBF neural network model and principal component analysis, ultimately achieving high prediction accuracy. Liu Xiaoyue [25], Cao Bo [26], Ren Shaowei [27], and others used the BP neural network to predict coal and gas outbursts and optimized the dimensions of the influencing factors through principal component analysis. By reducing the correlation among variables and selecting the main control factors, they improved the prediction efficiency of the entire model. The optimized models also achieved high prediction accuracy. Zhao Huatian [28], Zeng Weishun [29], Zhang Wenjuan [30], and others made full use of support vector machines (SVMs) to address the advantages of solving small sample problems and combined them with other optimization algorithms to predict coal and gas outbursts. Zhao Huatian and Zeng Weishun used particle swarm optimization to optimize SVM, searching for global optimal solutions from a global perspective and greatly reducing the probability of local optimal solutions. Zhang Wenjuan utilized the least squares method to optimize SVM, effectively removing noise from gas data and improving prediction accuracy. Wu Yaqin [31] and others combined genetic algorithms with simulated annealing algorithms to propose a genetic simulated annealing algorithm. They introduced adaptive learning rates into the BP neural network and further optimized the BP network using the GASA algorithm. They ultimately established an improved GASA—BP neural network model for outburst prediction. The accuracy of the predicted results of this model was verified through practical application in coal mines.Xuning Liu [32] proposed a hybrid prediction model that combines feature extraction and pattern classification for coal and gas outbursts. Experimental results on a coal and gas outburst dataset showed that, compared to other models from the current coal and gas outburst prediction models, this method significantly influenced various indicators. In the field of machine learning, the improvement in model accuracy through dataset optimization often surpasses the improvement achieved through algorithm enhancements [33]. However, in practical production activities, the occurrence

of coal and gas outburst accidents may result in limited and missing accident data due to the damage of monitoring devices. This leads to reduced model accuracy, overfitting issues, and other problems.

In view of this, this paper analyzes the correlation between different indicators and the risk level of coal and gas outbursts using the grey relational algorithm, aiming to select feature indicators and perform attribute reduction based on the mechanism of coal and gas outbursts. A data generation algorithm based on machine learning and data reconstruction is constructed to generate virtual data from the original data. XGBoost, SVM, and GBDT are selected as primary learners, and random forest is used as a secondary learner to construct a predictive model for the risk level of coal and gas outbursts based on stacking ensemble learning. This model predicts the magnitude of the risk level of coal and gas outbursts, and the results are compared. This work aims to reduce personnel casualties and related economic losses caused by coal mine accidents from coal and gas outbursts, and promote the construction of a smart mine safety system.

This paper introduces the basic theoretical part of the gas outburst risk warning model, including the process and principles of grey relational analysis and the XGBoost, SVM, GBDT, and random forest algorithms. Then, a gas outburst risk warning model is constructed based on stacking ensemble learning. The construction process for this model consists of three parts. The first step is to build a data generation model based on XGBoost-Regressor to expand the capacity of the original dataset. The second step is the selection of main control factors based on grey relational analysis. The factors with the highest grey correlation coefficients, indicating the greatest relevance to coal and gas outburst risks, are selected from a set of 20 indicators. The third step is to construct a warning model based on the stacking fusion algorithm framework. Through the implementation of XGBoost, SVM, GBDT, and random forest as the four learners, an effective warning for coal and gas outburst risks is achieved. Furthermore, sufficient experimental analysis is conducted, and the model's performance is analyzed to select the factors that most influence gas outburst risks. Finally, the experimental results are summarized and analyzed.

2. Basic Theory

2.1. Grey Relationship Analysis

Grey system theory can utilize limited small sample data to solve uncertainty problems. Grey relational analysis, a prominent technique in the theory, is used to measure the degree of correlation between influencing factors and the research issue. It finds wide application in various domains for system diagnosis and analysis. The main principle is to assess the correlation between factors based on the similarity or dissimilarity of their development trends. By conducting grey relational analysis, one can analyze the impact of sub-factors on the main factor, aiming to optimize the dimensions of the system.

Grey relational analysis typically begins with selecting a reference sequence. Due to the varying dimensions among different influencing factors, direct comparison is not feasible. Therefore, commonly used approaches such as mean normalization, initial value normalization, standardization, or extreme value normalization are employed to eliminate the dimensional differences. These techniques help convert the indicators into dimensionless values, facilitating further analysis and comparison.

For grey relational analysis on processed data, the grey relational coefficient is calculated using Formula (1).

$$y(x_0(k), x_i(k)) = \frac{a + \rho b}{|x_0(k), x_i(k)| + \rho b} \tag{1}$$

The a is the minimum value, b is the maximum value of the two levels, and ρ is the discrimination coefficient (generally 0.5) [34]. $y(x_0(k), x_i(k))$ is the grey correlation coefficient, representing the correlation degree value between the sequence and the reference sequence.

The grey correlation degree of each subsequence is calculated by Formula (2), and the grey correlation degree is set as r_i:

$$r_i = \frac{1}{n} y(x_0(k), x_i(k)) \tag{2}$$

According to the obtained grey correlation degree r_i, it is sorted according to the size of r_i. If $r_1 > r_2$, it means that r_1 and the mother sequence are more related to r_0 and more correlated. According to the sorting results, the factor with the largest grey correlation degree value is selected as the main control factor. As shown in Figure 1.
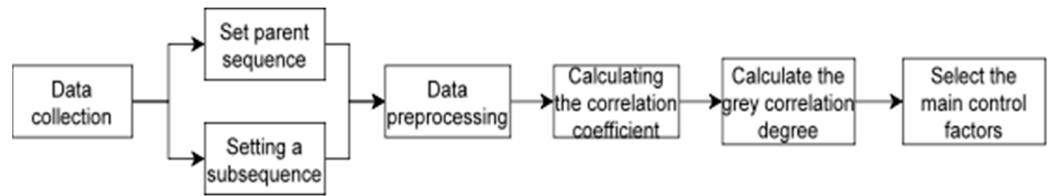

Figure 1. Grey correlation analysis flowchart.

2.2. Data Generation Algorithm

In machine learning, gas outburst is a prominent issue related to small sample problems. One important approach to improve model accuracy, prediction, and risk identification is through data augmentation and reconstruction. This paper constructs a gas outburst data generation model based on collected data to generate synthetic samples. This method can generate data samples that are consistent with the real data distribution, enhancing data and improving the effectiveness and quality of the model. The specific process of the data generation strategy consists of three steps: (1) training feature models, (2) sampling features to generate virtual data, and (3) generating the final synthesized data.

In the data generation algorithm, first, in the known data samples with certain features, a feature in the data sample is designated as the label, while the remaining features are treated as elements of the new feature vector to form new training samples. Based on the new training samples, data are randomly sampled and recombined by analyzing the correlations between various features, thus obtaining feature values for n features. Finally, by selecting individual features, using Feature 1 as input to the model for training and obtaining output F1, and using Feature 2 as input to the model for training and obtaining output F2, after completing the sampling process for all indicators, a temporary data sample can be obtained, thereby obtaining a complete dataset. Its data distribution characteristics are similar to the original data and have good representativeness. As shown in Figure 2.
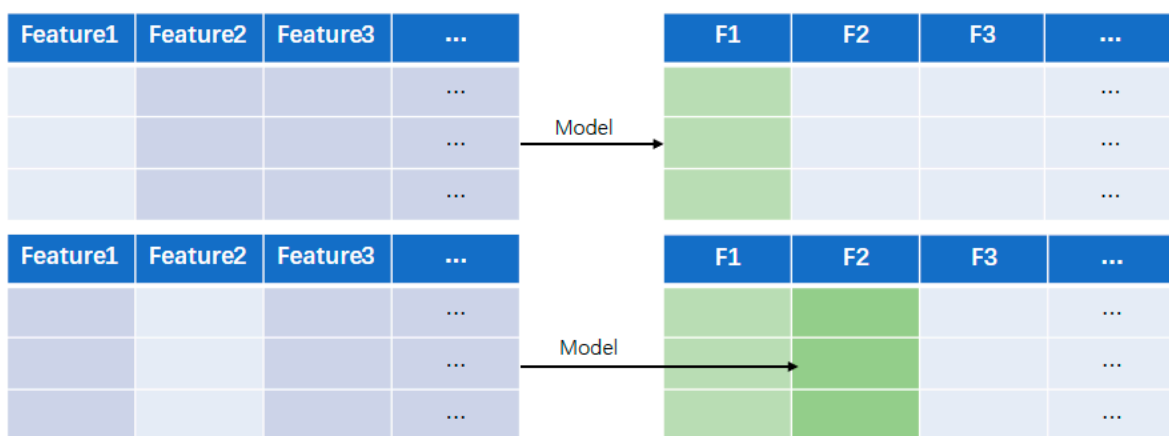

Figure 2. Virtual data generation process.

2.3. XGBoost

XGBoost (eXtreme Gradient Boosting) is a powerful supervised multi-parameter model that operates within the gradient boosting framework. It is an implementation of

the boosting algorithm, designed to handle both classification and regression problems. The fundamental concept behind XGBoost involves combining multiple weak learners to form a robust learner using specific techniques. This method utilizes multiple classification and regression trees in a collaborative manner, resulting in improved model performance. The following steps outline the working process of XGBoost:

Step 1: Calculate the predicted results of the model on the samples after t iterations, and then define a function that incorporates the model's loss function and a regularization term to suppress the complexity of the model.

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i) \tag{3}$$

Here, x_i represents the *i*-th feature in the feature vector, $\hat{y}_i^{(t)}$ represents the predicted value of sample *i* after *t* iterations, *k* is the number of base models, $f_k(x_i)$ represents the *k*th base model, and $\hat{y}_i^{(t-1)}$ represents the predicted value of sample after $t-1$ iterations. $f_k(x_i)$ is the model of the *t* th tree.

Step 2: Calculate the loss function of the model and the objective function consisting of the regularization term that inhibits the complexity of the model:

$$Obj^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) + \sum_{t=1}^{T-1} \Omega(f_t) \tag{4}$$

Here, y_i represents true value, $\sum_{i=1}^{n} l(y_i, \hat{y}_i)$ represents the loss function L, and Ω represents a regular term that suppresses the complexity of the model.

Step 3: Simplify the calculation to obtain the final objective function and solve the model. The objective function is close to the Taylor expression, so the objective function can eventually be simplified as follows.

$$Obj^{(t)} \simeq \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{5}$$

Here, g_i, h_i are the first derivative and the second derivative of the loss function L, $g_i = \partial_{\hat{y}^{(t-1)}} l\left(y_i, \hat{y}^{(t-1)}\right)$, $h_i = \partial^2_{\hat{y}^{(t-1)}} l\left(y_i, \hat{y}^{(t-1)}\right)$. Therefore, as long as the loss function is determined, then g_i and h_i are determined, and the objective function is determined.

2.4. SVM

Support vector machine (SVM) is a powerful generalized linear classifier primarily employed for binary classification in supervised learning. The underlying principle involves mapping the original data into a high-dimensional feature space through a series of transformations, enabling efficient classification within this transformed space. SVM exhibits strong generalization and self-learning capabilities, ensuring effective performance even with limited statistical sample datasets. The following steps outline the working process of SVM:

Step 1: Given the training set $T = \{(x1, y1), (x2, y2), \ldots, (xn, yn)\}$.

Step 2: Solve the quadratic programming problem, which is solved by $\alpha* = (\alpha_1^*, \ldots, \alpha_n^*)^T$.

$$\min \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j y_i y_j (x_i . x_j) - \sum_{i=1}^{n} a_i \tag{6}$$

$$s.t. \sum_i a_i y_i = 0, a_i \geq 0 \tag{7}$$

Here, x_i represents the *i*-th feature in the feature vector, and y_i represents true value.

Step 3: Calculate the parameter w and take a positive component α_i^*, and calculate the b.

$$w^* = \sum \alpha_i^* y_i x_i \tag{8}$$

$$b^* = y_j - \sum a_i^* y_i (x_i \cdot x_j) \tag{9}$$

Step 4: Structural decision boundary: $g(x) = (w^* \cdot x) + b^* = 0$, the decision function is thus obtained:

$$f(x) = \text{sgn}(g(x)) \tag{10}$$

2.5. GBDT

Gradient boosted decision trees (GBDT) is an iterative algorithm that combines the concepts of boosting and gradient descent. It leverages the forward distribution algorithm to train multiple weak learners, with each weak learner being constructed using a CART regression tree. By combining these weak learners through an additive model, it forms a powerful strong learner. The training process of each weak classifier is guided by the negative gradient of the loss from the previous weak classifier. This iterative optimization approach gradually reduces the loss, leading to the convergence towards the optimal solution. The principles of GBDT can be summarized as follows:

The model constants are initially given:

$$F_0(\vec{x}) = \underset{\gamma}{\text{argmin}} \sum_{i=1}^{n} L(y_i, \gamma) \tag{11}$$

Here, y_i represents true value, and γ represents the prior probability of a class. $\sum_{i=1}^{n} L(y_i, \gamma)$ represents the loss function.

For $m = 1$ to M, compute the pseudo-residual:

$$r_{im} = -\left[\frac{\partial L(y_i, F_{m-1}(\vec{x_i}))}{\partial F_{m-1}(\vec{x_i})} \right] (i = 1, 2, \ldots, n) \tag{12}$$

Using data to calculate the basis function for fitting residuals, in gradient enhancement, a decision tree is calculated based on pseudo residuals. The input space is divided into disjoint regions, and the decision tree can provide a certain type of deterministic prediction in each region.

Update the current model to:

$$F_m(x) = F_{m-1}(x) + \gamma t_m(x) \tag{13}$$

$$F_m(x) = F_{m-1}(x) + \sum_{j=1}^{J} \gamma_{jm} I(x \in R_{jm}) \tag{14}$$

$$\gamma_{jm} = \underset{w}{\text{argmin}} \sum_{\vec{x_i} \in R_{jm}}^{n} L(y_i, F_{m-1}(x_i) + w) \tag{15}$$

Here, x_i represents the i-th feature in the feature vector, and $t_m(x)$ is the basis function that fits the residual. $I(x)$ represents indication mark.

The final model as:

$$F_M(x) = F_0(x) + \sum_{m=1}^{M} \sum_{j=1}^{J} \gamma_{jm} I(x \in R_{jm}) \tag{16}$$

2.6. Random Forest

Random forest is an ensemble learning algorithm in the field of machine learning. It consists of multiple decision trees as classifiers. Each decision tree independently outputs

a class, and the final prediction is determined by taking the majority class among these decision tree outputs. Random forest integrates the predictions of multiple trees using the idea of ensemble learning. By combining the effects of random forest, we can obtain more robust and accurate classification results. Random forest is widely used in practical applications, especially for handling large-scale datasets and high-dimensional features. It can not only effectively handle classification problems but also be used for regression and feature selection tasks.

The root of the decision tree algorithm is information. The basic concepts of entropy and information gain can be understood through these three concepts to determine the order of feature selection in the decision tree.

$$H(x) = \sum_{i=1}^{n} p(x_i) I(x_i) = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i) \tag{17}$$

Here, $H(x)$ indicates entropy, which depends on x. The distribution, but with x It doesn't matter. In the decision tree. In the middle, the greater the entropy. The greater the category uncertainty, the smaller the reverse. Information gain is used to select feature indicators in the decision tree. The greater the information gain, the better the selectivity of features. A feature corresponds to multiple categories, and conditional entropy needs to be introduced in the calculation. The formula is as follows:

$$H(Y|X) = \sum p(x) H(Y|X = x) \tag{18}$$

In fact, the information gain also expresses the difference between the entropy of the set to be classified and the conditional entropy of selecting a feature. Therefore, the information gain formula is introduced:

$$IG(Y|X) = H(Y) - H(Y|X) \tag{19}$$

The prediction effect of ensemble learning is judged by the error rate. According to the Hoeffding inequality, the error rate of ensemble learning is:(where T is the number of decision trees):

$$P(H(X) \neq f(x)) = \sum_{k=0}^{T/2} \binom{T}{k} (1 - \varepsilon)^k \varepsilon^{T-k} \leq \exp\left(-\frac{1}{2} T (1 - 2\varepsilon)^2\right) \tag{20}$$

The algorithm steps are as follows:

Step 1: Using the Bagging method, randomly select N samples from N sample sets, and use these N samples to train a decision tree as samples for the root node of the decision tree.

Step 2: From the sample M Random selection among the features m Characteristics, satisfied $m << M$, Select a feature from m features using a certain strategy (information gain, Gini index, etc.) as the splitting feature of the node.

Step 3: Repeat step 2 to split the node until it cannot be split, forming a decision tree.

Step 4: Follow steps 1~2. Establish a large number of decision trees to form random forests.

3. An Account of Stacking-Framed Coal and Gas Outburst Risk Warning Model

All data undergo pre-training data cleaning, data filtering, and data sorting. The influence of dimensionality differences among different indicators is removed through standardization processing. Missing values are filled in using data cleansing techniques, and certain data points are identified or removed. Once data processing is complete, model training can commence.

Step 1: Training feature model. In the known presence ofIn the data sample of features, each dataset has its own characteristics. One feature of the given data sample is used as a

label, and the rest of the feature are used as elements of the new feature vector to form a new training sample.

Step 2: Sampling features to generate virtual data. From the collection, replace a value of the sample: the The value is the characteristic of the temporarily synthesized data sample. 1. The first eigenvalue, and then the remaining features are taken in turn, and finally you can getthe eigenvalues of the feature.

Step 3: The features to be generated in the temporary data sample 1. As an output, characteristics 2 and other features are used as input respectively... Put in the feature matrix and put The output value of the model is characteristic.1The virtual generation value of In the same sense, all virtual values are obtained.

Step 4: Based on the original data and the newly generated data, calculate the grey correlation coefficient between the degree of gas outburst and various factors.

$$y(x_0(k), x_i(k)) = \frac{a + \rho b}{|x_0(k), x_i(k)| + \rho b} \tag{21}$$

Step 5: Calculate the grey correlation between each subsequence, and sort each indicator according to the grey correlation.

$$r_i = \frac{1}{n} y(x_0(k), x_i(k)) \tag{22}$$

Step 6: Divide indicator groups based on the size of grey correlation degree, and use XGBoost, RF, GBDT, and SVM training data to comprehensively compare MAE, MSE, and R2 in each indicator. Select the data group with the smallest error and the best fit degree as the main control factor.

Step 7: Based on the main control factors of coal and gas outburst, SVM, XGBoost, GBDT, and RF training models are selected to predict the situation of gas outburst.

Step 8: Using the above predictions on the training set as the training set and the predictions on the test set as the test set, retrain the data to obtain new prediction results. Stacking combines steps 7 and 8 together, using the output of step 7 as the input to step 8 to obtain the final output result.

Step 9: Compare the training effects of SVM, XGBoost, RF, GBDT, and Stacking fusion models under the main control factors, and select the Stacking fusion model with the best prediction effect to construct a gas outburst risk warning model.

The occurrence of coal and gas outbursts is comparatively intricate. It is influenced by multiple factors and exhibits a certain degree of nonlinearity. This article employs grey relational analysis to explore the nonlinear relationship between coal and gas outburst issues. It aims to unveil the correlation between influencing factors and the occurrence of outbursts. A set of key controlling factors that impact gas outburst situations is selected. By doing so, the prediction work is streamlined, resulting in improved model forecasting performance.

Due to variations in gas outburst factors among different mines, there is a certain relationship between these factors and the mining environment and coal occurrence conditions. Therefore, when constructing a gas outburst risk warning model, the model's generalization ability should be considered. Integrated models have better generalization ability compared to single machine learning models, which can improve the accuracy of the model. Classic methods of integration include bagging, boosting, and stacking. The main idea of stacking ensemble learning is to combine multiple models and fully utilize their respective advantages to make final predictions and achieve the best results. As shown in Figure 3.
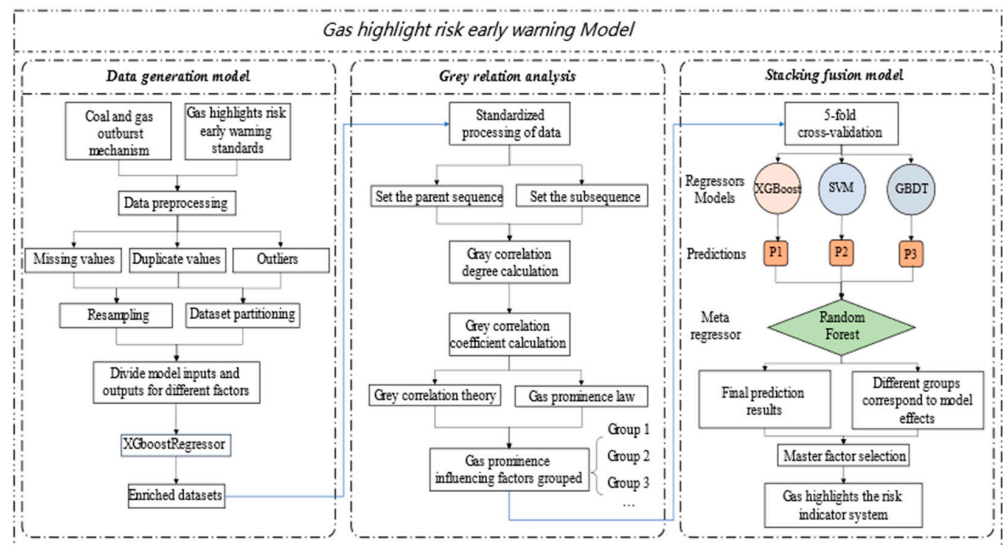
Figure 3. XGBoost–GR–stacking flowchart of gas outburst warning model.

4. Experimental Simulation

4.1. Data Source

The influential factors of gas outburst can include the static data of the mine and the dynamic data of real-time monitoring. The static parameters of the outburst mine partially reveal the potential patterns of gas outbursts. To study the laws of gas outbursts, this paper collected relevant data on gas outbursts from 26 mines in the Southwest China, North China, and Central China regions, as well as numerical simulation indicators. The related data indicators included mining depth, coal solidification coefficient, ash content, volatile matter, initial gas emission velocity, gas adsorption constants a and b, Δh_2, absolute gas outburst (m^3/min), relative gas outburst (m^3/t), K$_1$, Smax (kg/m), q (L/min), and ejection amount. At the same time, this paper used numerical simulation experiments to simulate the dynamic slope of gas pressure, geostress, goaf thickness, fault height, and fitting curve of pre-gas outburst data with similar ejection amounts. Part of the raw data is shown in Table 1.

Table 1. Part of the original data.

Mining Depth	Coefficient of Coal Firmness	Ash Content	Volatile Matter	Seam Thickness	Initial Velocity of Gas Release	Gas Adsorption Constant a	Gas Adsorption Constant b	Δh_2	Absolute Gas Emission	Relative Gas Emission
800	0.34	0.18	0.12	1.25	14.84	38.59	0.787	170	32.01	13.94
452.7	0.31	0.15	0.1	5.69	12.16	37.32	0.723	140	6.58	4.21
850	0.39	0.23	0.12	4	18.6	18.13	2.3445	180	34.8	15.09
600	0.49	0.12	0.08	1.7	28	30.303	1.3346	175	32.04	13.95
500	0.24	0.18	0.12	1.25	31	33.4832	1.6166	170	32.01	13.94
515	0.17	0.17	0.11	4.6	38	31.08	1.13	172	16.01	10.88
500	0.23	0.17	0.11	3.2	26.3	26.3459	1.2572	160	9.78	10.45
…	…	…	…	…	…	…	…	…	…	…

K$_1$	Smax	q	Initial gas pressure	Numerical simulation of gas pressure	Numerical simulation of ground stress	Numerical simulation of stone gate thickness	Numerical simulation of fault height	Dynamic slope
0.36	3.5	57	0.75	28	4	5	15	0.75
0.32	3.4	26.05	0.75	35	4	1	2	0.75
0.34	3.7	34	2.4	16	0.2	1	5	2.4
0.41	2.1	38	0.75	10	3	5	7.5	0.75
0.413	2.2	42	2.4	35	4	1	1.1	2.4
0.36	2.2	83	0.75	10	0.2	5	20	0.75
0.47	2	36	0.75	22	1	1	1.5	0.75
…	…	…	…	…	…	…	…	…

4.2. Data Generation

Part of the virtual data is shown in Table 2. The data samples contained in this table were part of the virtual samples generated based on the above methods, and their data distribution characteristics were similar to the original data, which had a good representation. This study tested the effect of the model based on the following data.

Table 2. Partial virtual data tables.

Mining Depth	Coefficient of Coal Firmness	Ash Content	Volatile Matter	Seam Thickness	Initial Velocity of Gas Release	Gas Adsorption Constant a	Gas Adsorption Constant b	Δh_2	Absolute Gas Emission	Relative Gas Emission
681.0	0.474	0.124	0.299	1.277	5.118	27.82	0.949	152.6	30.85	681.0
400.0	1.499	0.310	0.210	1.320	16.39	38.58	0.790	150.0	2.490	400.0
400.0	1.500	0.310	0.210	1.320	16.39	38.59	0.790	150.0	2.490	400.0
399.9	1.499	0.310	0.210	1.320	16.39	38.59	0.790	150.0	2.490	399.9
400.0	1.499	0.310	0.209	1.321	16.39	38.58	0.790	150.0	2.490	400.0
450.0	0.490	0.131	0.100	1.550	18.35	33.09	1.120	179.9	27.58	450.0
450.0	0.490	0.130	0.100	1.550	18.35	33.09	1.120	180.0	27.59	450.0
...	...	...	...	...	...	...	...	...	...	...

K_1	Smax	q	Initial gas pressure	Numerical simulation of gas pressure	Numerical simulation of ground stress	Numerical simulation of stone gate thickness	Numerical simulation of fault height	Dynamic slope
0.403	2.949	42.681	0.962	1.766	21.832	2.222	2.868	3.967
0.389	2.800	43.000	1.500	1.300	16.000	0.200	1.001	5.000
0.390	2.800	43.000	1.500	1.300	16.000	0.200	1.000	5.000
0.389	2.800	43.000	1.499	1.300	16.000	0.200	1.000	5.000
0.390	2.800	43.000	1.500	1.300	16.000	0.200	1.001	5.000
0.340	2.900	38.000	0.650	2.399	16.000	0.200	1.000	4.500
0.340	2.900	38.000	0.650	2.400	16.000	0.200	1.000	4.500
...	...	...	...	...	...	...	...	...

The diagram directly reflects the distribution of the original data and extended data. It also standardizes the original data and extended data of the 20 index groups to remove the dimensional influence. The experimental results are shown in Figure 4.
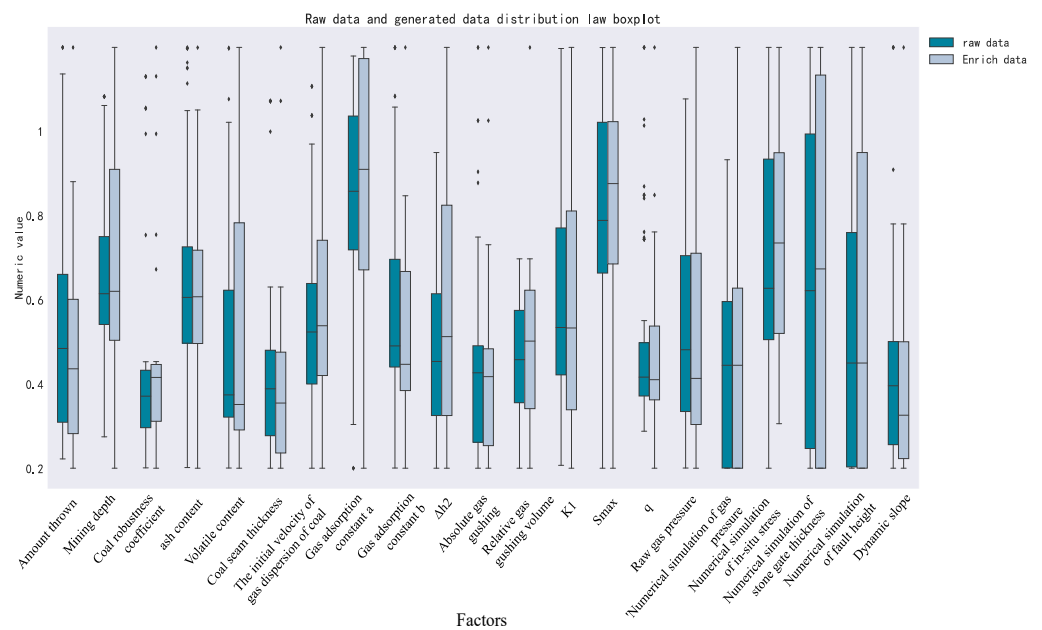

Figure 4. Box diagram of distribution of original data sample and expanded data.

By comparing the distribution of the original data and the extended data, it was found that the distribution interval of the 20 indicators of the extended data was basically consistent with the distribution interval of the original data, indicating that the extended data was consistent with the distribution of gas outburst indicators in the mine, had strong credibility, and could be used for the construction of the gas outburst warning model.

In order to verify whether the generated virtual data could effectively improve the training effect of the model in this paper, the original data samples and the expanded data samples were respectively placed in the stacking fusion model for training. Then, we analyzed the ROC curve of the stacking fusion model to prove the feasibility of the above methods. The experimental results are shown in Figure 5.
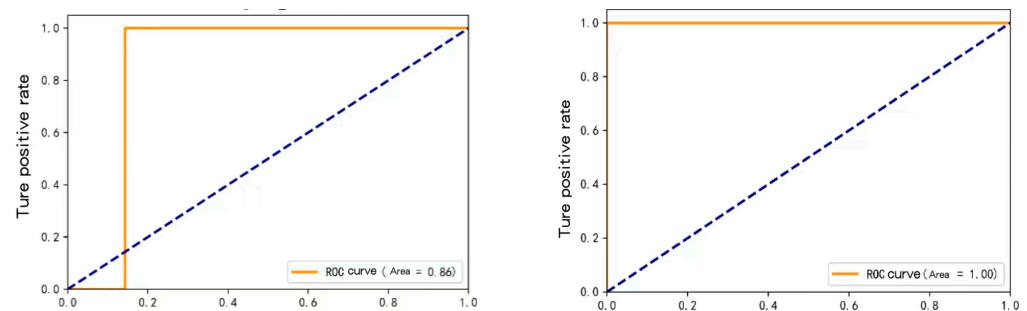
Figure 5. Comparison of the effects of the stacking model between original data samples and expanded data samples.

For comparing the ROC curve of the original data and the extended data, it is generally believed that the closer the area under the ROC curve is to 1, the better the classification prediction of this model is. When the value is greater than 0.5, the model is better than random guessing. If the model sets a reasonable threshold, the classifier can have predictive value. For the expanded data, the ROC curve area value of the stacking model was equal to 1, so the prediction results were perfect. This model belongs to a good predictor. For the raw data, where the ROC curve area of the stacking fusion model was equal to 0.86, the prediction effect was acceptable. The ROC curve shows that the expanded data model was better than that for the original data.

In addition, this paper trained XGBoost, GBDT, SVM, RF, and stacking prediction models based on the original and extended data. After visual comparison of the difference between the real value and the predicted value (Figures 6 and 7), it was found that the error between the real value and the predicted value of the expanded data was significantly smaller than that between the real value and the predicted value of the original data. The error of SVM-predicted value and true value was the largest for raw data and extended data, but the prediction error was larger for the raw data. The errors between the predicted values and the true values of the stacking model were the minimum for the original data and the expanded data, and the prediction effect was better for the expanded data. The actual values and errors of the five models showed that the extended data model was better than that for the original data.

All in all, the ROC curve comparison between the original data and the extended data and the errors between the real and predicted values of the XGBoost, GBDT, RF, SVM, and stacking models trained with the original data and the extended data, respectively, indicated that the model effect after data expansion was significantly improved compared with the model effect with the original data. This paper used extended data to train the model.
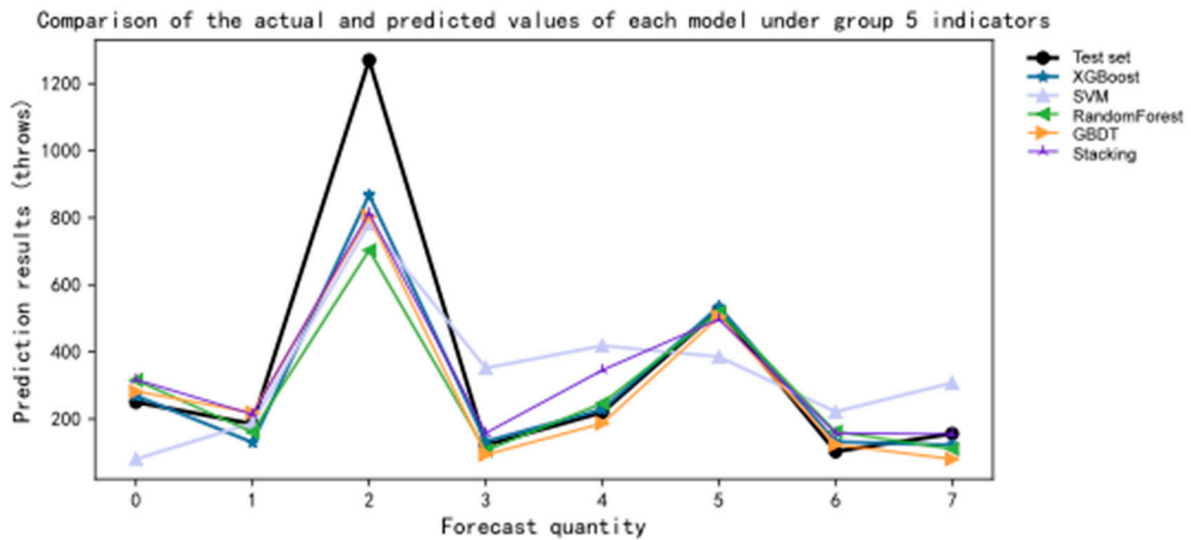
Figure 6. Comparison between the predicted value and the true value of each model for the original data.
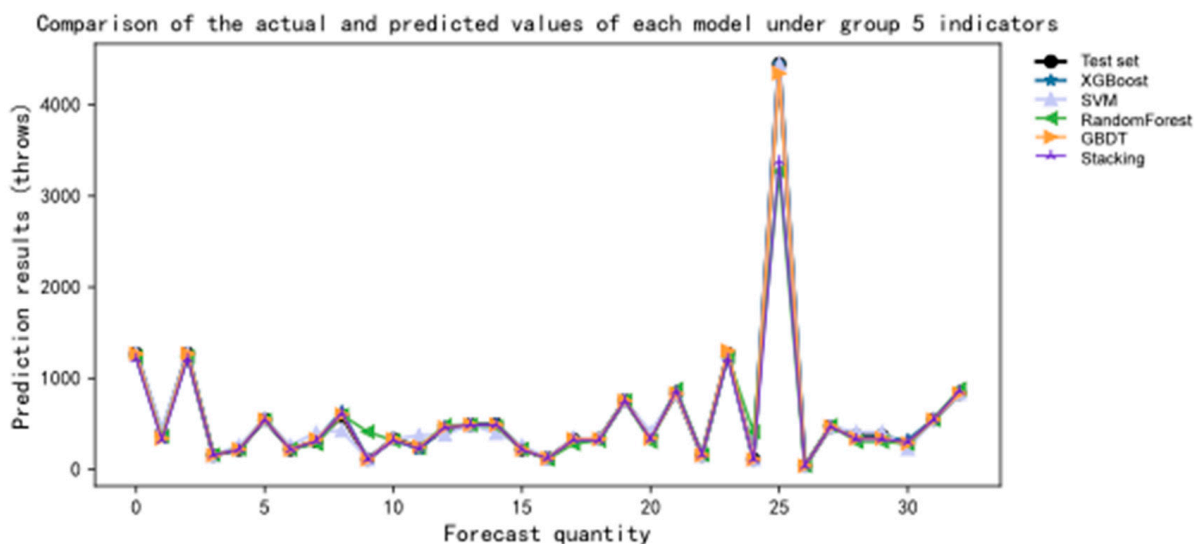

Figure 7. Comparison between the predicted value and the true value of each model for the extended data.

4.3. Analysis of Correlation Degree of Risk Factors of Gas Outburst

This paper explored the concept of throw-out quantity as the main factor, with other indicators as subsidiary factors. The dimensional unit of each data group was eliminated through standardization processing. By calculating the grey correlation coefficient using Formula (1), combined with Formula (2), the magnitude of the grey correlation degree between the grey subsequence and the main sequence was obtained. The results were then sorted based on the grey correlation degree, as shown in Table 3:

Table 3. Grey correlation degree ranking of subsequence and parent sequence.

Name	Grey Correlation Degree Value	Rank
Dynamic slope	0.830	1
Initial velocity of coal gas release	0.745	2
Initial velocity of coal gas release	0.742	3
Δh_2	0.728	4

Table 3. *Cont.*

Name	Grey Correlation Degree Value	Rank
q	0.727	5
Relative abundance of methane	0.709	6
Numerical simulation of fault height	0.699	7
Absolute gas emission rate	0.691	8
Coal firmness coefficient	0.678	9
Ash content	0.670	10
Thickness of coal seam	0.668	11
Exploitation depth	0.667	12
Numerical simulation of gas pressure	0.666	13
K_1	0.649	14
Initial gas pressure	0.646	15
Volatiles	0.638	16
Numerical simulation of ground stress	0.635	17
Gas adsorption constant a	0.593	18
Numerical simulation of stone gate thickness	0.578	19
Smax	0.572	20

The results in the table indicate a clear distinction in the grey correlation of all 20 indicators, ranging from 0.572 to 0.830. Among the 20 indicators, the grey correlation value between the dynamic slope and gas emission is the highest at 0.830. This suggests that there is a strong and close relationship between the dynamic slope and gas emission. Following that, the grey correlation values decrease significantly for the initial gas emission velocity of coal and the gas adsorption constant b, reaching 0.745 and 0.742, respectively, indicating a weakening correlation. Among the twenty indicators, the grey correlation between Smax and gas emission is the lowest, at 0.572, implying the weakest and least significant relationship between Smax and grey correlation. The value of 0.578 for the simulated thickness of the gate in numerical simulations indicates a loose relationship between the simulated gate thickness and gas emission.

According to the results of grey correlation degree calculated in Table 2, we set five conditions respectively for prediction, and then selected the main influencing factors according to the error of the predicted results. We set up five groups of experiments to study the correlation and influencing factors. The correlation degrees were, respectively, $r_i > 0.700$, $r_i > 0.670$, $r_i > 0.650$, and $r_i > 0.630$. The results are shown in Table 4:

Table 4. Classification of influencing factors based on correlation degree.

Group	Grey Relational Degree	Influence Factors
1	$r_i > 0.700$	Dynamic slope, initial velocity of coal gas release, gas adsorption constant, q, relative abundance of methane.
2	$r_i > 0.670$	Dynamic slope, the initial speed of gas dispersion of coal, gas adsorption constant, q, relative gas outflow, numerical simulation fault height, absolute gas outflow, coal toughness coefficient, ash content.
3	$r_i > 0.650$	Dynamic slope, initial velocity of coal gas release, gas adsorption constant, q, relative gas emission, numerical simulated fault height, absolute gas emission, coal firmness coefficient, ash content, coal seam thickness, mining depth, numerical simulated gas pressure.
4	$r_i > 0.630$	Dynamic slope, initial velocity of coal gas release, gas adsorption constant, q, relative gas emission amount, numerical simulated fault height, absolute gas emission amount, coal firmness coefficient, ash content, seam thickness, mining depth, numerical simulated gas pressure, K1, original gas pressure, volatile content, numerical simulated ground stress.
5	$r_i > 0.570$	All factors.

The impact of various factors on the level of gas outburst hazard differed based on the grouping results shown in the above table. Furthermore, there existed either strong or weak relationships among these factors. By conducting error analysis and analyzing the

fitting degree of the models for each indicator group, the key controlling factor indicator group was determined. The comparative effects of each model are shown in Table 5.

Table 5. Comparison of model results.

Model	XGBoost			SVM			RF		
Index	**MAE**	**MSE**	$\mathbf{R^2}$	**MAE**	**MSE**	$\mathbf{R^2}$	**MAE**	**MSE**	$\mathbf{R^2}$
Group 1	0.091	0.091	0.933	0.152	0.152	0.888	0.061	0.061	0.955
Group 2	0.031	0.031	0.981	0.094	0.094	0.944	0.031	0.031	0.978
Group 3	0.091	0.091	0.933	0.242	0.303	0.776	0.091	0.091	0.933
Group 4	0.152	0.152	0.900	0.424	1.152	0.244	0.091	0.091	0.940
Group 5	0.121	0.121	0.920	0.333	0.333	0.781	0.091	0.091	0.940

Model	GBDT			Stacking		
Index	MAE	MSE	R^2	MAE	MSE	R^2
Group 1	0.061	0.061	0.955	0.061	0.061	0.955
Group 2	0.031	0.031	0.981	0.031	0.031	0.981
Group 3	0.091	0.091	0.933	0.091	0.091	0.932
Group 4	0.061	0.061	0.960	0.121	0.121	0.920
Group 5	0.091	0.091	0.940	0.152	0.152	0.900

For the five models (XGBoost, GBDT, RF, SVM, and stacking), the table above presents metrics such as MAE, MSE, and R^2, which reflect the performance of the models. MAE and MSE represent the errors of the models, and a lower value indicates higher accuracy in the predictions. The R^2 coefficient represents the goodness of fit of the models. Within a certain range, a higher value indicates a better fit and a more effective prediction. Based on the data in the table:

For the first set of indicators, the XGBoost predictive model yields an MAE, MSE, and R^2 of 0.091, 0.091, and 0.933, respectively. The SVM predictive model scores an MAE, MSE, and R^2 of 0.152, 0.152, and 0.888, respectively. The RF predictive model, on the other hand, performs with an MAE, MSE, and R^2 of 0.061, 0.061, and 0.955. As for the GBDT predictive model, it achieves an MAE, MSE, and R^2 of 0.061, 0.061, and 0.955. Lastly, the stacking predictive model exhibits an MAE, MSE, and R^2 of 0.061, 0.061, and 0.955.

For the second index group, the MAE, MSE and R^2 of the XGBoost prediction model are 0.031, 0.031 and 0.981, respectively, while the MAE, MSE and R^2 of the SVM prediction model are 0.094, 0.094 and 0.944. The MAE, MSE and R^2 of the RF prediction model are 0.031, 0.031 and 0.978, respectively, and the MAE, MSE and R^2 of the GBDT prediction model are 0.031, 0.031 and 0.981, respectively. The MAE, MSE, and R^2 of the stacking prediction model are 0.031, 0.031, and 0.981, respectively.

For the third group of indicators, the XGBoost predictive model exhibits an MAE, MSE, and R^2 of 0.091, 0.091, and 0.933, respectively. The SVM predictive model shows an MAE, MSE, and R^2 of 0.242, 0.303, and 0.776, respectively. For the RF predictive model, the MAE, MSE, and R^2 are 0.091, 0.091, and 0.933, respectively. The GBDT predictive model demonstrates an MAE, MSE, and R^2 of 0.091, 0.091, and 0.933, respectively. Lastly, the stacking predictive model yields an MAE, MSE, and R^2 of 0.091, 0.091, and 0.932, respectively.

For the fourth group of indicators, the XGBoost predictive model achieves an MAE, MSE, and R^2 of 0.152, 0.152, and 0.9, respectively. The SVM predictive model presents an MAE, MSE, and R^2 of 0.424, 1.152, and 0.244, respectively. For the RF predictive model, the MAE, MSE, and R^2 are 0.091, 0.091, and 0.94, respectively. The GBDT predictive model demonstrates an MAE, MSE, and R^2 of 0.061, 0.061, and 0.96, respectively. Lastly, the stacking predictive model yields an MAE, MSE, and R^2 of 0.121, 0.121, and 0.92, respectively.

For the fifth group of indicators, the XGBoost predictive model exhibits an MAE, MSE, and R^2 of 0.121, 0.121, and 0.92, respectively. The SVM predictive model shows an MAE, MSE, and R^2 of 0.333, 0.333, and 0.781, respectively. For the RF predictive model, the MAE, MSE, and R^2 are 0.091, 0.091, and 0.94, respectively. The GBDT predictive model

demonstrates an MAE, MSE, and R^2 of 0.091, 0.091, and 0.94, respectively. Lastly, the stacking predictive model yields an MAE, MSE, and R^2 of 0.152, 0.152, and 0.9, respectively. As shown in Figure 8.
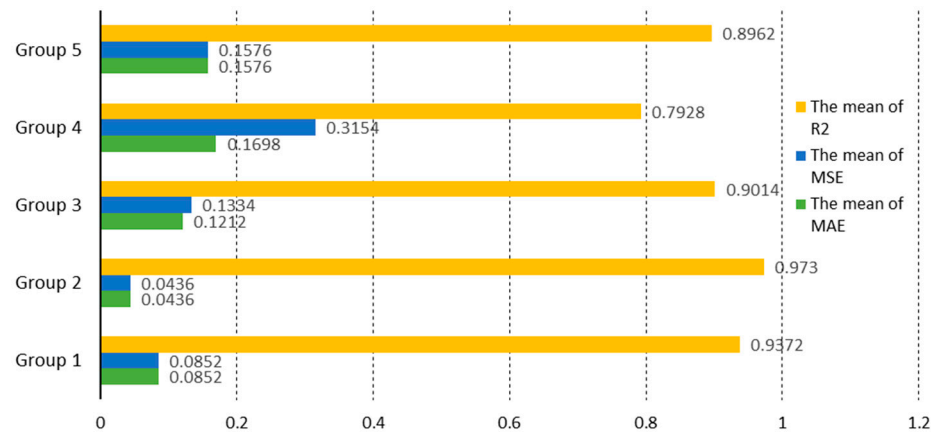

Figure 8. Average indicators of each group.

By plotting bar charts for the average MAE, MSE, and R^2 of the five datasets, it is visually evident that the second group of metrics performs optimally. Specifically, the MAE, MSE, and R^2 for this group are 0.0436, 0.0436, and 0.973, respectively. On the other hand, the fifth group exhibits the poorest performance with MAE, MSE, and R^2 values of 0.1576, 0.1576, and 0.8962, respectively. As shown in Figure 9.
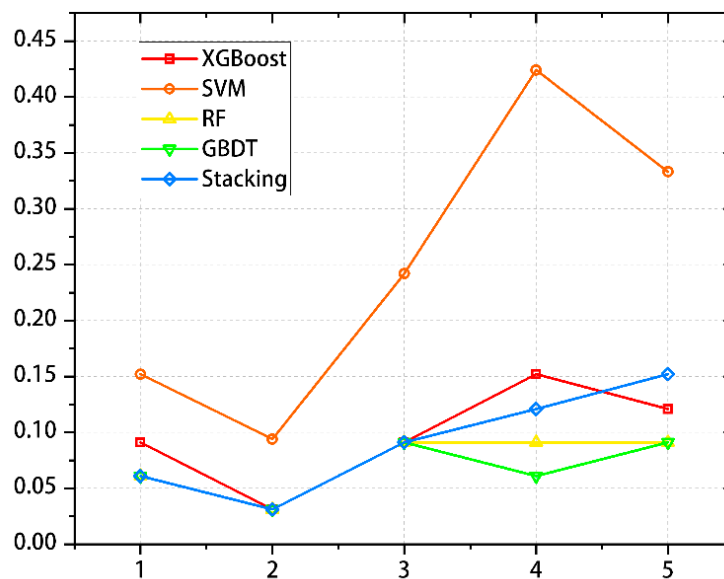

Figure 9. MAE index diagram for each mode.

By drawing five models, it was concluded that the change in MAE for the different data groups directly reflected the prediction effects of the different indicator groups in each model. By comparing the five points on the same line, it was found that the MAE value of the second indicator group was the smallest, while the MAE value of the fourth indicator group was the largest in SVM and XGBoost. The fifth indicator group had the largest MAE value in stacking, RF, and GBDT. As shown in Figure 10.

By drawing five models, it was concluded that the change about MSE values for the different data groups directly reflected the prediction effects of the different indicator groups in each model. By comparing the five points on the same line, it was found that the MSE value of the second indicator group was the smallest, while the MSE value of the

fourth indicator group was the largest in SVM and XGBoost. The MSE value of the fifth indicator group in stacking, RF, and GBDT was the largest. As shown in Figure 11.
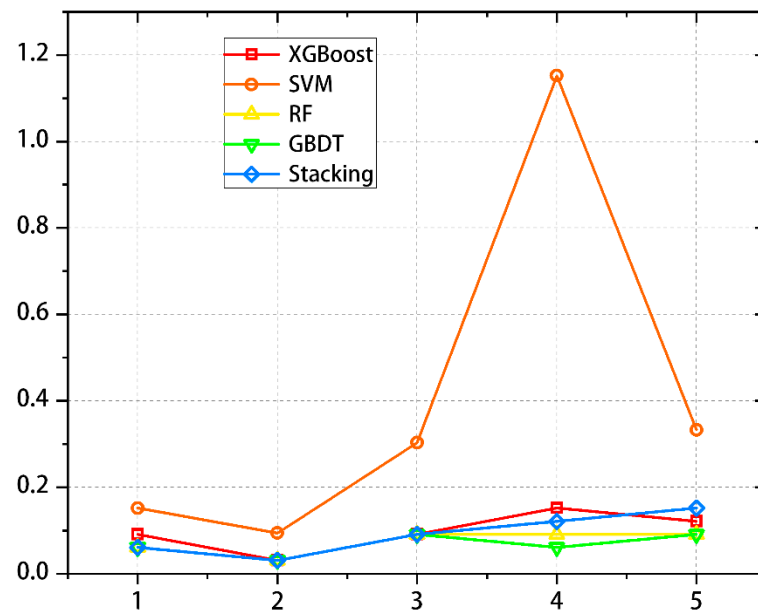
Figure 10. MSE index diagram for each model.
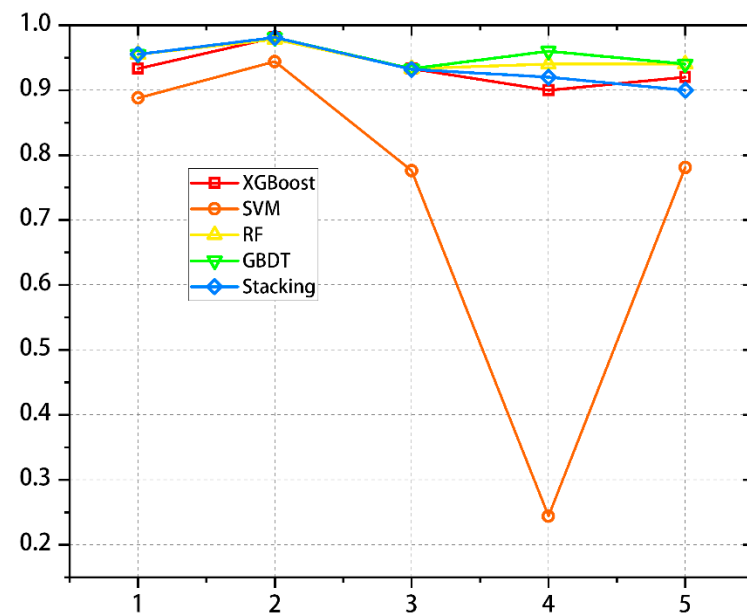
Figure 11. R^2 indicators for each model.

By drawing five models, it was concluded that the change in R^2 values for the different data groups directly reflected the prediction effects of the different indicator groups in each model. By comparing the five points on the same line, it was found that the R^2 value of the second indicator group was the largest, while the R^2 value of the fourth indicator group was the largest in SVM and XGBoost. The R^2 value of the fifth indicator group in stacking, RF, and GBDT was the largest.

After a comprehensive analysis of the aforementioned results, it was observed that the average MSE and MRE of the five models, with the second set of indicators as the main controlling factor, were the lowest, while the average R^2 was the highest. This indicated that, when the second set of indicators was used as the main controlling factor, the models

had the lowest errors and the best fitting performance. Therefore, it was possible to select the datasets with a grey correlation degree greater than 0.670 as the main controlling factors, namely, "dynamic slope", "q", "absolute gas emission volume", "coal seam thickness", "coal hardness coefficient", "relative gas emission volume", "volatile matter", and "gas adsorption constant b". Furthermore, among the first three sets of indicators, the stacking fusion model showed the best performance. As for the fourth and fifth sets of indicators, GBDT and RF demonstrated relatively good performance, but the overall performance was inferior to that of the stacking fusion model of the first three sets of indicators.

4.4. Analysis of Gas Outburst Risk Early Warning Model

The present study conducted a series of experiments to investigate the training performance of different models when input with gas outburst factor data under various degrees of relevance. Specifically, the gas outburst data were categorized into several groups based on their grey relevance, including factors with relevance degrees $r_i > 0.700$, $r_i > 0.670$, $r_i > 0.630$, $r_i > 0.570$, and all factors combined. The GBDT, RF, SVM, XGBoost, and stacking ensemble models were employed for training. In order to visually compare the differences between the actual values and predicted values, line graphs were plotted as shown in Figures 12–19.
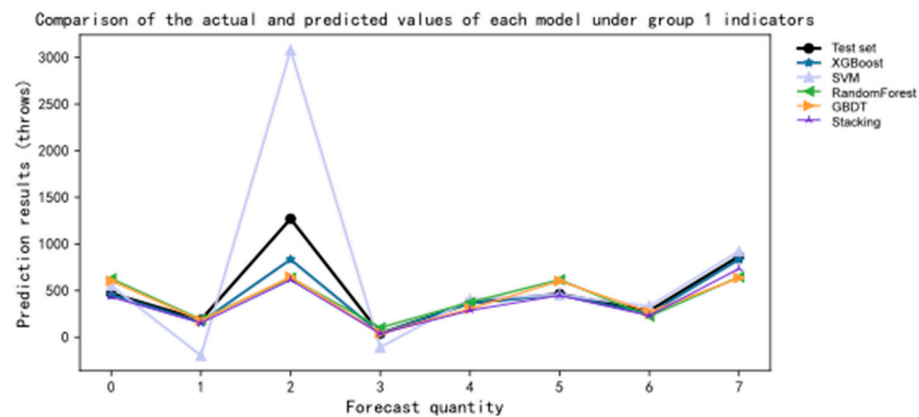
Figure 12. Comparison between the predicted value and the true value of each model under the first set of indicators in the original data.
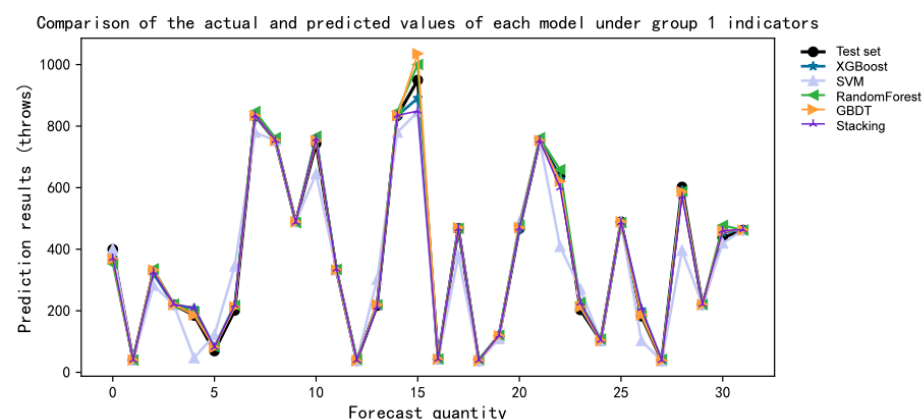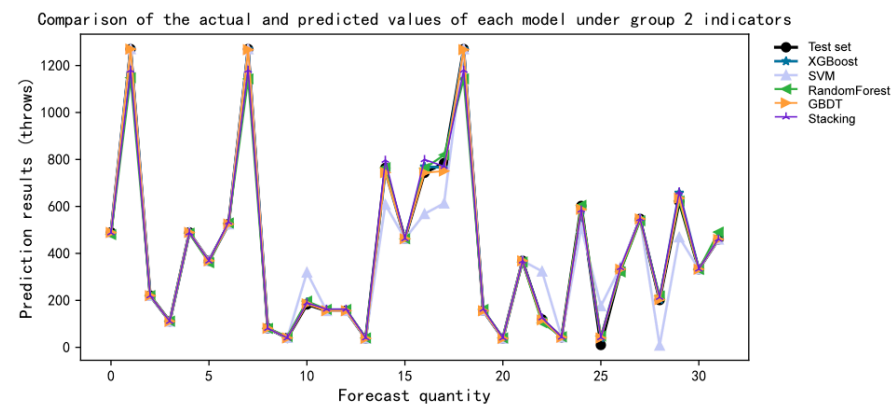

Figure 13. Comparison between the predicted value and the true value of each model of the original data under the first set of indicators in the extended data.

For the comparison of errors between actual values and predicted values, Figures 6 and 7 represent the complete set of indicators, which is the fifth group of indicators. The agreement between the actual values and predicted values for the expanded data in all five indicator groups was higher than that for the original data. This indicates that the predictive performance of the models was better for the expanded dataset compared to the original dataset.

Among the original and expanded datasets, the second group of indicators exhibited the highest agreement between the actual values and predicted values among the XGBoost, SVM, RF, GBDT, and stacking prediction models. This suggests that selecting the second group of indicators as the main control factor leads to the best predictive performance. Among the XGBoost, RF, GBDT, SVM, and stacking models, it is observed that, particularly for models trained on the second group of indicators from the expanded data, the stacking fusion model shows a closer fit between predicted values and actual values. The higher the accuracy of the model, the better the predictive performance.
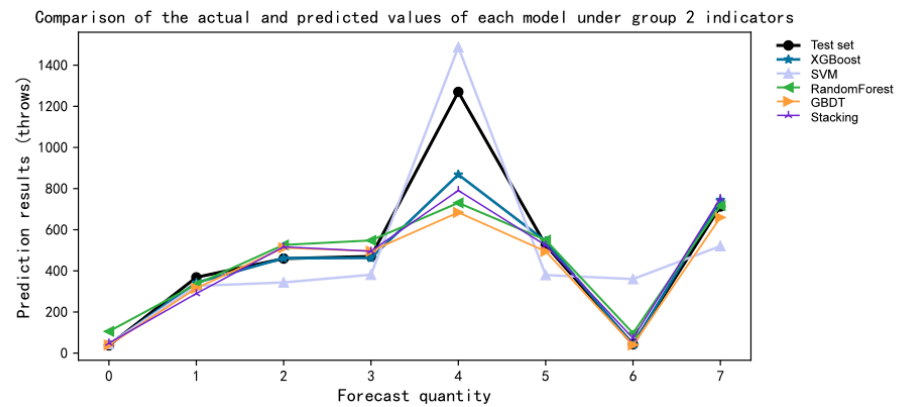
Figure 14. Comparison between the predicted value and the true value of each model of the original data under the second group of indicators in the original data.

Figure 15. Comparison between the predicted value and the true value of each model of the original data under the second group of indicators in the extended data.
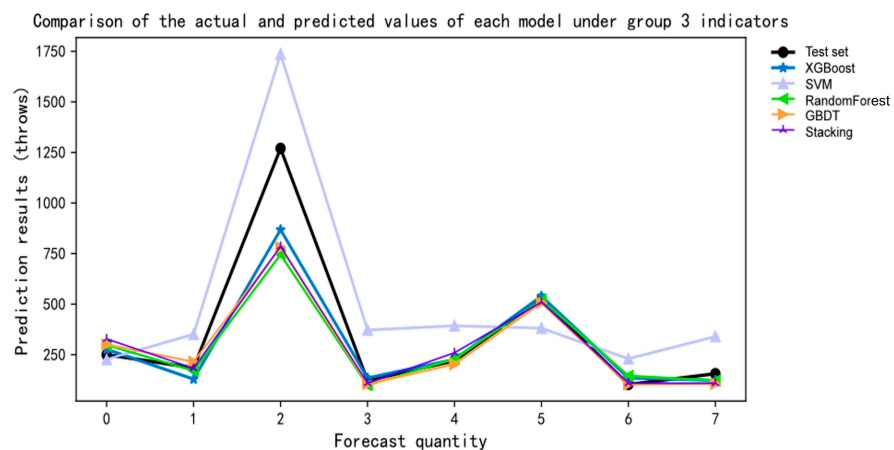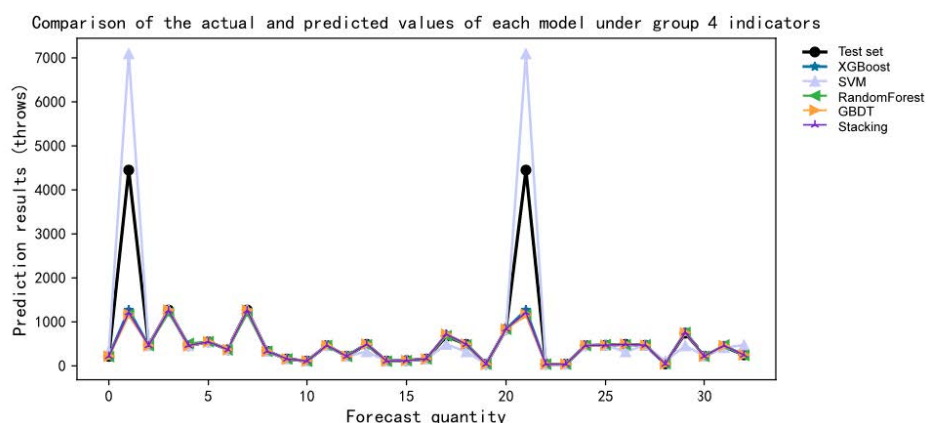

Figure 16. Comparison between the predicted value and the real value of each model of the original data under the third group of indicators in the original data.
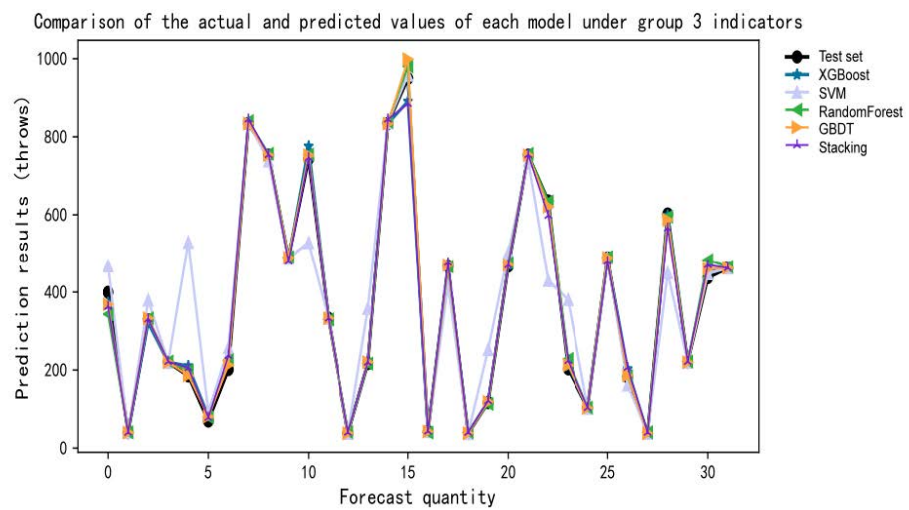
Figure 17. Comparison between the predicted value and the true value of each model of the original data under the third group of indicators in the extended data.
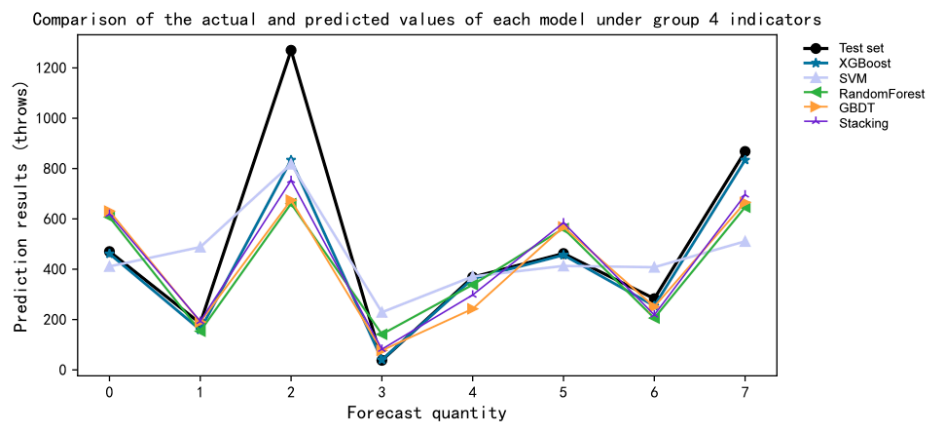
Figure 18. Comparison between the predicted value and the true value of each model of the original data under the fourth group of indicators in the original data.

Figure 19. Comparison between the predicted value and the true value of each model of the original data under the fourth group of indicators in the extended data.

In the course of the experiment, we utilized various evaluation metrics to comprehensively assess the performance and effectiveness of the model. These metrics included mean squared error (MSE), mean absolute error (MAE), and the coefficient of determination (R^2), among others.

Based on the comparison and analysis of the experimental results, stacking and RF exhibited the best predictive performance among the first set of metrics, while SVM performed the poorest. In terms of the second set of metrics, RF and stacking showed the most favorable outcomes with the selected controlling factors, while SVM performed the worst. However, overall, the predictive efficacy with the second set of metrics as controlling factors surpassed that of the first set. Regarding the third set of metrics, stacking demonstrated relatively good predictive performance. However, for models other than stacking, there were significant errors when comparing the actual values with the predicted values, making them unsuitable for forecasting purposes. As for the fourth and fifth sets of metrics, SVM again displayed the worst predictive performance, and stacking also exhibited inferior results compared to the previous sets. Additionally, by examining the data in Tables 4 and 5, we found that when selecting data with r > 0.700 as input for the model after augmentation, all models performed relatively well except for SVM. However, the comparison between actual values and predicted values in the analysis table indicated that stacking and RF achieved superior results with minimal error. Conversely, SVM performed the worst in terms of the comparison. Therefore, we concluded that stacking demonstrates greater stability and reliability compared to the RF model. Taking into consideration all the aforementioned factors, we chose stacking as the final gas outburst risk warning model.

5. Conclusions

1. This paper proposes a data generation model based on XGBoost to address the issue of coal and gas outburst risk warning. Virtual samples are generated to expand the dataset. By comparing the original data with the expanded data, it is found that the expanded samples outperform the original data in multiple methods. After data expansion, the model's predicted ROC curve area (AUC) value increased from 0.86 to 1.00, indicating a significant improvement in prediction effectiveness using the expanded data model.

2. The process of gas emission is influenced by various factors, and the relationship between each factor and the emission rate is nonlinear. This paper proposes the use of grey correlation analysis to select the main controlling factors based on the magnitude of the grey correlation degree. The experiments conducted indicate that the model with a grey correlation degree ranging from 0.67 to 0.70 achieves the best predictive performance, with average MSE, MRE, and R2 values of 0.0436, 0.0436, and 0.973, respectively. Therefore, the model with the least prediction error and the optimum model fit is identified. The group of factors includes "Dynamic Slope Indicator", "q", "Absolute Gas Emission Rate", "Coal Thickness", "Coal Firmness Coefficient", "Relative Gas Emission Rate", "Volatile Matter", and "Gas Adsorption Constant b".

3. This paper proposes the XGBoost–GR–stacking model for addressing the issue of coal and gas outburst risk warning. The XGBoost algorithm is utilized to generate data, while the GR algorithm is employed for feature selection. Furthermore, a prediction model based on the stacking fusion algorithm is established. The results show that the MSE, MRE, and R2 predictions of this model are 0.031, 0.031, and 0.981, respectively, which are superior to those of the XGBoost, GBDT, RF, and SVM models. This indicates that the proposed model exhibits lower prediction errors and a higher fitting degree, making it highly applicable in the domain of gas outburst warning.

References

1. Zhang, C.; Wang, P.; Wang, E.; Peng, S.; Yin, S.; Jiang, Q.; Liu, M. Development history and prospect of the mechanism of coal and gas outburst in China over 70 years. *Coal Geol. Explor.* **2023**, *51*, 59–94.
2. Chen, L.; Chen, J. Study on the prediction effect of coal and gas outburst based on data imputation-machine learning. *China Saf. Sci. J.* **2022**, *18*, 69–74.
3. Lin, H.; Zhou, J.; Jin, H.; Li, S.; Zhao, P.; Liu, S. A collaborative prediction method for coal and gas outburst hazard level based on feature selection and machine learning. *J. Min. Saf. Eng.* **2023**, *40*, 361–370.
4. Ping, Y.; Bo, L.; Xue, W.; Hao, S.; Qiao, C.; Jiang, X. Study on the Evolution Mechanism of Coal and Rock Seepage under Cyclic Loading and Unloading of Pore Pressure. *J. China Univ. Min. Technol.* **2023**, *52*.
5. Skoczylas, N. Laboratory study of the phenomenon of methane and coal outburst. *Int. J. Rock Mech. Min. Sci.* **2012**, *55*, 102–107. [CrossRef]
6. Zhao, W.; Cheng, Y.; Jiang, H.; Jin, K.; Wang, H.; Wang, L. Role of the rapid gas desorption of coal powders in the development stage of outbursts. *J. Nat. Gas Sci. Eng.* **2016**, *28*, 491–501. [CrossRef]
7. Wold, M.B.; Connell, L.D.; Choi, S.K. The role of spatial variability in coal seam parameters on gas outburst behaviour during coal mining. *Int. J. Coal Geol.* **2008**, *75*, 1–14. [CrossRef]
8. Dazhao, S.; Minggong, G.; Gang, Y.; Liming, Q. Mechanism of Coal and Gas Outburst Induced by Non synchronous Deformation of Soft and Hard Coal Stratification. *Coal Sci. Technol. Mag.* **2023**, *2023*, 44.
9. Zhang, C.; Wang, P.; Wang, E.; Xu, J.; Li, Z.; Liu, X.; Peng, S. 70 year development process and prospects of coal and gas outburst mechanisms in China. *Coal Geol. Explor.* **2023**, *2023*, 51.
10. Hu, Q. Mechanical action mechanism of coal and gas outburst process. *J. Coal Sci.* **2008**, *3*, 1368–1372.
11. Lijun, Z.; Wentuan, Q.; Wei, Y. Research progress on coal and gas outburst mechanism based on geological structure perspective. *Saf. Coal Mines* **2022**, *2022*, 53.
12. Liang, W.; Bi, S.; Qing, T.; Yi, L. Experimental Study on Strata Fracture Evolution Characteristics of middling coal in the Process of Coal and Gas Outburst. *J. Saf. Sci. Technol.* **2023**, *2023*, 19.
13. Peng, S.; Yang, W.; Zhou, B.; Xu, J.; Cheng, L.; Yang, X. Study on the spatiotemporal evolution law of gas pressure during coal and gas outburst under true three-dimensional stress state. *Chin. J. Rock Mech. Eng.* **2020**, *39*, 1762–1772.
14. Guo, P. Study on the Development Mechanism of Coal and Gas Outburst Induced by Layer Fracture. Ph.D. Thesis, China University of Mining and Technology, Xuzhou, China, 2014.
15. Xu, M.; Dong, K.; Dong, Y. Microelement destruction and outburst mechanism of coal-rock and gas. *J. Xi'an Univ. Sci. Technol.* **2014**, *34*, 249–254.
16. Shan, D. Investigation of sensitive indicators and critical values for coal seam outbursts dominated by gas as the main factor. *Coal* **2021**, *30*, 20–22+93.
17. Si, H.; Zhao, J.; Hu, Q. Analysis of coal and gas outburst accidents under the theory of big data. *J. Xi'an Univ. Sci. Technol.* **2018**, *38*, 515–522+537.
18. Wen, C. Attribute recognition model and examples for risk assessment of tunnel gas outburst. *J. China Coal Soc.* **2011**, *36*, 1322–1328.
19. Wang, G.; Cheng, W.; Xie, J.; Zhou, G. Analysis of the role of gas content in the outburst process. *J. China Coal Soc.* **2011**, *36*, 429–434.
20. Cao, S.; Zhang, Z.; Li, Y.; Ping, G.; Yan, B. Experimental study on adsorption and desorption gas deformation characteristics of dangerous coal prone to outbursts. *J. China Coal Soc.* **2013**, *38*, 1792–1799.
21. Yang, D.; Peng, K.; Zheng, Y.; Chen, Y.; Zheng, J.; Wang, M.; Chen, S. Study on the characteristics of coal and gas outburst hazard under the influence of high formation temperature in deep mines. *Energy* **2023**, *268*, 126645. [CrossRef]
22. Li, Y.; Zhang, Y.; Zhang, Z.; Jiang, B. Experimental study on initial characteristics of coalbed methane desorption. *J. China Coal Soc.* **2013**, *38*, 15–20.
23. Xiang, Z. Optimization for Prediction Indexes of Coal and Gas Outburst in Coal Roadway Excavation Working Face Based on Grey Synthetically Relational Degree. *Adv. Mater. Res.* **2012**, *1792*, 524–527.

24. Zhou, X.; Sun, J. Gas Outburst Quantity Prediction Based on Improved BP Neural Network Using Principal Component Analysis. *Min. Saf. Environ. Prot.* **2018**, *45*, 43–47.
25. Liu, X.; Li, P. Soft Measurement Modeling of Coal and Gas Outburst Based on PCA and IFOA-BP. *Min. Res. Dev.* **2018**, *38*, 109–115.
26. Cao, B.; Bai, G.; Li, H. Prediction Analysis of Gas Content Based on PCA-GA-BP Neural Network. *China Saf. Sci. J.* **2015**, *11*, 84–90.
27. Ren, S. Application Research of PSO-DE Hybrid Algorithm Optimizing BP Neural Network in Coal and Gas Outburst Prediction. Ph.D. Thesis, Taiyuan University of Technology, Taiyuan, China, 2015.
28. Zhao, H. Application Research on Mine Gas Concentration Prediction Based on Support Vector Machine. *Coal Sci. Technol.* **2018**, *4*, 41–44.
29. Zeng, W.; Han, X.; Liu, J. Study on Gas Outburst Layer Based on Particle Swarm Optimization Support Vector Machine. *Coal Geol. China* **2018**, *30*, 44–46.
30. Zhang, W. Research on Gas Concentration Prediction Based on Support Vector Machine and Immune Genetic BP. Master's Thesis, Xi'an University of Science and Technology, Xi'an, China, 2017.
31. Wu, Y.; Li, H.; Xu, D. Coal and Gas Outburst Prediction Algorithm Based on IPSO-Powell Optimized SVM. *Ind. Autom. Min. Metall.* **2020**, *46*, 46–53.
32. Liu, X.; Li, Z.; Zhang, Z.; Zhang, G. Coal and gas outbursts prediction based on combination of hybrid feature extraction DWT+FICA–LDA and optimized QPSO-DELM classifier. *J. Supercomput.* **2022**, *78*, 2909–2936. [CrossRef]
33. Lu, G.; Guo, W.; Hu, H. Prediction of coal and gas outburst intensity based on KPCA-CS-ELM coupling model. *J. Appl. Funct. Anal.* **2018**, *20*, 218–226.
34. Qian, Y.; Qiu, L. Quantitative study on the value of resolution coefficient in grey correlation analysis. *Stat. Decis.* **2019**, *35*, 10–14.