fire

MDPI

Article

Research on Coal and Gas Outburst Prediction and Sensitivity Analysis Based on an Interpretable Ali Baba and the Forty Thieves–Transformer–Support Vector Machine Model

Yanping Wang, Zhixin Qin *, Zhenguo Yan, Jun Deng, Yuxin Huang, Longcheng Zhang, Yuqi Cao and Yiyang Wang

College of Safety Science and Engineering, Xi'an University of Science and Technology, Xi'an 710054, China; wang_yanping@xust.cn (Y.W.); yanzg@xust.edu.cn (Z.Y.); dengj518@xust.edu.cn (J.D.); 20120089017@stu.xust.edu.cn (Y.H.); 23220089040@stu.xust.edu.cn (L.Z.); 19403070424@stu.xust.edu.cn (Y.C.); 22320226003@stu.xust.edu.cn (Y.W.)

* Correspondence: 23220089060@stu.xust.edu.cn

Abstract: Coal and gas outbursts pose significant threats to underground personnel, making the development of accurate prediction models crucial for reducing casualties. By addressing the challenges of highly nonlinear relationships among predictive parameters, poor interpretability of models, and limited sample data in existing studies, this paper proposes an interpretable Ali Baba and the Forty Thieves–Transformer–Support Vector Machine (AFT-Transformer-SVM) model with high predictive accuracy. The Ali Baba and the Forty Thieves (AFT) algorithm is employed to optimise a Transformer-based feature extraction, thereby reducing the degree of nonlinearity among sample data. A Transformer-SVM model is constructed, wherein the Support Vector Machine (SVM) model provides negative feedback to refine the Transformer feature extraction, enhancing the prediction accuracy of coal and gas outbursts. Various classification assessment methods, such as TP, TN, FP, FN tables, and SHAP analysis, are utilised to improve the interpretability of the model. Additionally, the permutation feature importance (PFI) method is applied to conduct a sensitivity analysis, elucidating the relationship between the sample data and outburst risks. Through a comparative analysis with algorithms such as eXtreme gradient boosting (XGBoost), k-nearest neighbour (KNN), radial basis function networks (RBFNs), and Bayesian classifiers, the proposed method demonstrates superior accuracy and effectively predicts coal and gas outburst risks, achieving 100% accuracy in the sample dataset. The influence of parameters on the model is analysed, highlighting that the coal seam gas content is the primary factor driving the outburst risks. The proposed approach provides technical support for coal and gas outburst predictions across different mines, enhancing emergency response and prevention capabilities for underground mining operations.

Keywords: transformer-based feature extraction; classification assessment methods; interpretable model; SHAP analysis; permutation feature importance

check for updates

Academic Editor: Zihe Gao

Received: 26 December 2024
Revised: 17 January 2025
Accepted: 19 January 2025
Published: 22 January 2025

Citation: Wang, Y.; Qin, Z.; Yan, Z.; Deng, J.; Huang, Y.; Zhang, L.; Cao, Y.; Wang, Y. Research on Coal and Gas Outburst Prediction and Sensitivity Analysis Based on an Interpretable Ali Baba and the Forty Thieves–Transformer–Support Vector Machine Model. *Fire* **2025**, *8*, 37. https://doi.org/10.3390/fire8020037

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Coal remains a vital component of global energy systems, playing a significant role in ensuring energy security and driving industrial progress. However, with the depletion of shallow coal resources and increasing mining depths, coal and gas outburst accidents have become critical challenges that threaten the normal operation of mines [1,2]. As of 2019, over 878 coal and gas outburst incidents were reported across 22 underground coal

mines in Australia [3]. On 12 January 2024, a coal and gas outburst accident in the Tianan Coal Mine, Henan Province, China, resulted in 16 fatalities and 5 injuries. Similarly, on 7 June 2024, a gas outburst occurred in the Pingdingshan Coal Mine, Yunnan Province, China, causing three fatalities and five injuries. Despite the rapid advancements in artificial intelligence and machine learning technologies in recent years, which have introduced numerous new methods and theories to address the challenges of coal and gas outbursts, such accidents continue to occur frequently. The prevention and control of coal and gas outbursts remain a significant challenge, and urgent measures are required to address these issues [4,5].

To tackle the problem of coal and gas outbursts, significant scientific research and engineering practices have been conducted, leading to notable achievements. Early international research widely recognised the comprehensive action theory, which posits that coal and gas outbursts are determined by three factors: ground stress, coal properties, and gas content [6,7]. J. Hanes et al. [8] conducted studies in Australia on geological conditions, coal properties, and gas parameters, concluding that coal and gas outbursts are jointly triggered by stress and gas content, with gas content being the dominant factor. Similarly, K. Sato et al. [9] utilised digital seismographs across entire coal mines in Japan to analyse the impact of geological variations on gas outbursts. J. Shepherd et al. [10] studied parameters such as ground stress and gas content in Germany, identifying them as the most critical factors influencing coal and gas outbursts.

With the rapid development of artificial intelligence and machine learning models, numerous new theories have been applied to the study of coal and gas outbursts in mines. Peng Ji et al. [11] developed a coal mine data model based on the HPO-BiLSTM algorithm, achieving a method for coal and gas outburst early warning. Junqi Zhu et al. [12] constructed a coal and gas outburst risk identification model using the RS-GA-BP hybrid model, which significantly improved the risk identification speed. Other researchers have employed mathematical theories and data mining techniques to establish coal and gas outburst risk evaluation and identification systems. For instance, Xie Xuecai et al. [13] utilised data mining and the Apriori algorithm to analyse the causes of coal and gas outbursts, developing a Bayesian network model to conduct a sensitivity analysis of accident occurrences. Wei Wang et al. [14] applied extension theory to construct a risk prediction and risk grading indicator system for coal and gas outbursts, successfully predicting risks in 12 high-gas mines. David R. Hanson et al. [15] analysed MSHA accident data and applied various algorithms to coal samples in Pennsylvania, examining the probability of outburst accidents based on stratigraphic chemistry and lithofacies data. Furthermore, some researchers have explored traditional algorithms such as decision tree models. Zheng Xiaoliang et al. [16] combined meta-heuristic algorithms with the XGBoost theory to achieve quantitative analyses of gas outburst predictions. Zhonghui Li et al. [17] investigated a risk assessment for coal and gas outbursts based on logistic regression models, constructing a non-contact EMR index and achieving a regression prediction accuracy of 94%. Finally, geophysical methods have also been employed to develop various evaluation approaches and equipment. For example, V. Frid et al. [18] monitored high-frequency electromagnetic waves emitted from rock fractures to detect outburst risks and compared these with laboratory results, bridging the gap between laboratory and field predictions. Similarly, Janathan P. Mathews et al. [19] utilised X-ray computed tomography to study the expansion and contraction of gas absorption and desorption in coal under confining pressure, providing novel methodologies for coal and gas outburst research.

In practical coal mining operations, predictive models for coal and gas outbursts must effectively address challenges such as limited sample data, significant nonlinear relationships between parameters, and poor model interpretability, thereby achieving

high generalisation capability and prediction accuracy. To address these challenges, this study constructs an interpretable Ali Baba and the Forty Thieves–Transformer–Support Vector Machine (AFT-Transformer-SVM) model based on multidimensional data from actual mining faces and designs a series of computational experiments. By employing a SHAP analysis to interpret feature contributions and combining TP, TN, FP, and FN metrics for comprehensive performance evaluation, the model's interpretability is significantly enhanced. Additionally, the permutation feature importance (PFI) method is adopted to quantify the sensitivity of raw data features, enabling the development of a coal and gas outburst early warning model under real-world conditions. Through comparative validation with various classical algorithms and field data, the proposed model demonstrates superior prediction accuracy and efficiency, offering robust support for the development of high-performance, interpretable coal and gas outburst prediction models.

2. Materials and Methods

Figure 1 represents the primary model proposed in this study, which integrates the coal and gas outburst prediction model with the feature sensitivity analysis model. Initially, nine features were selected, encompassing three dominant factors: the physical and mechanical properties of coal, gas-related factors, and ground stress. These features were input into the AFT-Transformer-SVM model for analysis. The Ali Baba and the Forty Thieves (AFT) algorithm was employed to optimise the feature extraction capability of the Transformer module. During this process, the prediction accuracy of the coal and gas outburst model, derived from the SVM, was utilised as a convergence function to provide negative feedback, thereby adjusting the Transformer module's layer structure. This iterative optimisation ultimately facilitated accurate feature extraction by the Transformer.

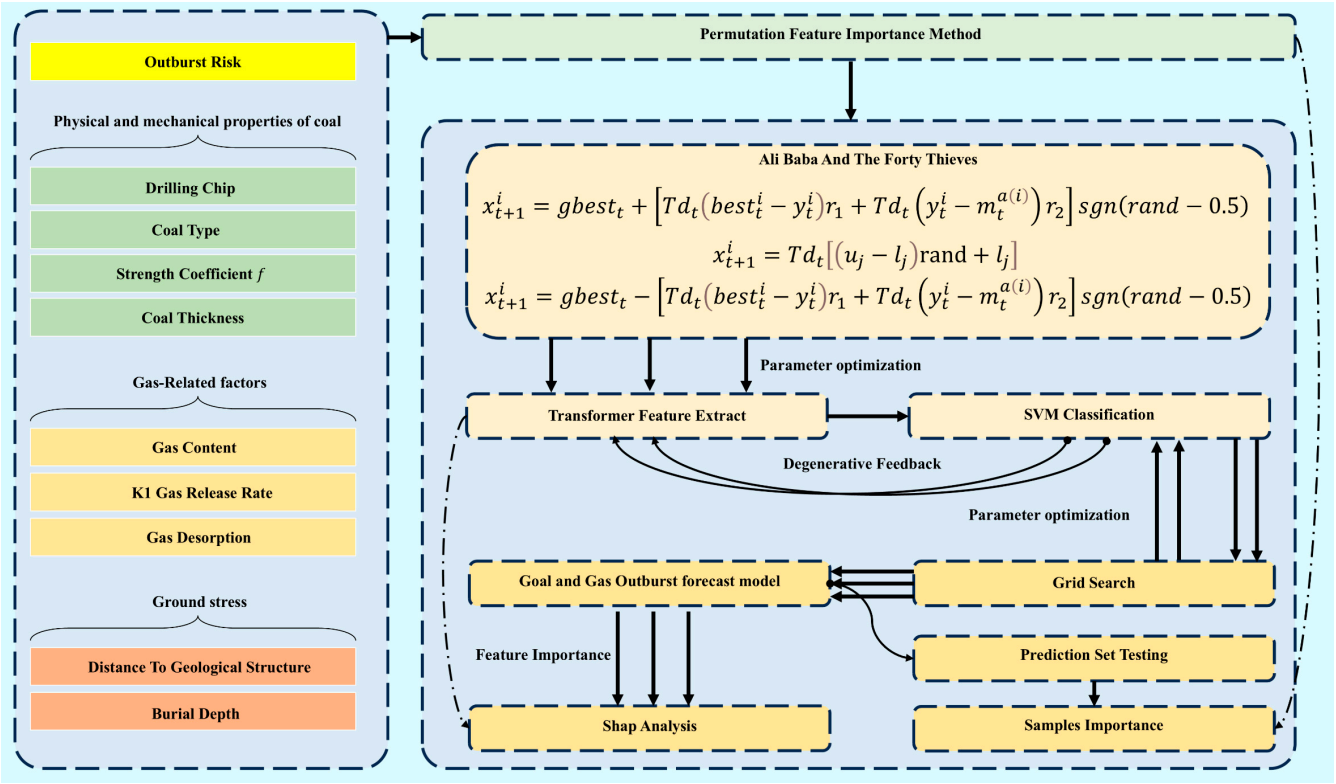


Figure 1. Schematic diagram of the overall model workflow.

Subsequently, a grid search algorithm was utilised to optimise the SVM layer structure, enabling coal and gas outburst testing on the test dataset. A SHAP analysis [20] was applied to examine the distribution of features extracted by the Transformer, offering interpretability for the current model. To quantify the contribution of the original sample data, the permutation feature importance (PFI) method [21] was employed to randomly shuffle the columns of the original feature data. The model was then retrained and predictions were performed. The prediction accuracy was used as a criterion to evaluate the contribution of the original data to the model.

2.1. Collection of Raw Data

To adapt to the conditions of on-site working environments, the primary factors influencing coal and gas outbursts were selected. These factors include gas-related parameters, the physical and mechanical properties of coal, and ground stress. The selected sample features consist of nine key parameters: coal failure type, initial velocity of gas desorption from coal, coal firmness coefficient, gas content in coal seams, K1 gas desorption volume, drilling debris volume, distance to geological structures, burial depth of the coal, and coal seam thickness. A total of 569 data samples were collected from an intelligent mine analysis platform for testing and validation purposes [22].

(1) Gas-Related Factors

Gas content: The gas content within coal seams is recognised as one of the critical factors contributing to gas-related disasters. Coal seams with high gas content, particularly those with loose structures and significant porosity, are prone to gas accumulation, thereby increasing the risk of outbursts. The gas storage capacity of coal seams is closely associated with the permeability of the surrounding rock. When the surrounding rock is dense and exhibits low permeability, gas release becomes restricted, further heightening the probability of outbursts.

K1 gas release rate: The K1 value represents the capacity of coal samples to desorb gas from the coal body, reflecting the gas storage conditions within the coal seam. A higher K1 value indicates a stronger gas release capability, which typically corresponds to a looser coal structure that facilitates gas release, thus elevating the risk of gas outburst accidents.

Gas desorption: The initial velocity of gas desorption reflects the rate at which gas is released from the coal body. This value is measured by applying a specific pressure to a coal sample and observing the pressure change within one minute. A higher gas desorption velocity indicates a faster release of gas, increasing the likelihood of gas outbursts.

(2) Physical and Mechanical Properties of Coal

Coal type: Under the influence of geological activities, including structural stress and ground stress, the original structure of coal may be damaged, resulting in the formation of numerous pore spaces. Coal failure types are classified into intact coal, fractured coal, highly fractured coal, pulverised coal, and completely pulverised coal, depending on the degree of damage. The extent of structural damage directly affects the mechanical properties and gas storage conditions of coal, thereby influencing the risk of gas outbursts.

Strength coefficient: The strength coefficient of coal is utilised to measure its compressive strength, hardness, and brittleness. A higher value indicates greater resistance to compressive forces and less susceptibility to damage. Conversely, a lower strength coefficient implies that the coal is more vulnerable to external forces, facilitating gas release under pressure and, thereby, increasing the probability of outbursts.

Drilling chip volume: Drilling chip volume refers to the amount of coal powder generated during the drilling process in coal seams. This parameter is closely related to the strength and brittleness of the coal. Weaker or more brittle coal seams produce larger

volumes of drilling chips, indicating a higher propensity for structural damage, thereby elevating the risk of gas outbursts.

Coal thickness: The thickness of coal seams directly determines the volume of gas stored within the coal body. Thicker coal seams generally contain more gas. When the coal body exhibits low strength, the accumulated gas is more likely to breach the constraints of the coal structure and escape, resulting in gas outbursts.

(3) Ground Stress

Distance to geological structures: Coal seams located near geological structures, such as faults and folds, are subjected to significant structural stress, increasing the risk of gas outbursts. As mining operations progress, the redistribution of structural stress may lead to gas accumulation. This is particularly evident in stress-concentrated regions, where the likelihood of outbursts is significantly heightened.

Burial depth: With increasing mining depth, ground stress intensifies progressively. Deeply buried coal seams are subjected to higher ground stress, resulting in tighter coal structures that hinder the release of gas. When gas accumulation reaches a critical threshold, an outburst disaster may occur. Therefore, the burial depth of coal seams and the associated ground stress are critical factors influencing the risk of gas outbursts.

2.2. AFT-Transformer-SVM Model for Coal and Gas Outburst Prediction

Figure 2 illustrates the Transformer-SVM coal and gas outburst prediction model, designed for multidimensional, nonlinear, and small-sample prediction tasks. The diagram demonstrates the connection between the raw data, the Transformer model framework, and the SVM classification model. Initially, raw sample data are fed into the Transformer model for training. The Transformer model is primarily composed of multi-head attention and positional embedding. Positional embedding explicitly incorporates sequence position information, while multi-head attention models the correlation within the sequence by feeding the input data into multiple attention heads to learn different feature representations. These features are then concatenated to allow the model to extract more comprehensive and multidimensional feature information. Subsequently, the extracted features are passed to a Flatten Layer, which converts the sequence into a one-dimensional vector, flattening the multidimensional features into a single vector. The vector is then processed through a Dense Layer, where nonlinear mapping is applied to extract higher-level and abstract features. Following this, the SVM model adopts a degenerative feedback Transformer structure to enhance feature extraction and improve the prediction performance. In the diagram, the left section represents the raw data, consistent with Figure 1, and depicts the original sample. The middle section represents the Transformer feature extraction module, showcasing various layers of the Transformer structure. The right section illustrates the SVM classification module, where the extracted features are learnt and compared against the outburst risk parameter to evaluate the prediction accuracy.

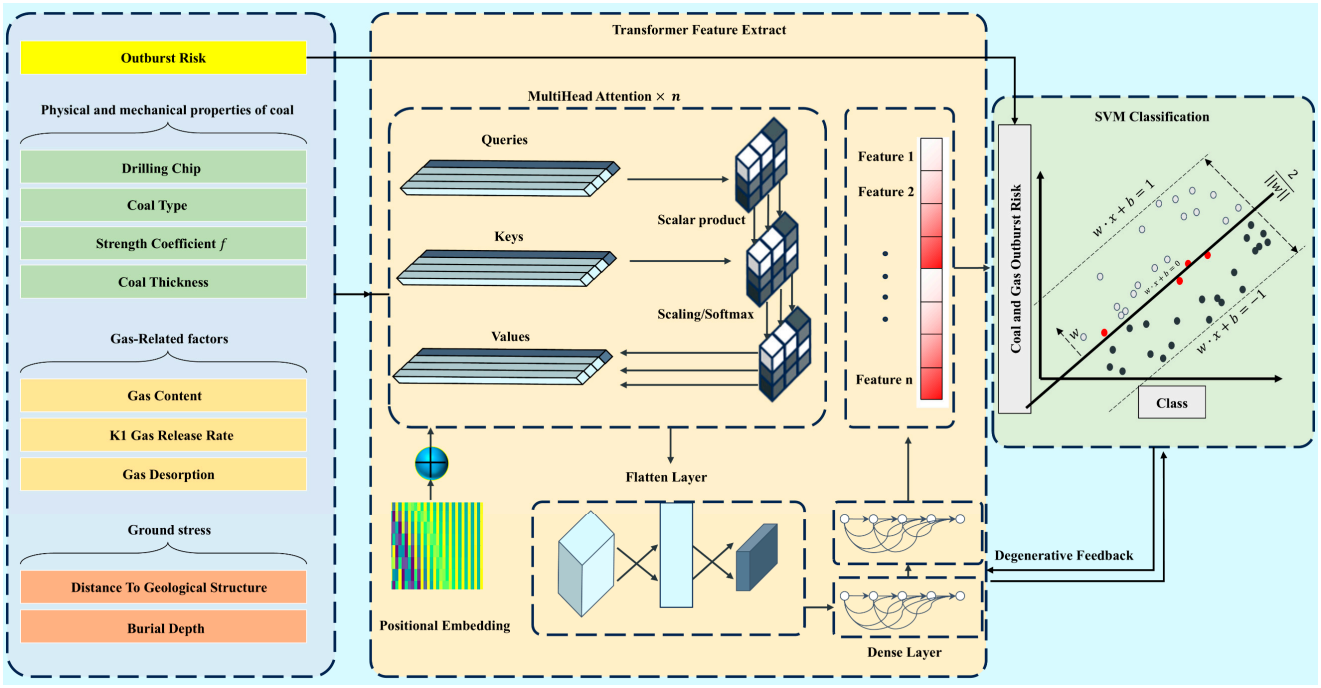


Figure 2. Schematic diagram of the Transformer-SVM model structure.

2.2.1. Ali Baba and the Forty Thieves Optimisation Algorithm

The AFT algorithm is an intelligent optimisation method suitable for solving multidimensional objective functions to obtain optimal solutions [23]. The fundamental principles of the algorithm are described as follows:

(1) Acquiring Information and Pursuing "Alibaba":

The algorithm first gathers information and pursues the optimal solution, as described in Equation (1):

$$x_{t+1}^{i} = gbest_t + \left[Td_t\left(best_t^{i} - y_t^{i}\right)r_1 + Td_t\left(y_t^{i} - m_t^{a(i)}\right)r_2\right]sgn(rand - 0.5) \quad (1)$$

Here, x_{t+1}^{i} denotes the position of the ith individual in the $t+1$ iteration; $gbest$ represents the global best position, indicating the global optimal solution identified during the current search process; Td_t is the tracking distance, which controls the magnitude of position updates; y_t^{i} is the current position of individual i in the tth iteration; $best_t^{i}$ is the historical best position of the individual i; and $m_t^{a(i)}$ is the reference position of the individual i.

(2) Random Exploration and Deception:

Individuals are deceived and randomly explore the solution space, as described in Equation (2):

$$x_{t+1}^{i} = Td_t\left[(u_j - l_j)\,\mathrm{rand} + l_j\right] \quad (2)$$

Here, u_j and l_j represent the upper and lower bounds, respectively, of the jth dimension in the search space.

(3) Balancing Global and Local Search:

The algorithm ensures a balance between global exploration and local exploitation to prevent premature convergence to local optima. This balance is achieved as described in Equation (3):

$$x_{t+1}^{i} = gbest_t - \left[Td_t\left(best_t^{i} - y_t^{i}\right)r_1 + Td_t\left(y_t^{i} - m_t^{a(i)}\right)r_2\right]sgn(rand - 0.5) \quad (3)$$

This optimisation approach enables efficient exploration of the solution space, ensuring a balance between exploitation and exploration for achieving the global optimal solution.

2.2.2. Transformer Feature Extraction

In the Transformer model, a positional encoding layer and self-attention mechanism are introduced to effectively process the sequential information of the input data. Initially, a positional embedding layer is incorporated, which utilises an embedding mechanism to assign a unique vector to each input, enabling the model to perceive positional information. Through an additive mechanism, positional embeddings are combined with the original features, allowing the Transformer model to fully comprehend the spatial relationships within the sequential data.

Given the input sequence $X = [x_1, x_2, \ldots, x_n]$, where $x_i \in \mathbb{R}^d$ represents the feature vector at the ith position of the input, the positional encoding $P = [p_1, p_2, \ldots, p_n]$ generates the positional embedding vector $p_i \in \mathbb{R}^d$ for each position i. The positional encoding is then additively combined with the input features, as described in Equation (4):

$$X' = X + P \tag{4}$$

where X' represents the input data after the addition of positional encoding.

Next, a multi-head self-attention layer is introduced. By parallel computing multiple attention weight matrices, the model is enabled to capture critical information from the input data across different subspaces simultaneously. Through the stacking of two self-attention layers, the model effectively learns the dependencies among features at various levels, enhancing its capability to recognise complex patterns. The computation for each attention head is as follows:

(1) Calculation of Query (Q), Key (K), and Value (V):

For each attention head h, the Query Q_h, Key K_h , and Value V_h are computed as follows:

$$Q_h = X' W_Q^{(h)} \tag{5}$$

$$K_h = X' W_K^{(h)} \tag{6}$$

$$V_h = X' W_V^{(h)} \tag{7}$$

where $W_Q^{(h)}$, $W_K^{(h)}$, and $W_V^{(h)}$ are the learnt weight matrices.

(2) Attention Score:

The attention score is calculated using the Query and Key as follows:

$$Attention(Q_h, K_h) = \frac{Q_h K_h^T}{\sqrt{d_k}} \tag{8}$$

In the equation, d_k represents the dimensionality of the Key vector, and the normalisation factor $\sqrt{d_k}$ is used to prevent the values from becoming excessively large.

(3) Weighted Sum Calculation:

The attention scores are used to compute the weighted sum of the values, as shown in Equation (9).

$$Output_h = softmax\left(\frac{Q_h K_h^T}{\sqrt{d_k}}\right) V_h \tag{9}$$

(4) Multi-Head Attention Combination:

The outputs from all attention heads are concatenated and passed through a linear transformation to obtain the final attention output, as shown in Equation (10):

$$MultiHeadOutput = Concat(Output_1, \ldots, Output_h)W_O \tag{10}$$

Here, W_O is a learnt linear transformation matrix.

Finally, a fully connected layer and classification output are applied. Based on the self-attention mechanism, the extracted features are flattened through a flattening layer and further subjected to a series of fully connected (Dense) layers for nonlinear mapping. The output layer utilises the softmax activation function for multi-class classification, providing the final prediction for the coal and gas outburst discrimination task, as shown in Equations (11) and (12):

(5) Flattening and Fully Connected Layer:

The flattened output Z is mapped through the first fully connected layer:

$$F_1 = Relu(ZW_1 + b_1) \tag{11}$$

Here, W_1 and b_1 represent the weight matrix and bias term of the first fully connected layer, respectively, and Relu is the nonlinear activation function.

(6) Output Layer:

Multi-class classification is performed through the second fully connected layer and the softmax activation function:

$$\hat{y} = softmax(F_1W_2 + b_2) \tag{12}$$

Here, W_2 and b_2 are the weight matrix and bias term of the second fully connected layer, respectively, and $\hat{y}$ represents the predicted probability distribution output by the model.

2.2.3. Grid-Optimised Support Vector Machine Algorithm

Grid search is an exhaustive search method used to optimise the hyperparameters of a Support Vector Machine (SVM). The performance of an SVM model relies heavily on the proper selection of hyperparameters, such as the kernel function parameters and the regularisation parameter. The primary objective of grid optimisation is to identify the optimal combination of these hyperparameters to maximise the model's performance on the validation set.

(1) Fundamental Formula of SVM

Since coal and gas outburst data are nonlinear, even after feature extraction using the Transformer model, a Gaussian radial basis function (RBF) kernel is required to map the data to a higher-dimensional space. The formula for the Gaussian RBF kernel is given in Equation (13):

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\| \mathbf{x}_i - \mathbf{x}_j \|^2}{2\sigma^2}\right) \tag{13}$$

(2) Grid-Optimised SVM

Grid optimisation identifies the optimal hyperparameter combination that achieves the best model performance. It involves optimising the regularisation parameter C and the kernel parameter σ. The steps for grid optimisation are as follows:

Define the hyperparameter search space:

$$\mathcal{H} = \{(C, \gamma) \mid C \in \mathcal{C}, \gamma \in \Gamma\} \tag{14}$$

Here, $\mathcal{C}$ represents the set of candidate values for C, and Γ represents the set of candidate values for the RBF kernel parameter $\gamma = 1/\left(2\sigma^2\right)$.

Cross-validation evaluation: For each hyperparameter combination $(C, \gamma) \in \mathcal{H}$, k-fold cross-validation is performed to evaluate the model's performance:

$$\mathrm{CV}_{\mathrm{score}}(C, \gamma) = \frac{1}{k} \sum_{i=1}^{k} \mathrm{Score}_i(C, \gamma) \tag{15}$$

Here, $\mathrm{Score}_i(C, \gamma)$ represents the model's accuracy on the ith fold of the data, and k denotes the number of folds in the cross-validation process.

Hyperparameter selection: All hyperparameter combinations are iterated, and the combination that maximises the cross-validation score is selected as follows:

$$(C^*, \gamma^*) = \arg \max_{(C,\gamma)\in\mathcal{H}} \mathrm{CV}_{\mathrm{score}}(C, \gamma) \tag{16}$$

Model training with optimal hyperparameters: Finally, the SVM model is retrained on the training set using the optimal hyperparameters (C^*, γ^*):

$$\hat{f}(x) = \mathrm{SVM}(x; C^*, \gamma^*) \tag{17}$$

2.3. Data Sensitivity Analysis

In the process of data analysis, a combination of SHAP analysis and permutation feature importance (PFI) methods is utilised to evaluate the importance of data features.

2.3.1. Permutation Feature Importance

The PFI method is applied within the Transformer-SVM model by randomly permuting the values of the specific feature j. The permuted data are then used for coal and gas outburst prediction with the Transformer-SVM model to assess the model's performance. The prediction accuracy is analysed to measure the impact of different features on the overall prediction performance. The fundamental formulas are as follows:

First, let $M(model, X)$ denote the prediction accuracy of the model on data X, and let $X_{perm}^{(j)}$ represent the new matrix obtained by randomly permuting the jth column of the feature matrix X. The model performance on the permuted data is expressed using Equation (18):

$$M_j^{(perm)} = M\left(AFT - Transformer - SVM, X_{perm}^{(j)}\right) \tag{18}$$

Based on the performance degradation, the reduction in model performance is defined by Equation (19):

$$PFI_j = M_{base} - M_j^{(perm)} \tag{19}$$

2.3.2. SHAP Analysis for Interpreting the Coal and Gas Outburst Model

The SHAP analysis is used to explain the influence of parameters extracted by the Transformer-SVM model on the model's predictions. The Shapley value from game theory is employed to determine the contribution of each participant (feature) to the final prediction of the coal and gas outburst model. The model is expressed as shown in Equation (20):

$$f(x) = SVM(T(x)) \tag{20}$$

Here, $T(x)$ represents the process in which the Transformer maps the original sample x to the new feature $p = T(x)$.

For a single sample $x \in \mathbb{R}^d$ and the model $f(\cdot)$, let the feature set $F = \{1, 2, \ldots, d\}$. The SHAP value $\phi_j(x)$ for the jth feature is calculated as defined in Equation (21):

$$\phi_j(x) = \sum_{S \subseteq F\{j\}} \frac{|S|!(d - |S| - 1)!}{d!} [f(S \cup \{j\} - f(S)] \tag{21}$$

where $|S|$ represents the size of the feature subset S; and $f(S)$ denotes the model's prediction for sample x using only the feature subset S.

In practical applications, to reduce the computational complexity of enumerating all subsets S, the kernel SHAP method is employed to approximate the SHAP values and minimise the computational burden.

3. Results and Discussion

This study utilised a dataset of 569 samples to construct the AFT-Transformer-SVM model using the TensorFlow library within Python 3.9. A series of computational comparison experiments were designed to assess and validate the coal and gas outburst risk based on real-world data, providing technical support for safe production in actual mining operations.

Table 1 presents a subset of the sample data used for the analysis, which is consistent with Figure 1 and the descriptions provided earlier. Nine variables were selected as sample features, with the outburst risk levels represented by drilling operations at the mining site. No specific measures taken were considered as no outburst risk, represented by 0.1. Conducting 12–15 sets of pressure relief boreholes indicated a moderate outburst risk, represented by 0.6. Drilling 20 sets of pressure relief boreholes was considered to represent a severe outburst risk, represented by 1.

Table 1. Collected sample data table.

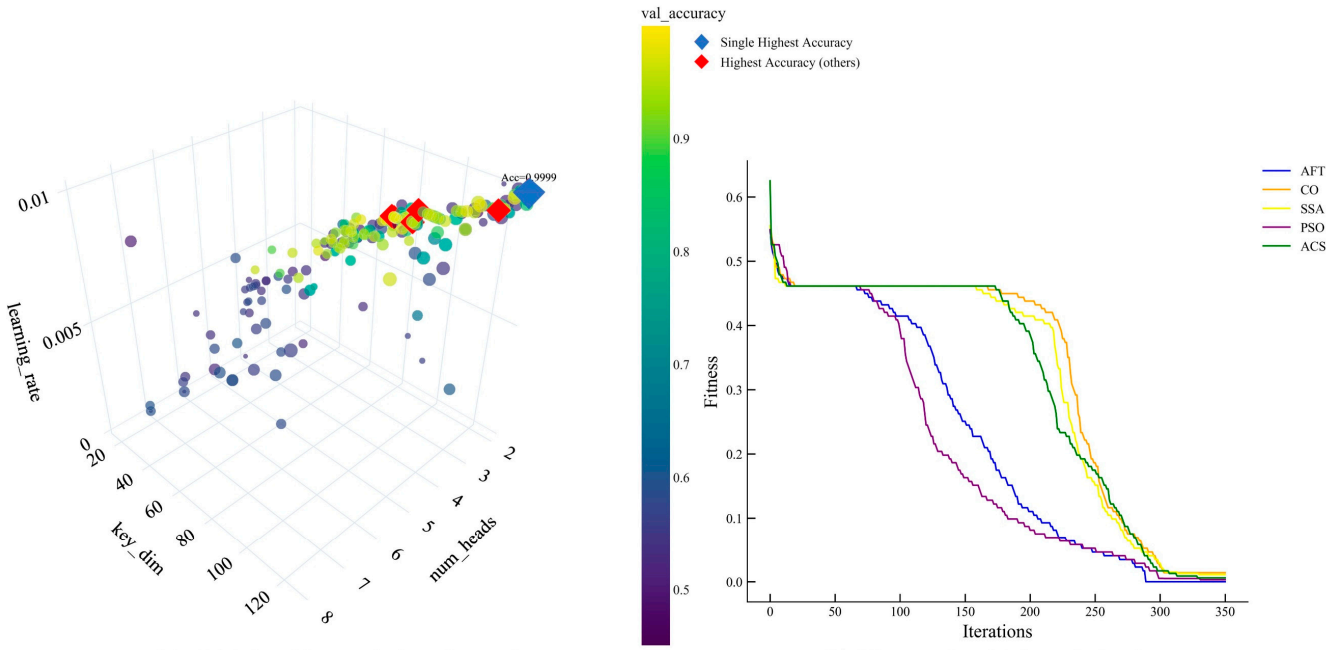
Coal Type	Gas Desorption Δp	Strength Coefficient f	Gas Content /$m^3 \cdot t^{-1}$	$K1$ Gas Release Rate $mL \cdot (g \cdot min^{0.5})^{-1}$		
2	16.86	0.31	7.13	0.4	0.27	0.29
2	17.83	0.30	10.62	0.42	0.21	0.24
2	17.78	0.35	7.52	0.22	0.16	0.18
3	17.09	0.32	11.44	0.11	0.25	0.21
...	...	...	...	...	...	...

Distance to Geological Structure/m	Burial Depth /m	Coal Thickness/m	Drilling Chip/$kg \cdot m^{-1}$			Outburst Risk
32	366	0.8	3.7	3.4	3.9	1
0	358	2.2	3.1	3.6	3	0.6
42	352	2.2	3.5	3.4	3.8	0.1
62	370	2.5	3.2	3.2	3.2	0.6
...	...	...	...	...	...	...

3.1. Optimisation of the Transformer-SVM Model Using Different Algorithms

Firstly, the original data were split in a 7:3 ratio, with the first 70% used for the model optimisation. Various optimisation algorithms were employed to optimise the parameters of the Transformer-SVM model, including units_fc1, num_heads, key_dim, and learning_rate.

The optimisation process was conducted using a fitness function defined as (1−accuracy). A population of 30 individuals was selected, and the optimisation process was iterated 350 times, as illustrated in Figure 3.



(a) AFT algorithm optimizes the results **(b) Fitness of multiple optimization algorithms**

Figure 3. Comparative results of various optimisation algorithms.

Figure 3a presents the optimisation results using the AFT algorithm, where the point with the highest accuracy reached 0.999. The size of the points in the figure represents units_fc1, the x-axis represents num_heads, the y-axis represents key_dim, the z-axis represents learning_rate, and the colour bar indicates the prediction accuracy. The final optimised parameters were determined as follows: [num_heads: 4, key_dim: 98, learning_rate: 0.009449, units_fc1: 64].

Figure 3b shows the relationship between the fitness function and the number of iterations for various optimisation algorithms. The blue curve represents the AFT algorithm, where the fitness value reached 0.00012 after 289 iterations, making it the algorithm with the smallest fitness value. This was followed by the PSO algorithm, with a fitness value of 0.0032, and then the CO, SSA, and ACS optimisation algorithms.

Through experimental comparisons, each algorithm was run ten times, and the optimal fitness values of the various optimisation algorithms were selected for a statistical analysis. The results are presented in Table 2, where the AFT algorithm demonstrated superior performance in terms of range, minimum, maximum, and standard deviation compared to the other algorithms. However, it exhibited a relatively higher variance. Following the AFT algorithm were the PSO and SSA algorithms, with the CO and ACS algorithms performing the least effectively.

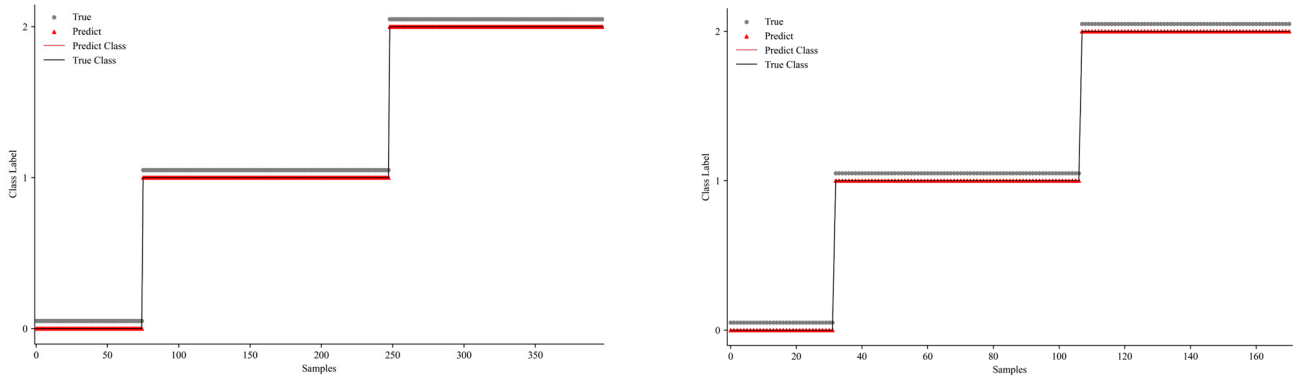
Table 2. Comparative results of various optimisation algorithms.

Optimization Algorithms	Variance	Range	Minimum	Maximum	Standard Deviation
AFT	0.00004	0.0004	0.000110	0.00015	0.000011
CO	0.00001	0.003	0.013	0.016	0.000884
SSA	0.00003	0.003	0.006	0.009	0.000936
PSO	0.00002	0.0018	0.0032	0.005	0.000503
ACS	0.00013	0.012	0.016	0.028	0.003606

These results validate the efficiency and accuracy of the AFT algorithm in optimising the parameters of the Transformer-SVM model for coal and gas outburst predictions. Therefore, the AFT algorithm was chosen as the final intelligent optimisation algorithm.

3.2. AFT-Transformer-SVM Model for Coal and Gas Outburst Risk Prediction

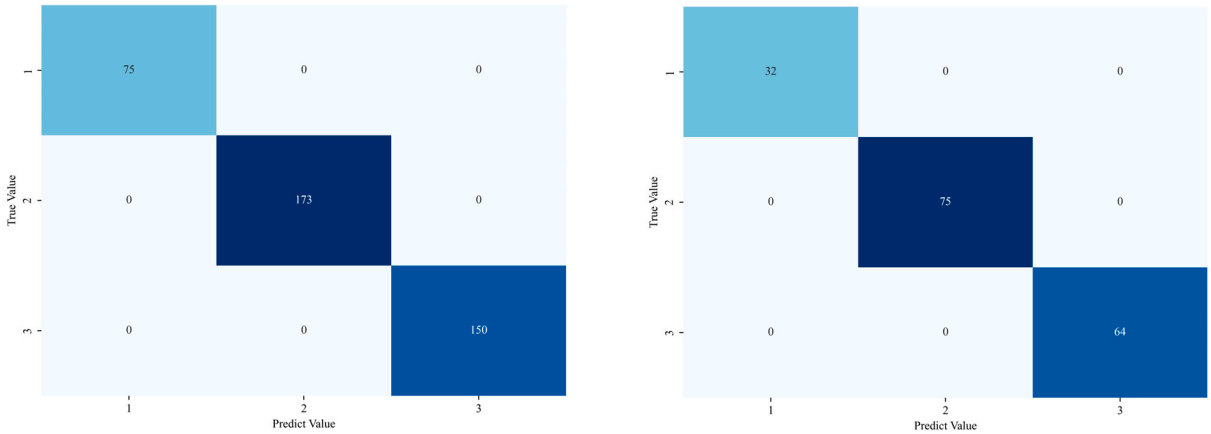
To evaluate the performance of the model, 70% of the sample data were used as training data, while the remaining 30% were used as testing data to predict coal and gas outburst risks. Figure 4 illustrates the prediction results of the AFT-Transformer-SVM model. In the figure, the true values are represented by grey circles, while the prediction results are represented by red triangles. The "Predict Class" line indicates the predicted results, and the "True Class" line represents the actual results. In the sample data, the two lines overlap completely, indicating that the prediction results are 100% accurate. Figure 4a shows the results for the training set, and Figure 4b displays the results for the test set.



(**a**) train set predict results (**b**) test set predict results

Figure 4. Prediction results of the AFT-Transformer-SVM model.

Figure 5 presents the corresponding confusion matrices, where all predictions lie along the diagonal, signifying perfect prediction accuracy. The three colours in the figure represent three categories: 1, 2, and 3. The numerical values in the table indicate the number of instances in each category. Figure 5a corresponds to the results for the training set, while Figure 5b corresponds to the results for the test set.



(**a**) train set predict results (**b**) test set predict results

Figure 5. Confusion matrix of the prediction results for the AFT-Transformer-SVM model.

Table 3 presents the sensitivity analysis of the prediction results, including eight parameters: recall, precision, F1-score, accuracy, sensitivity, specificity, AUC, and Kappa. These metrics involve four key parameters: True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN). TP represents instances where the actual label is positive, and the model predicts positive; FN represents instances where the actual label is positive, but the model predicts negative; FP represents instances where the actual label is negative, but the model predicts positive; and TN represents instances where the actual label is negative, and the model predicts negative [24–26].

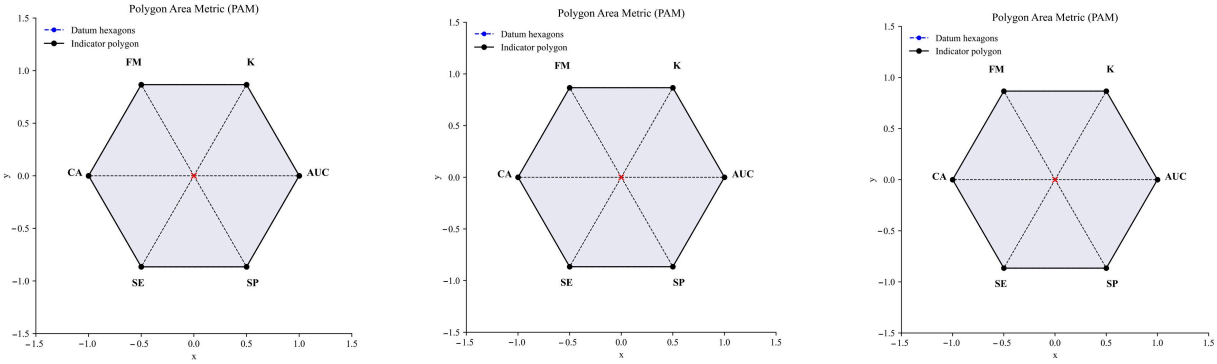
Table 3. Sensitivity analysis of prediction results.

Recall/%	Precision/%	F1-Score/%	Accuracy/%	Sensitivity/%	Specificity/%	AUC/%	κ/%
100	100	100	100	100	100	100	100

$Recall = \frac{TP}{TP+FN}$, representing the proportion of positive samples correctly predicted as positive. $Percision = \frac{TP}{TP+FP}$, representing the proportion of correctly predicted positive samples among all predicted positive samples. $F1 = 2 \times \frac{Percision \times Recall}{Percision+Recall}$, combining precision and recall to indicate the proportion of perfect classification. $Accuracy = \frac{TP+TN}{All\ Samples}$, indicating the proportion of correctly classified samples, whether positive or negative. Sensitivity is similar to recall. $Specificity = \frac{TN}{TN+FP}$, representing the proportion of negative samples correctly classified as negative. AUC refers to the area under the ROC curve, representing the model's discriminative ability across different thresholds. $\kappa = \frac{p_0 - p_e}{1-p_e}$, where p_0 denotes the observed accuracy, and p_e represents the expected accuracy under random guessing, indicating the model's accuracy after accounting for random agreement.

The prediction results shown in Table 3 indicate that the model has perfectly learnt the features of the sample data and performed accurate predictions.

Figure 6 illustrates the Polygon Area Metric (PAM) results of the AFT-Transformer-SVM model, with subplots a, b, and c representing the polygon area predictions for the three categories. In the figure, FM represents F_measure, or the F1-Score. K represents κ. SP denotes specificity. SE denotes sensitivity. CA represents accuracy. The prediction results for all three categories are entirely correct, confirming the feasibility of the model.



(**a**) (**b**) (**c**)

Figure 6. PAM results of AFT-Transformer-SVM predictions.

Figure 7 shows the ROC curve of the model's predictions. The actual curve starts from the point (0, 0), progresses to (0, 1), and ends at (1, 1), indicating that the model's predictions perfectly align with the actual results. The curves for Class 1, Class 2, and Class 3 overlap and converge at the top-right corner, further demonstrating the model's outstanding performance.

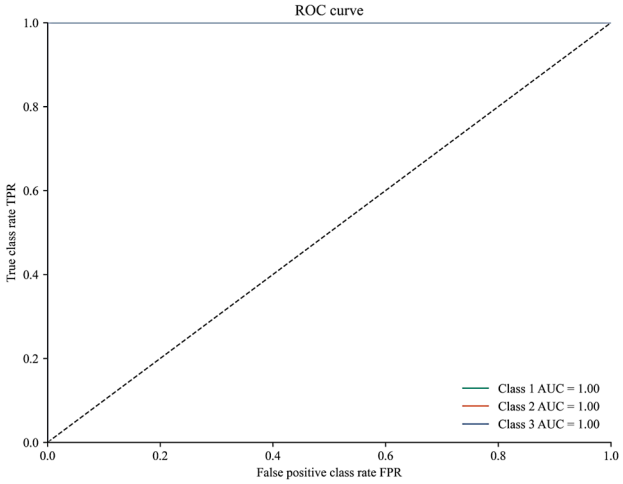


Figure 7. ROC curve of prediction results.

To interpret the model, a SHAP value analysis was conducted, and the results are shown in Figure 8. The figure illustrates three features extracted by the Transformer model. The SHAP analysis demonstrates clear feature distribution trends. For Feature 2, the relationship between the SHAP values and feature values is relatively monotonic: higher feature values (red) positively contribute to SHAP values, while lower feature values (blue) suppress SHAP values. This indicates that the model's treatment of Feature 2 is close to linear, where higher feature values have the greatest positive impact on the prediction results. For Feature 0, a reverse monotonic relationship is observed between the SHAP values and feature values: lower feature values (blue) positively contribute to the SHAP values, while higher feature values (red) suppress the SHAP values. The distribution trend is clear, indicating that the model processes Feature 0 in a nearly linear manner. For Feature 1, the SHAP value exhibits a certain level of complexity, where higher feature values (red) suppress the SHAP values, while middle SHAP values are not fully monotonic, showing both positive and negative SHAP values. This suggests that the model processes Feature 1 with some degree of nonlinearity.

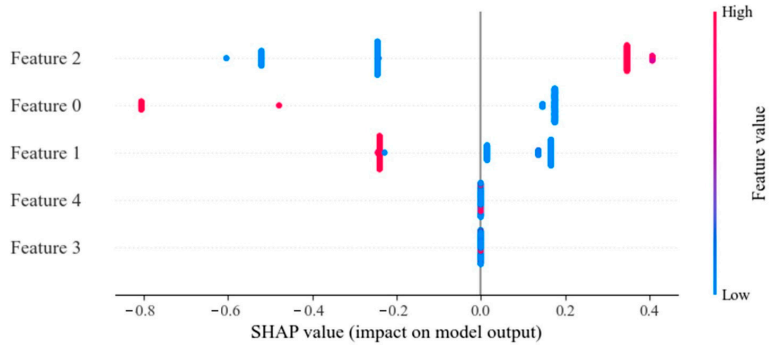


Figure 8. SHAP analysis of feature extraction in the AFT-Transformer-SVM model.

Figure 9 presents the SHAP analysis results for the original data. The figure shows the relationships between the feature values of various raw data variables and their SHAP values. Most variables exhibit highly nonlinear relationships, with the SHAP values distributed between −0.5 and 0.3, indicating a relatively small influence. Among these, gas content is the feature with the widest distribution and the highest degree of nonlinearity, followed by gas release rate, with other variables showing smaller contributions. A comparison between Figures 8 and 9 demonstrates that feature extraction reduces the nonlinearity

of the original multidimensional data and increases the association between feature values and SHAP values.

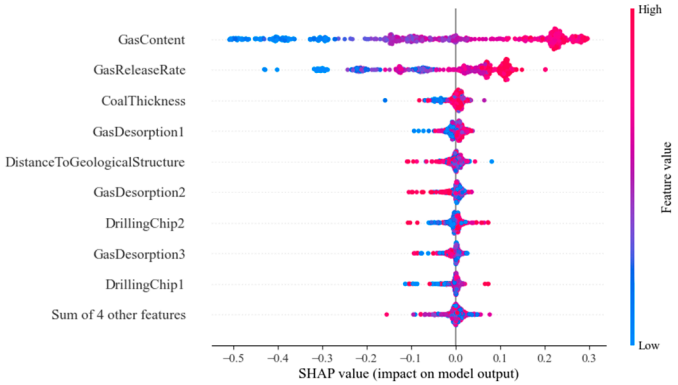


Figure 9. SHAP analysis of original sample features.

Figure 10 shows the waterfall plots for the training and testing samples. Figure 10a represents the training set features. In samples 0–100, Feature 3 and Feature 1 contribute strongly negatively to the model output (blue), while Feature 2 contributes positively (red). For samples 100–200, the positive contributions of Feature 3 and Feature 2 gradually increase, with the overall colour shifting towards blue, and Feature 1 shows a more substantial positive contribution. In samples 200–300, the contributions of Feature 1 and Feature 2 increase significantly (evidenced by the expansion of the red area), driving a marked rise in prediction values. In samples 300–400, the contributions of all three features increase (indicated by the red area), resulting in an upward trend in the model output. Figure 10b illustrates the feature contributions in the test set, which exhibit a similar overall trend to the training set. However, a noticeably varying region is observed within samples 0–50, where the model output shows an increase in this range.

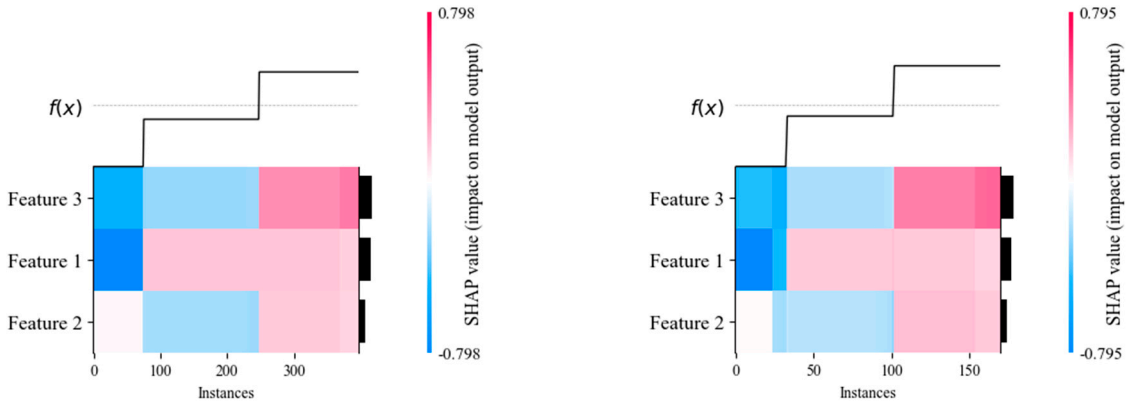


Figure 10. Waterfall plot of extracted sample features in training and testing sets.

3.3. Comparative Analysis of Different Algorithms

The XGBoost, KNN, RBFN, and Bayesian classifier algorithms were selected for comparison to analyse the prediction accuracy of different algorithms applied to coal and gas outburst sample data. The results are shown in Figure 11. The XGBoost algorithm achieved a training set accuracy of 93.47%, corresponding to Figure 11a, and a testing set accuracy of 82.46%, corresponding to Figure 11b. The KNN algorithm achieved a training set accuracy of 100%, corresponding to Figure 11c, and a testing set accuracy of 84.21%, corresponding to Figure 11d. The RBFN algorithm achieved a training set accuracy of 100%, corresponding to

Figure 11e, and a testing set accuracy of 85.96%, corresponding to Figure 11f. The Bayesian classifier algorithm achieved a training set accuracy of 100%, corresponding to Figure 11g, and a testing set accuracy of 83.04%, corresponding to Figure 11h.

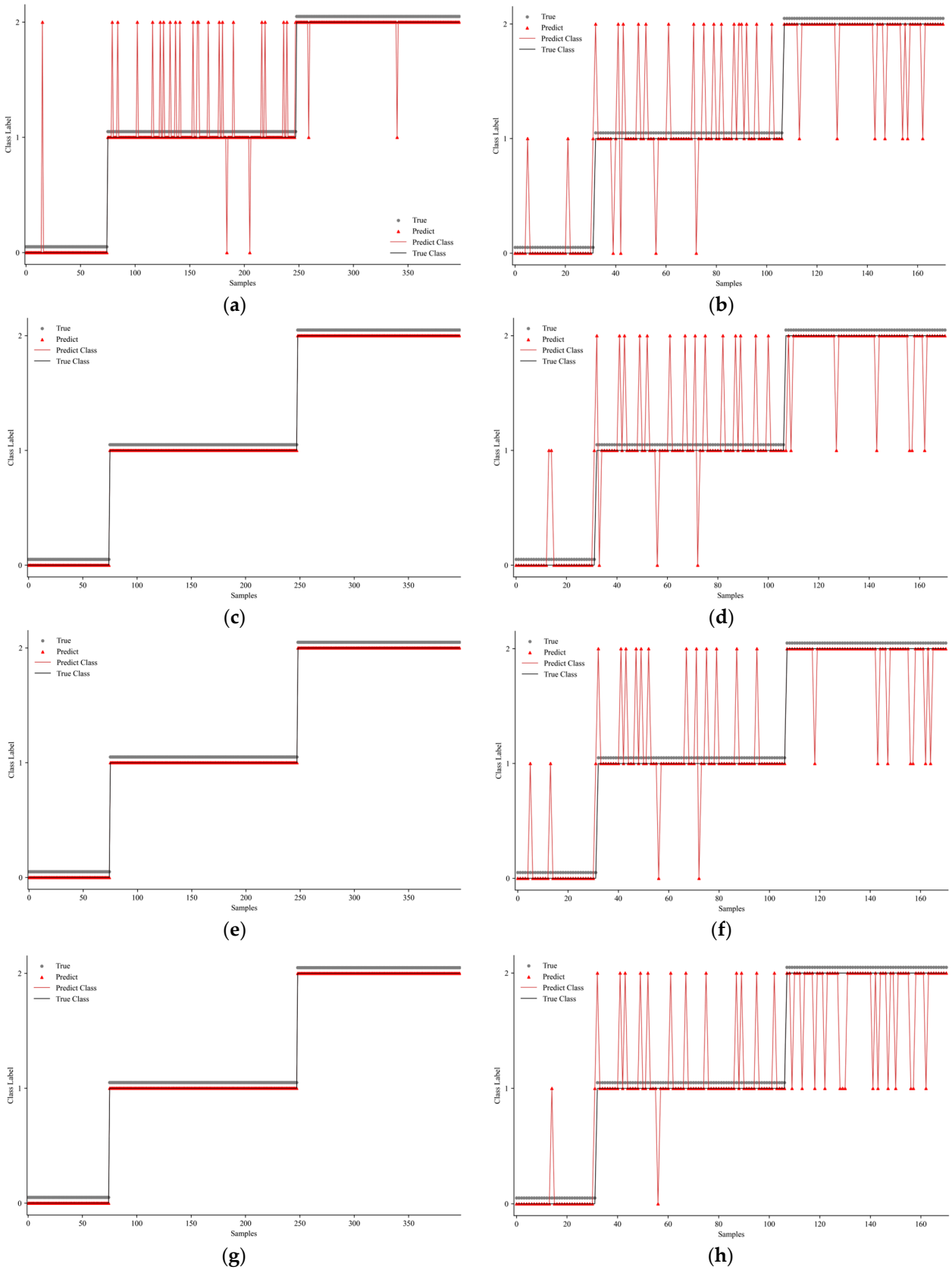


Figure 11. Prediction results of different comparison algorithms.

Figure 12 illustrates the ROC curves of different comparison algorithms. The overall AUC of the XGBoost algorithm was 0.94, with the AUCs of the three classes being 0.99, 0.89, and 0.94, represented in cyan, orange, and blue, respectively, as shown in Figure 12a. The KNN algorithm achieved an overall AUC of 0.93, with class-specific AUCs of 0.98, 0.88, and 0.95, as shown in Figure 12b. The RBFN algorithm achieved an overall AUC of 0.93, with class-specific AUCs of 0.97, 0.88, and 0.93, as shown in Figure 12c. The Bayesian classifier achieved an overall AUC of 0.93, with class-specific AUCs of 0.98, 0.89, and 0.91, as shown in Figure 12d.

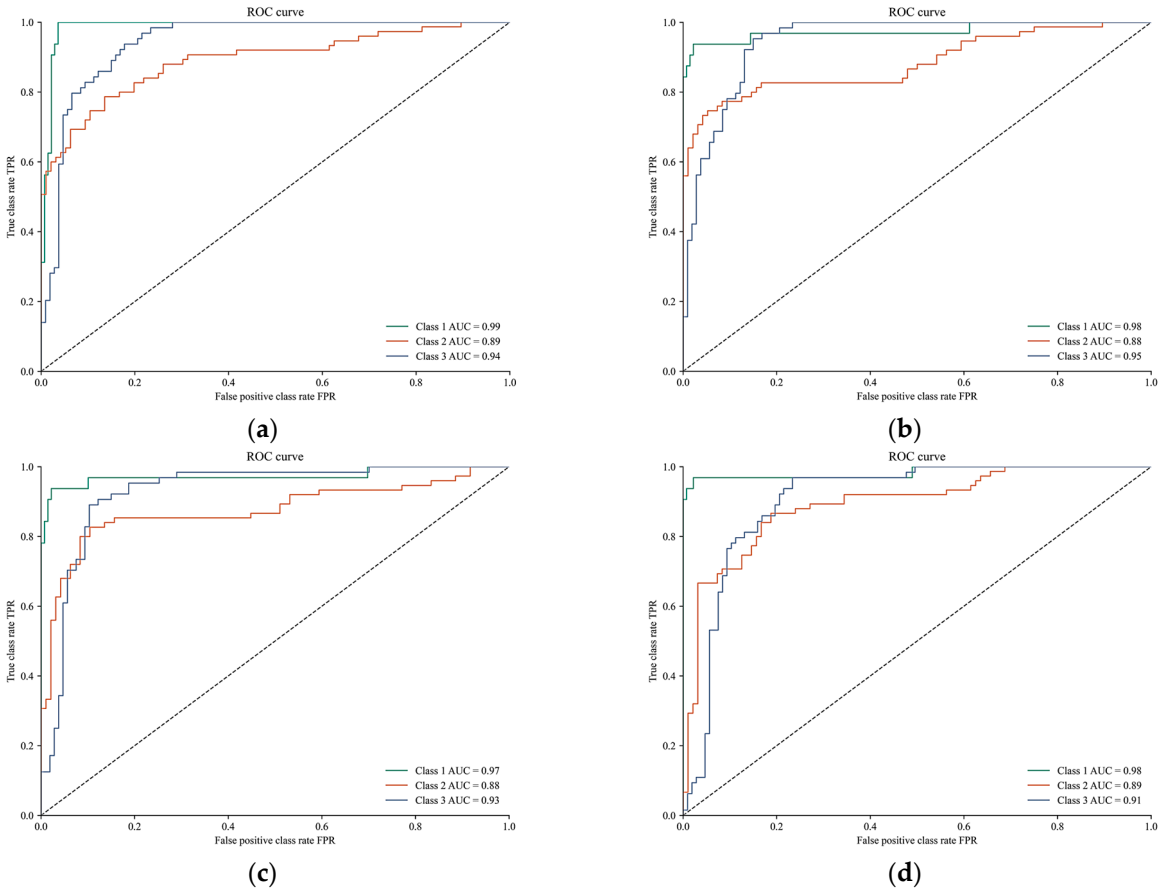


Figure 12. ROC curves of different comparison algorithms.

The comparative analysis of the four algorithms demonstrates that the RBFN algorithm achieved the highest prediction accuracy, followed by the KNN, Bayesian classifier, and, finally, the XGBoost algorithm. Although all four algorithms achieved 100% accuracy on the training set, none reached 100% accuracy on the testing set, further validating the superiority of the AFT-Transformer-SVM model.

Figure 13 presents the PAM (Polygon Area Metric) results of the four algorithms, corresponding to Figures 11 and 12, to analyse the sensitivity of the prediction accuracy of the different models. The corresponding data are summarised in Table 4. In Table 4, the comparative analysis of various data features reveals that all four algorithms performed poorly on the second and third classes of features while achieving better results on the first class of features. This finding indicates that the nonlinear characteristics of the data have not been fully extracted, and uncertainties remain in the data features.

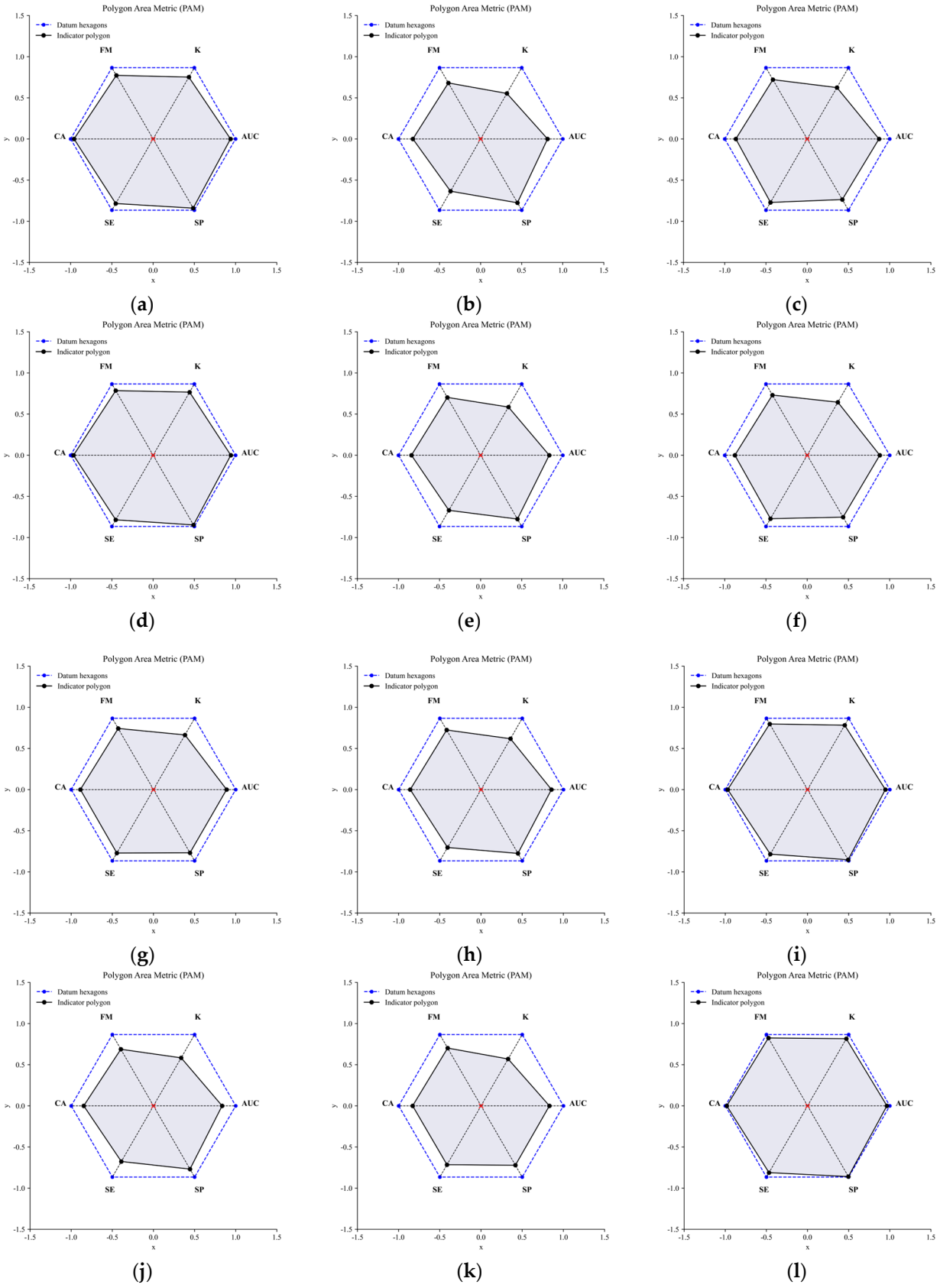


Figure 13. PAM results of different comparison algorithms.

Table 4. PAM values of different algorithms.

Algorithm	FM/%	κ/%	AUC/%	SP/%	SE/%	CA/%
XGBoost	89.23	86.70	93.87	97.12	90.62	95.91
	78.57	63.85	81.46	89.58	73.33	82.46
	83.21	72.07	87.05	85.05	89.06	86.55
KNN	90.62	88.47	94.23	97.84	90.62	96.49
	81.12	67.61	83.46	89.58	77.33	84.21
	84.44	74.34	87.99	86.92	89.06	87.72
RBFN	92.06	90.27	94.59	98.56	90.62	97.08
	83.56	71.33	85.46	89.58	81.33	85.96
	85.71	76.64	88.92	88.79	89.06	88.89
Bayesian classifiers	95.24	94.16	96.52	99.28	93.75	98.25
	81.05	65.71	83.00	83.33	82.67	83.04
	79.37	67.33	83.46	88.79	78.12	84.80

In Table 4, the PAM values for the first, second, and third classes of features are listed from top to bottom. Referring to Figure 12, the blue outer ring represents the ideal PAM value. None of the four algorithms achieved the optimal PAM value across all three classes. Specifically, the XGBoost algorithm achieved a PAM of 85.07% for the first class, 61.02% for the second class, and 70.27% for the third class. The KNN algorithm achieved a PAM of 86.05% for the first class, 64.79% for the second class, and 72.38% for the third class. The RBFN algorithm achieved a PAM of 88.07% for the first class, 68.62% for the second class, and 74.54% for the third class. The Bayesian classifier achieved a PAM of 92.52% for the first class, 63.63% for the second class, and 64.44% for the third class.

Through a comparative analysis with multiple algorithms, the AFT-Transformer-SVM algorithm demonstrated superior capabilities in extracting multidimensional nonlinear data features and achieving high prediction accuracy for coal and gas outbursts. The model achieved a testing set accuracy of 100%.

3.4. Data Validation for Other Mining Areas

To enhance the credibility of the model, based on the existing model, 66 sets of gas outburst test data from different time periods in a coal mine in Shaanxi, China, were selected as validation data. A validation experiment was designed. The prediction results of the model are shown in Figures 14–17. Among them, Figure 14 shows the prediction results of the AFT-Transformer-SVM model, where two test cases show slightly overestimated predictions, while the remaining values are predicted accurately, achieving an overall prediction accuracy of 96.97%, which is relatively high. Figure 15 shows the confusion matrix of the AFT-Transformer-SVM model's prediction results. From the figure, it can be seen that the model's prediction errors occur in two instances: one outburst risk point predicted as Class 2 and another outburst risk point predicted as Class 3, while the remaining values are all correctly predicted.

Figures 16 and 17 show the PAM results and ROC curve of the prediction results, respectively. In Figure 16a, the polygonal area is 96.42%; in Figure 16b, the polygonal area is 91.96%; and in Figure 16c, the polygonal area is 95.65%. The specific data can be referenced in Table 5. Classification 1 in Table 5 corresponds to Figure 16a, classification 2 corresponds to Figure 16b, and classification 3 corresponds to Figure 16c. Figure 17 presents the ROC curve of the prediction results, where all three classification results are close to the upper-left corner, indicating that the prediction results are relatively accurate.

Through the analysis, it has been found that the AFT-Transformer-SVM model achieved a prediction accuracy of 96.97% in a mining area in Shaanxi, China, which demonstrates a high level of accuracy and verifies the reliability of the model.

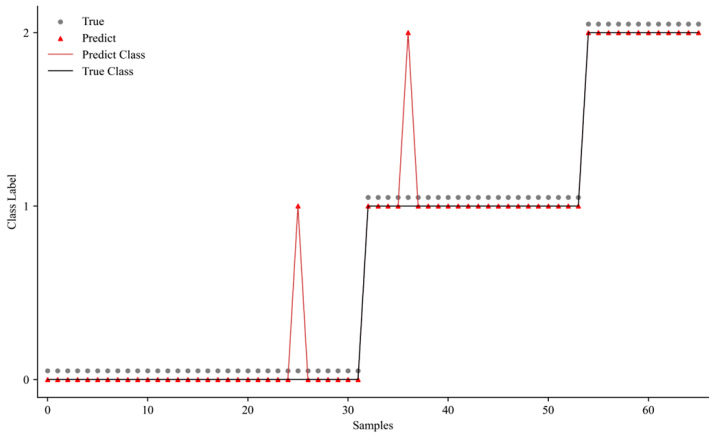


Figure 14. Prediction results of the AFT-Transformer-SVM model.

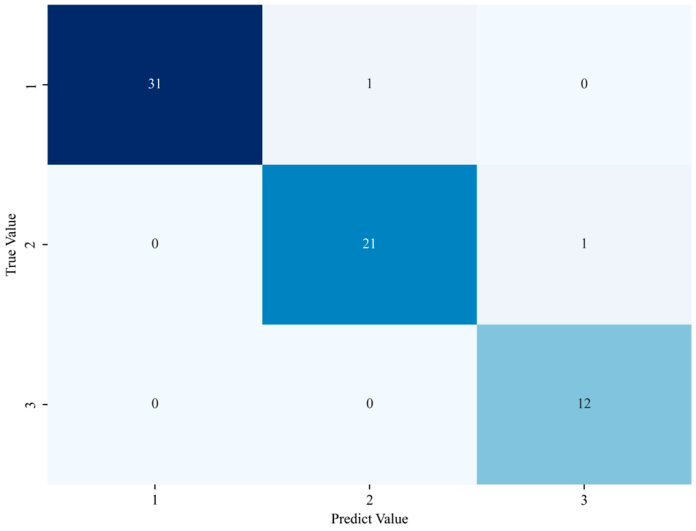


Figure 15. Confusion matrix of the prediction results for the AFT-Transformer-SVM model.

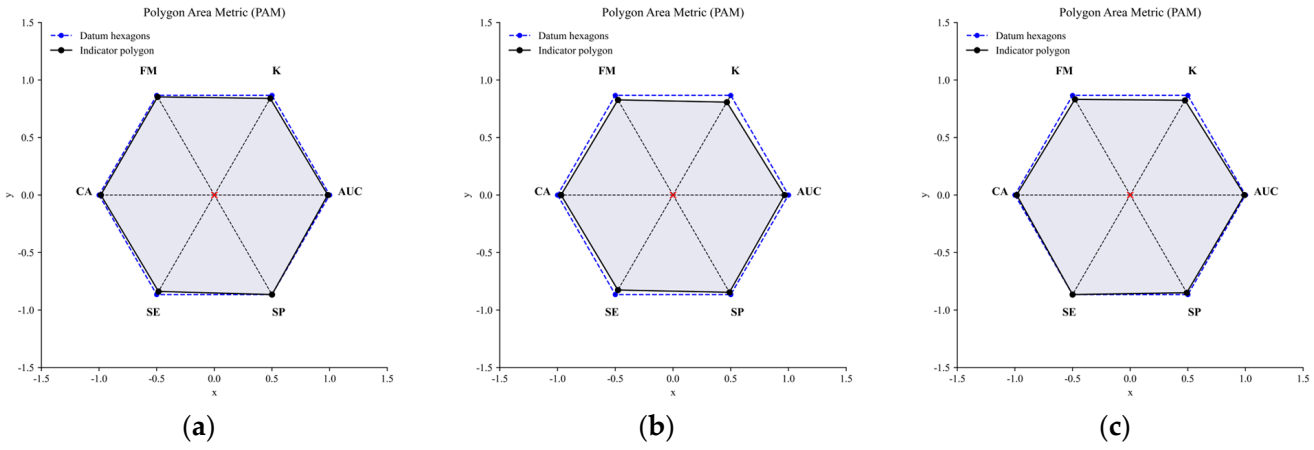


Figure 16. PAM results of AFT-Transformer-SVM predictions.

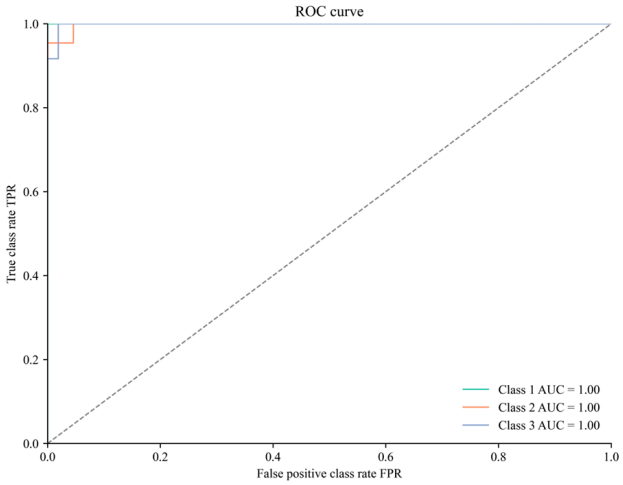


Figure 17. ROC curve of prediction results.

Table 5. PAM values of different algorithms.

Class	FM/%	κ/%	AUC/%	SP/%	SE/%	CA/%
1	98.41	96.96	98.44	100.00	96.88	98.48
2	95.45	93.18	96.59	97.73	95.45	96.97
3	96.00	95.07	99.07	98.15	100.00	98.48

3.5. Analysis of the Importance of Original Sample Data

PFI was employed to shuffle the original sample data and analyse the importance of the original features. Figure 18 illustrates the impact of different variables on the prediction results after shuffling. The variable with the greatest impact was gas content, followed by gas desorption, K1 gas release rate, drilling chip, distance to geological structure, burial depth, coal thickness, coal type, and strength coefficient. The box plots of the various parameters are marked in the figure. Table 6 provides the results of ten shuffles for each sample feature, listing the top five features with the highest importance. Among them, gas content had the highest feature importance, exhibiting the most significant impact on the prediction results, with the most apparent variations.

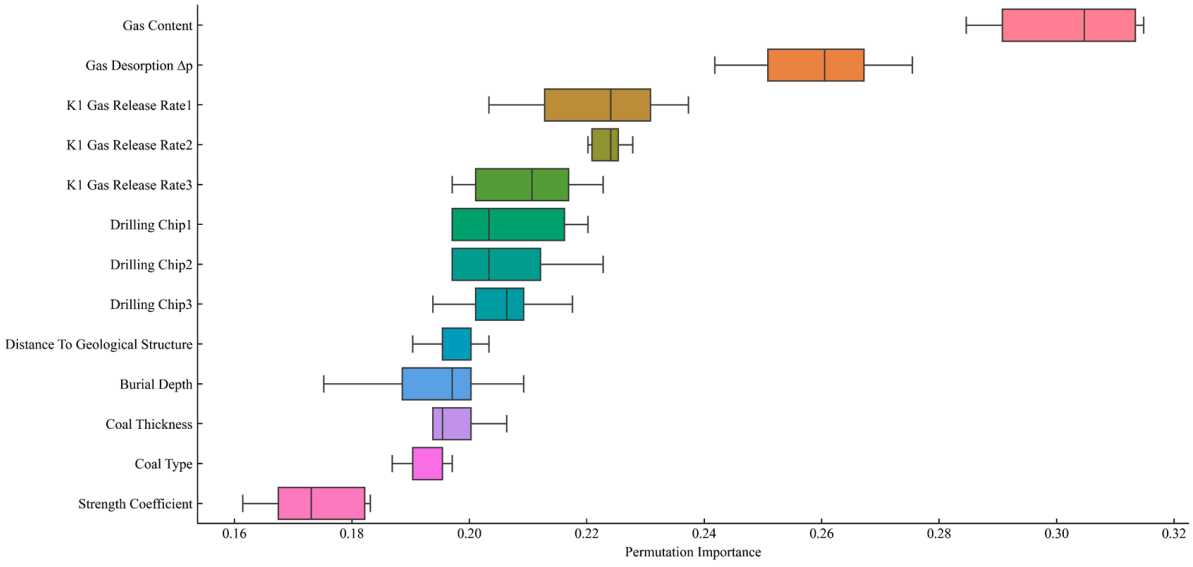


Figure 18. Box plots of feature importance based on PFI.

Table 6. Feature importance analysis after 10 shuffles.

Gas Content	0.20	0.39	0.40	0.28	0.30	0.29	0.30	0.30	0.31	0.29
Gas Desorption	0.26	0.26	0.24	0.24	0.26	0.25	0.26	0.24	0.27	0.25
Gas Release Rate1	0.23	0.21	0.22	0.22	0.23	0.23	0.23	0.20	0.21	0.22
Drilling Chip1	0.21	0.21	0.20	0.19	0.21	0.19	0.20	0.22	0.19	0.19
Distance to Geological Structure	0.20	0.19	0.19	0.20	0.20	0.20	0.20	0.20	0.20	0.20

Through this analysis, it was observed that in on-site sample testing data, the variable with the greatest impact on outburst risk was gas content, followed by gas desorption, K1 gas release rate, drilling chip, distance to geological structure, burial depth, coal thickness, coal type, and strength coefficient, which aligns with the characteristics of the real-world environment.

4. Conclusions

This study addresses the issue of gas outbursts in coal mine excavation faces by proposing a gas prediction and early warning method based on the interpretable AFT-Transformer-SVM model. A comprehensive framework for model construction, data analysis, and algorithm validation was developed. The findings of the study are as follows:

(1) This study innovatively integrates the AFT algorithm-optimised Transformer with the SVM model. By leveraging positional embedding, multi-head self-attention mechanisms, and fully connected layers, the model's ability to extract nonlinear and multidimensional data features of gas outbursts was significantly improved. The application of intelligent optimisation algorithms enabled the efficient parameter adjustment, ensuring the model's stability and accuracy in real-world operational environments.

(2) Based on SHAP and PFI analysis methods, the study conducted an in-depth examination of the impact of various features on model prediction, identifying the critical role of factors such as gas content, gas desorption rate, and drilling chip volume in assessing gas outburst risks. By incorporating multiple classification methods, including TP, TN, FP, and FN, the interpretability of the model was enhanced, providing a technical foundation for explaining gas prediction models in coal mines.

(3) A comparative analysis with traditional algorithms, such as XGBoost, KNN, RBFN, and Bayesian classifiers, demonstrated the significant advantages of the AFT-Transformer-SVM model in handling multidimensional nonlinear data and achieving high prediction accuracy. Experimental results indicated that the model achieved 100% prediction accuracy and excellent sensitivity on the test set, outperforming the comparison algorithms.

In conclusion, an effective method for predicting gas outburst risks in coal mines is provided by this study, which improves both the model accuracy and interpretability while offering valuable technical support for safe production in actual mining operations. On this basis, the model is embedded into the intelligent mine analysis platform, where it is periodically trained and used for the real-time prediction of coal and gas outburst hazards, integrated with on-site data. Future research will further combine theoretical analysis with field data to enrich real-time monitoring indicators, optimise monitoring methods, and advance the development of intelligent coal mine safety management.

Author Contributions: Conceptualisation, Y.W. (Yanping Wang) and J.D.; methodology, Z.Q.; software, Z.Q.; validation, Z.Q., Y.H. and Y.W. (Yiyang Wang); formal analysis, Y.C.; investigation, L.Z.; resources, Y.W. (Yanping Wang); data curation, Z.Q.; writing—original draft preparation, Y.W. (Yanping Wang); writing—review and editing, Y.W. (Yanping Wang); visualisation, Y.H.; supervision, J.D.; project

administration, Z.Y.; funding acquisition, J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Research and Development Program of Shaanxi Province, grant numbers 2020GY-139 and 2022GY-150.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data are not publicly available due to commercial confidentiality, as they contain information that could compromise the privacy of the research participants.

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

References

1. Liang, Y.; Zheng, M.; Li, Q.; Mao, S.; Li, X.; Li, J.; Zhou, J. A review on prediction and early warning methods of coal and gas outburst. *J. China Coal Soc.* **2023**, *48*, 2976–2994. [CrossRef]
2. Zhang, C.; Wang, P.; Wang, E.; Xu, J.; Li, Z.; Liu, X.; Peng, S. Coal and gas outburst mechanism: Research progress and prospect inChina over the past 70 years. *Coal Geol. Explor.* **2023**, *51*, 59–94.
3. Black, D.J. Review of Coal and Gas Outburst in Australian Underground Coal Mines. *Int. J. Min. Sci. Technol.* **2019**, *29*, 815–824. [CrossRef]
4. Wang, E.; Zhang, G.; Zhang, C.; Li, Z. Research Progress and Prospect on Theory and Technology for Coal and Gas Outburst Control and Protection in China. *J. China Coal Soc.* **2022**, *47*, 297–322.
5. Zhu, Z.; Zhang, Y.; Wang, S.; Zhao, B. Effect of Plasma-Activated Water on the Settling Characteristics of Ultrafine Kaolinite. *Process Saf. Environ. Prot.* **2024**, *192*, 613–620. [CrossRef]
6. Nakajima, I.; Ujihira, M. *Considerations for Acoustic Emission and Gas Emission in Gas Outburst Processes*; VRMA: Morgantown, WV, USA, 1989.
7. Feyt, G.N. Selecting Methods for Preventing Sudden Outbursts Based on Comprehensive Criteria of the Loss of Stability and Avalanche Destruction of the Seam. *Nauchn. Soobshch.-Inst. Gorn. Dela A. A. Skochinskogo (USSR)* **1980**, *186*, 39–46.
8. Hanes, J.; Lama, R.D.; Shepherd, J. Research into the Phenomenon of Outbursts of Coal and Gas in Some Australian Collieries. In Proceedings of the 5th ISRM Congress, Melbourne, Australia, 10 April 1983.
9. Sato, K.; Fujii, Y. Source Mechanism of a Large Scale Gas Outburst at Sunagawa Coal Mine in Japan. *Pure Appl. Geophys.* **1989**, *129*, 325–343. [CrossRef]
10. Shepherd, J.; Rixon, L.K.; Griffiths, L. Outbursts and Geological Structures in Coal Mines: A Review. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1981**, *18*, 267–283. [CrossRef]
11. Ji, P.; Shi, S.; Shi, X. Research on Early Warning of Coal and Gas Outburst Based on HPO-BiLSTM. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 2529808. [CrossRef]
12. Zhu, J.; Zheng, H.; Yang, L.; Li, S.; Sun, L.; Geng, J. Evaluation of Deep Coal and Gas Outburst Based on RS-GA-BP. *Nat. Hazards* **2023**, *115*, 2531–2551. [CrossRef]
13. Xie, X.; Shu, X.; Fu, G.; Shen, S.; Jia, Q.; Hu, J.; Wu, Z. Accident Causes Data-Driven Coal and Gas Outburst Accidents Prevention: Application of Data Mining and Machine Learning in Accident Path Mining and Accident Case-Based Deduction. *Process Saf. Environ. Prot.* **2022**, *162*, 891–913. [CrossRef]
14. Wang, W.; Wang, H.; Zhang, B.; Wang, S.; Xing, W. Coal and Gas Outburst Prediction Model Based on Extension Theory and Its Application. *Process Saf. Environ. Prot.* **2021**, *154*, 329–337. [CrossRef]
15. Hanson, D.R.; Lawson, H.E. Using Machine Learning to Evaluate Coal Geochemical Data with Respect to Dynamic Failures. *Minerals* **2023**, *13*, 808. [CrossRef] [PubMed]
16. Zheng, X.; Lai, W.; Zhang, L.; Xue, S. Quantitative Evaluation of the Indexes Contribution to Coal and Gas Outburst Prediction Based on Machine Learning. *Fuel* **2023**, *338*, 127389. [CrossRef]
17. Li, Z.; Wang, E.; Ou, J.; Liu, Z. Hazard Evaluation of Coal and Gas Outbursts in a Coal-Mine Roadway Based on Logistic Regression Model. *Int. J. Rock Mech. Min. Sci.* **2015**, *80*, 185–195. [CrossRef]
18. Frid, V.; Vozoff, K. Electromagnetic Radiation Induced by Mining Rock Failure. *Int. J. Coal Geol.* **2005**, *64*, 57–65. [CrossRef]
19. Mathews, J.P.; Campbell, Q.P.; Xu, H.; Halleck, P. A Review of the Application of X-Ray Computed Tomography to the Study of Coal. *Fuel* **2017**, *209*, 10–24. [CrossRef]
20. Parsa, A.B.; Movahedi, A.; Taghipour, H.; Derrible, S.; Mohammadian, A. Toward Safer Highways, Application of XGBoost and SHAP for Real-Time Accident Detection and Feature Analysis. *Accid. Anal. Prev.* **2020**, *136*, 105405. [CrossRef] [PubMed]
21. Lerner, B.; Guterman, H.; Aladjem, M.; Dinstein, I.; Romem, Y. Feature Extraction by Neural Network Nonlinear Mapping for Pattern Classification. In Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria, 25–29 August 1996; Volume 4, pp. 320–324.

22. Sun, L. Application Resrarch on the Coal and Gas Outburst Prediction Based on Gray Correlation Analysis and PSO-SVM. Master's Thesis, China University of Mining and Technology, Beijing, China, 2019.
23. Braik, M.; Ryalat, M.H.; Al-Zoubi, H. A Novel Meta-Heuristic Algorithm for Solving Numerical Optimization Problems: Ali Baba and the Forty Thieves. *Neural Comput. Appl.* **2022**, *34*, 409–455. [CrossRef]
24. Sokolova, M.; Japkowicz, N.; Szpakowicz, S. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In *AI 2006: Advances in Artificial Intelligence*; Sattar, A., Kang, B., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2006; Volume 4304, pp. 1015–1021. ISBN 978-3-540-49787-5.
25. Zhu, W.; Zeng, N.F.; Wang, N. Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS. In Proceedings of the NESUG Proceedings: Health Care and Life Sciences, Baltimore, MD, USA, 14–17 November 2010.
26. Tharwat, A. Classification Assessment Methods. *Appl. Comput. Inform.* **2021**, *17*, 168–192. [CrossRef]

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