The Evaluation of the Effects of Regulatory Variables on the Gravity of Coal Mine Disasters: A Machine Learning and SHAP Analysis Based Method

Mengxiu Li[1,2], Xihua Zhou[1,2], Hongyang Li[3,4]
[1]College of Safety Science & Engineering, Liaoning Technical University, Huludao 125105, Liaoning, China
[2]Laboratory of Key Laboratory of Mine Thermodynamic Disaster and Control of Ministry Education, Liaoning Technical University, Huludao 125105, Liaoning, China
[3]China Coal Technology and Engineering Group Shenyang Research Institute, Fushun 113122, Liaoning, China
[4]State Key Laboratory of Coal Mine Disaster Prevention and Control, Fushun 113122, Liaoning, China

Corresponding author: Mengxiu Li (e-mail: lmx13944350385@163.com)

ABSTRACT Accurate assessment of accident severity and regulatory deficiencies is essential for improving safety management and intelligent decision-making in complex industrial systems. This study proposes an interpretable machine learning framework to evaluate the influence of regulatory variables on coal mine accident severity by integrating data-driven prediction with SHAP-based feature attribution. Based on grounded-theory analysis of 382 accident investigation reports, more than 800 textual statements were condensed into 62 initial concepts, 13 secondary indicators, and five categories of regulatory factors, including design schemes, management systems, technical documents, organizational measures, and process methods. These variables were encoded and incorporated into Logistic Regression, C4.5, CART, CHAID, and Random Forest models for comparative analysis. Experimental results demonstrate that the Random Forest model achieves the best predictive performance in terms of accuracy and AUC, while SHAP analysis provides quantitative interpretation of the contribution and interaction of regulatory variables. The findings indicate that organizational measures, technical documentation, process methods, and management systems are the dominant determinants of accident severity, enabling transparent risk assessment and targeted intervention strategies. The proposed framework establishes an effective methodology for interpretable predictive analytics, intelligent safety monitoring, and data-driven decision support in complex engineering environments, offering valuable references for distributed sensing systems, industrial information fusion, and intelligent monitoring architectures related to Electromagnetic Waves, Antennas and Propagation engineering applications.

INDEX TERMS interpretable machine learning; SHAP analysis; accident severity prediction; intelligent safety monitoring; industrial information fusion

I. INTRODUCTION

COAL remains a critical component of the global energy system and plays a substantial role in economic development. However, coal mining involves inherent safety risks, and recurrent mine accidents have caused severe human and environmental losses. Against this background, this study examines how coal mine safety regulations influence accident prevention and accident severity [1]. Safety management literature highlights the centrality of rules and regulations in the prevention of accidents and reduction of their effects [2]. Nevertheless, the academic discussion has failed to differentiate between non-compliance attributed to poor practice and non-compliance due to poor regulatory systems [3]. Newer research articles like L. Wang et al have pointed to significant impacts of regulatory systems, safety settings, and investments on the safety of coal mines operations[4]. Y. Zhang underlined the fact that the so-called unsafe conditions of rules and regulations are the most important indicators influencing the so-called unsafe behavior of operators as well as the so-called unsafe conditions of equipment and the environment [5]. Based on these findings, Shi et al found unsafe behaviors, equipment failure and weak management systems as the leading causes of accidents [6]. In contrast, Laurence presented a counterargument that excessively prescriptive mining regulations could impede coming up with proper safety and health regulations and that they should focus on performance-based regulations based on overall duty of care [7]. Comparative researches, including Q. Liu, have demonstrated that the unified prescriptive system of coal mine safety control in China might be less efficient than the more flexible system in Australia[8]. Hopkins, despite this, noted the complementary aspect of adherence to the rules and the management of risks, stating that none of them should be ignored in order to prevent serious accidents in an organization [9]. Tan et al and Wu et al. suggested creating a coal mine safety supervision system that was in accordance with guidelines, policies and technical standards to establish a strong framework that would constrain corporate behaviour [10, 11]. Although these contributions have been made, the literature has not yet provided a complete view of the role that rules and regulations play in determining the different factors that lead to coal mine accidents.

Given the severity of coal mine accidents, there is an urgent need to develop effective approaches for predicting accident severity under different regulatory conditions. Conventional evaluation methods like Analytic Hierarchy Process and Fuzzy Comprehensive Evaluation have certain limitations when it comes to measuring complex relationships and forecasting the probability of accidents [12, 13]. The further development of artificial intelligence has seen the increased use of machine learning algorithms like neural networks, random forests, Bayesian networks, and support vectors machines in the prediction of coal mine accident risks [14-17]. As an example, Xu et al applied machine learning to measure ground subsidence due to coal mining and to predict the relationship between maximum subsidence and its determinants [18]. In the same way, Miao et al explored the indicators of risk prediction of coal mine rock bursts and compared the predictive capability of various algorithms at different risk levels [19]. Also, Bi et al provided a hybrid machine learning solution to determine the variables that influence coal mine gas emissions [20]. Taken together, the research presented in these studies shows the usefulness of machine learning models in the prediction of coal mine accident risks. Nevertheless, there are not many research works that have used the SHAP algorithm in interpretable machinelearning models to explain the risk factors of coal mine accidents. The value of the SHAP approach has been illustrated by research conducted by Ibrahim et al, Qiu and Zhou, and Xiaoliang to explain risk factors related to coal mine accidents [21-23]. To address this gap, this paper will build a theoretical framework to systematically examine the position of regulatory frameworks on coal mine accidents and combine machine learning approaches with the SHAP algorithm to enhance the predictive power and interpretability of accident risk estimation. The paper offers the fresh insight and solutions to the issue of prevention and control of coal mine accidents. Besides, it adds to the sustainable development of the coal sector by forecasting the impact of regulatory frameworks on the frequency and level of such accidents. It is essential in enhancing coal mine safety management and decreasing the risk of accidents.

II. MATERIAL AND METHODS

The present research paper takes a systematic methodological approach in studying the connection between the severity of coal mine accidents and regulatory factors. Figure 1 shows how the whole analytical framework is structured, starting with data preparation and containing important steps like data transformation, feature selection, model building, and model interpretation. The first step involved the systematic gathering of 382 case records of Chinese coal mine accidents that were affected by regulatory factors during the period 2012-2022. The raw data was coded and conceptualized through grounded theory, which identified the regulatory characteristics that impact the severity of coal mine accidents and served as the basis of further analysis. Then, the categorical variables were converted into Boolean expressions to make them applicable to future modeling. This study also used filter-based feature selection algorithms to determine variables that are closely related to accident severity and remove those with low association. Then, various classification models were developed based on the machine learning methods, such as logistic regression, C4.5 decision trees, CART decision trees, CHAID decision trees, and random forests. Not only do these models determine the factors that contribute to the severity of the accidents but they also offer scientific means of prediction and risk reduction in the field of coal mine safety management. Lastly, SHAP method was used to examine the interpretability of the model further. It was possible to quantify the particular role played by every feature in making the predictions of the model based on the Shapley value of each feature, which contributed to the increased transparency and credibility of the model.

A. DATA PREPARATION

1) Selection of Indicator Characteristics

Since the academic society does not completely comprehend what constitutes the regulatory frameworks in the coal mining industry, exploratory qualitative research is necessary. Grounded theory is distinguished by a dynamic way of collecting data compared to other qualitative research methodologies. The study uses an exploratory approach relying on the grounded theory to examine the raw data. To begin with, the qualitative data contained in the raw data is evaluated systematically, summarized, conceptualized and classified. Aftewards, the information is continuously compared and reduced and relationships between different concepts and categories are drawn, which eventually results into the creation of the theoretical framework of factors that influence them [24]. According to grounded theory, the qualitative analysis proceeded through open coding, axial coding, and selective coding [25]. The current research focuses on Chinese coal mine accidents associated with regu-

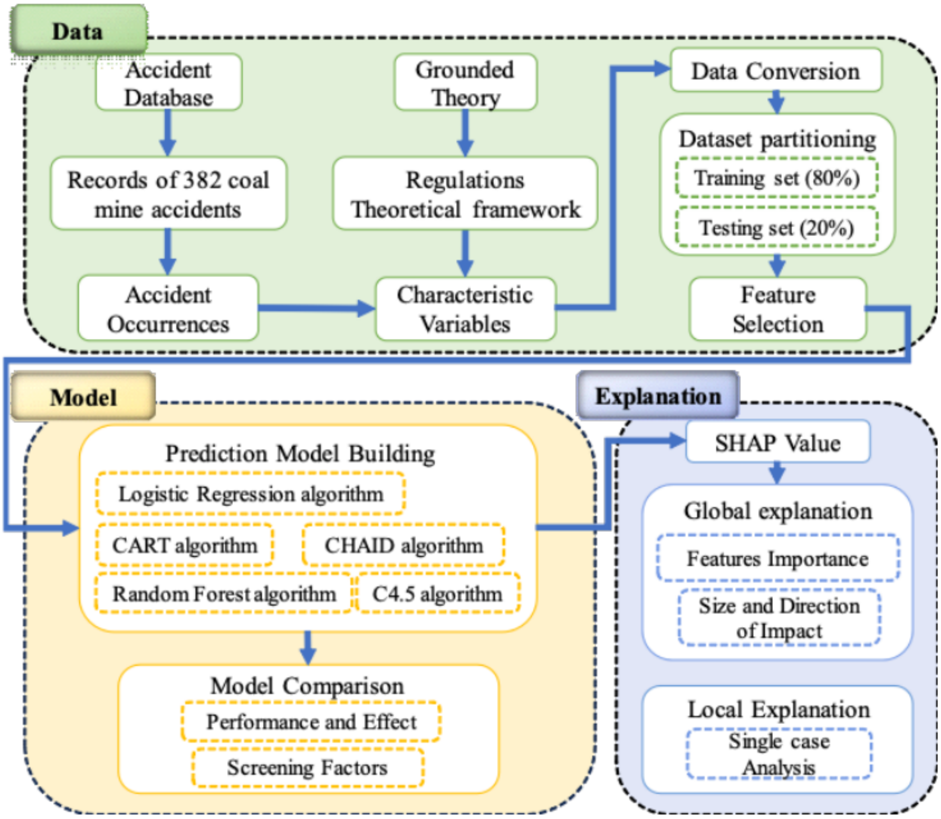

FIGURE 1. Analytical framework of this study

latory deficiencies during 2012-2022. Accident severity and fatalities were analyzed using annual investigation reports issued by national and provincial mine safety supervision agencies. Through open coding, more than 800 textual statements and their associated concepts were initially identified. After screening and synthesis, these statements were condensed into 62 initial concepts and 13 secondary indicators. During selective coding, the 13 secondary indicators were further integrated into five core categories: design schemes, management systems, technical documents, organizational measures, and process methods. In the quantitative stage, each secondary indicator was operationalized as a binary variable: a value of 1 was assigned only when the accident report explicitly described the corresponding deficiency, and 0 otherwise. To improve reproducibility, the coding rules were defined on the basis of report-based evidence rather than subjective wording alone. The revised Table 1 provides the operational definitions of the secondary indicators.

TABLE 1. Description of Variables

Variable type	Category	Variable name	Operational coding rule	Range of values
Implicit variable	–	Whether or not major casualties	–	0–1
Independent variable	Accident occurrence	Years	–	2012–2022
		Season	–	1–4
		Regions	–	1–11
	Design scheme	No formal design scheme	Use only when the accident report explicitly states that no relevant formal design scheme or design document was prepared for the operation or hazard-control activity.	0–1
		Untimely or unreasonable design updates	Use when a design existed but was not updated in time to reflect actual site conditions or when its technical arrangement was unreasonable for the observed risk.	0–1
		Deficient design schemes	Use when a design scheme existed but omitted key safety content, lacked necessary parameters/calculations, or was inconsistent with actual working conditions.	0–1

TABLE 1. Description of Variables (continued)

Variable type	Category	Variable name	Operational coding rule	Range of values
Independent variable	Management system	Absence of management system	Use when no formal management system or institutional arrangement existed for the relevant safety activity.	0–1
		Unreasonable management system design	Use when a system existed but its provisions were incomplete, contradictory, or unsuitable for the task.	0–1
		Dysfunctional implementation of management systems	Use when a formal system existed but failed in implementation or was ineffective in practice.	0–1
	Technical document	No technical document support	Use when relevant technical documents were absent.	0–1
		Incomplete technical documents	Use when documents existed but lacked essential content, procedures, or parameters.	0–1
		Inaccurate technical documents	Use when documents contained incorrect, outdated, or site-inconsistent information.	0–1
	Organizational measure	Insufficient organizational measures	Use when measures were missing or inadequate in scope, staffing, supervision, training, coordination, or emergency preparation.	0–1
		Ineffective organizational measures	Use when measures formally existed but did not function effectively in practice.	0–1
	Process method	Inappropriate process methods	Use when the selected operating process was technically unsuitable for the relevant hazard conditions.	0–1
		Unreasonable or illegal process methods	Use when the operating method violated required procedures, safety standards, or legal requirements.	0–1

B. DATA ANALYSIS

The research has analyzed 382 cases of coal mine accidents that took place during the period of 2012-2022 which included the time distribution of the events, regional features, and the role of regulatory factors. As shown in Figure 2, the results indicate that there has been a great decline in the number of accidents with large-scale casualties in the past few years. It can be seen that the number of major accidents (more than 10 deceased) fell each year throughout the period of the study. This tendency is probably closely connected with the toughening of national work safety regulations and the constant improvement of coal mine safety technologies.

In accordance with the classification by regions of the National Bureau of Statistics of China, the 31 provinces, autonomous regions and municipalities of China are categorized into four regions including eastern, central, western and northeast China.

According to Figure 3, eastern China has the lowest occurrence of coal mine accidents, possibly because of its stronger economic base and more mature safety management capacity. The central region shows the lowest average fatality level. By contrast, northeast China exhibits a higher average number of fatalities. This regional pattern may reflect not only differences in regulatory implementation but also region-specific mining conditions, such as deeper seams, more complex geological settings, gas and water hazards, mine age, and resource depletion. Because such physical variables were not explicitly modeled in this study, the regional result should be interpreted as a descriptive association rather than as evidence of purely regulatory gaps [26, 27].

Coal mine accidents are associated with cyclical variables, including seasonality. Figure 4 shows how coal mine accidents vary across different seasons. The increased market demand and concentration of production tasks in the third quarter lead to more frequent mining activity, which may elevate accident risk. Nevertheless, the average death rate per accident is rather low with a large number of accidents in the third quarter. It appears that when there is a high level of production activity, coal mining companies could have tighter safety monitoring policies to achieve safe production. Also, third quarter accidents usually have less casualties, which could be due to the strong enforcement of the safety control and supervision measures in season.

According to Figure 5, accidents due to technical documentation causes are the most common, however the average number of fatalities is much smaller compared to those caused by other causes. The result implies that inaccuracies or lack of information in technical documentation could be the main reason behind operational mistakes. Workers will not be able to have the correct standards to refer to so they are prone to making errors that might cause accidents. By contrast, despite the fact that accidents due to design-scheme factors are rare, the casualties are usually the most severe ones when they do happen. It probably is so because design scheme errors may be the fundamental cause of the operational process overall safety. Such accidents usually have systemic origins that result in massive deaths and injuries. This research indicates the need to focus on the precision of technical documentation and the quality of design schemes in coal mine safety management.

C. MODELING METHODS

1) Machine Learning Algorithm Comparison

The research paper involves the comparative multi-algorithm analysis in order to forecast the casualties of coal mining accidents due to regulatory factors. The algorithms that were used in the given analysis are logistic regression, three kinds of decision tree algorithms and random forest. The logistic regression is a statistics technique that has been applied to predict probabilities in binary classification issues. Even though the word regression is part of its name, logistic regression can be viewed as a classification algorithm. This model employs a logistic function, namely, the sigmoid function to transform the result of linear regression into

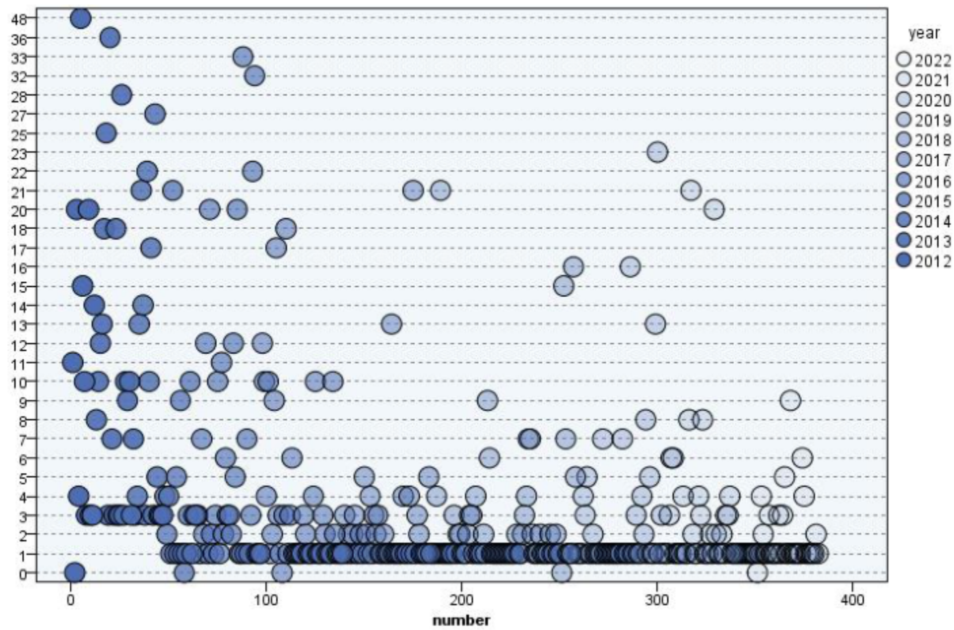

FIGURE 2. Trends in coal mine accidents, 2012-2022

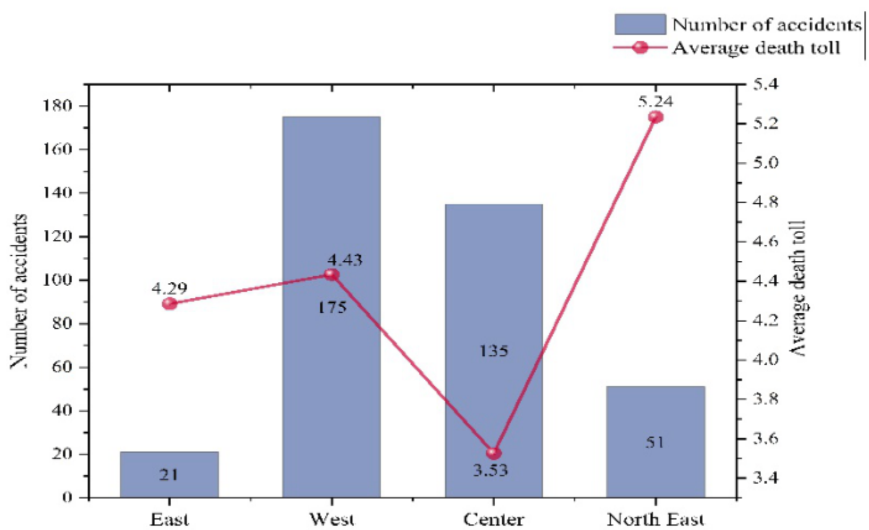

FIGURE 3. Coal mine accidents by region

the range of [0, 1] which estimates the probability of a specific event happening [28]. Decision tree algorithms belong to the category of machine-learning methods that are applied to classification and regression problems. C4.5 is an improved version of the ID3 algorithm and is mostly used in classification problems. CART (Classification and Regression Trees) algorithm can be applied to classification as well as regression problems. It measures node impurity through the Gini impurity index and the most appropriate attribute in terms of data segmentation is chosen, which leads to an intuitive tree structure. The CHAID (Chi-squared Automatic Interaction Detector) algorithm relies on the chi-squared test to determine the independence between attributes, thus it is especially useful in classification tasks. CHAID enables multivariate splits using several variables per node and produces interpretable tree models. Random Forest is a type of the ensemble learning algorithm, which builds several decision trees and averages their results to enhance the quality and strength of the predictive model. In the course of training, every tree is based on a random sample of the data set, and randomness is used in choosing the attributes to prevent overfitting [29].

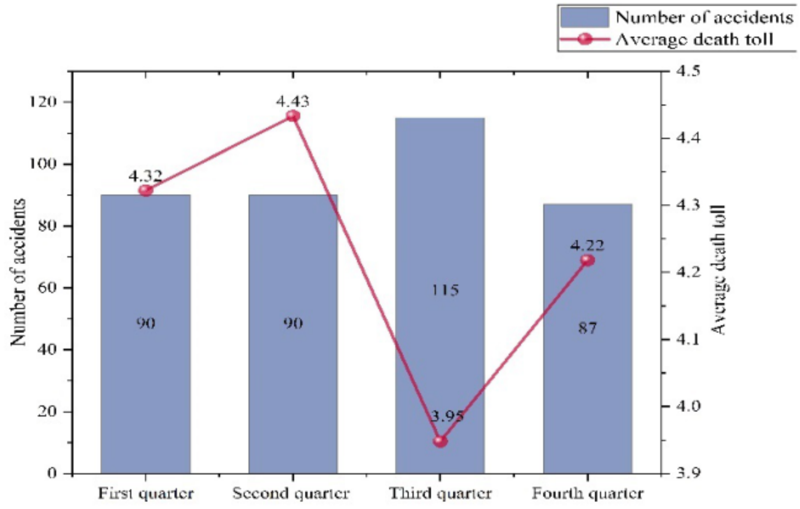

FIGURE 4. Coal mine accidents by quarter

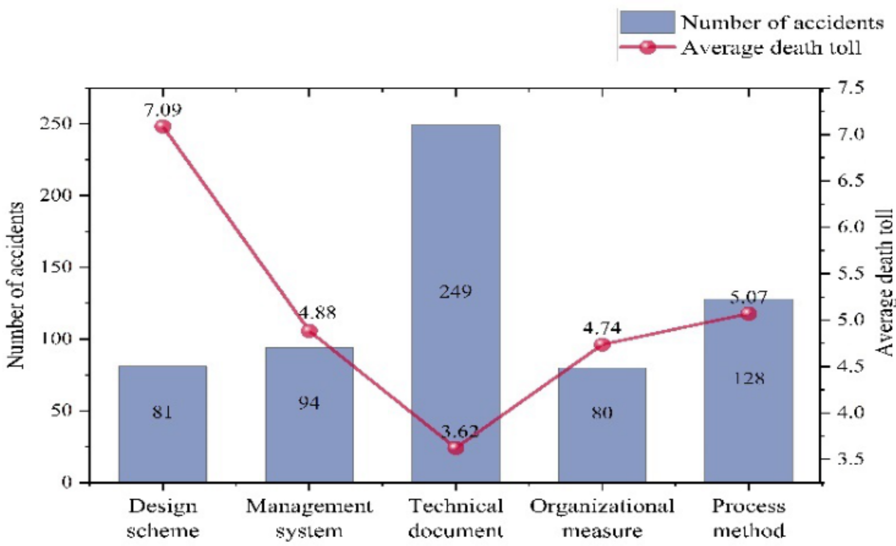

FIGURE 5. Coal mine accidents by regulatory factors

2) Categorical Evaluation Metrics

In order to test the generalization performance of a classification model, we depend on a sequence of successful experimental estimation techniques and appropriate evaluation measures. Accuracy, precision, recall and F1 score are major performance metrics used in classification tasks. The metrics are calculated using the confusion matrix of a binary classification problem, as presented in Table 2.

TABLE 2. Binary Confusion Matrix

		Predicted Values	
		Negative (0)	Positive (1)
Actual Values	Negative (0)	TN	FP
	Positive (1)	FN	TP

Note: True Positives (TP): Observations predicted as positive and actually positive.
True Negatives (TN): Observations predicted as negative and actually negative.
False Positives (FP): Negative observations are incorrectly predicted as positive.
False Negatives (FN): Positive observations are incorrectly predicted as negative.

The most obvious performance measure is accuracy, which reflects the fraction of samples that were correctly categorized by the model in the entire set of samples. It can be derived by using the confusion matrix.

$$\text{accuracy} = \frac{TN + TP}{FP + TN + TP + FN} \tag{1}$$

Precision (also called positive predictive value) is the ratio between the number of predictions labeled as positive and the number that were actually positive and measures the extent to which the model can predict the positive class. By comparison, recall is the percentage of all true positive examples that are correctly labeled as positive and measures the capacity of the model to classify the positive class. The respective formulas are:

$$\text{precision} = \frac{TP}{FP + TP} \tag{2}$$

$$\text{recall} = \frac{TP}{FN + TP} \tag{3}$$

The F1 score is the harmonic average of the precision and recall and can be considered a holistic measure of performance because it incorporates the two metrics. The F1 score is weighted towards precision and recall, however, one of them has a stronger effect on the result than the other. The formula is given below:

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{4}$$

The proportions of correct and incorrect predictions per class can also be deduced based on the confusion matrix, with four possible results: The true positive rate indicates the proportion of positive samples that are correctly predicted.

$$TPR = \frac{TP}{TP + FN} \tag{5}$$

The false negative rate is the fraction of positive samples that are erroneously classified as negative.

$$FNR = \frac{FN}{TP + FN} \tag{6}$$

False positive rate is the percentage of negative samples that are incorrectly labeled as positive.

$$FPR = \frac{FP}{FP + TN} \tag{7}$$

The true negative rate tells us what percentage of the negative samples will be correctly predicted as negative.

$$TNR = \frac{TN}{FP + TN} \tag{8}$$

These metrics offer an integrated structure to assess and compare the results of various classification models. The application of the measures will help researchers to develop a better insight into how a model is likely to perform on a given dataset, hence allowing them to make the right choices in selecting and optimizing models.

3) SHAP Interpretability Analysis

SHAP (Shapley Additive exPlanations) is a widely used explainability technique in machine learning introduced by Lundberg et al. [30]. The method is based on Shapley values from game theory and is used to quantify the contribution of each feature to the model output. By computing Shapley values, SHAP improves the transparency and interpretability of the model [31]. Mathematically, the Shapley value ϕ_i is calculated as follows:

$$\phi_i(v) = \sum_{\substack{s \subseteq \{x_1,\ldots,x_p\} \\ s \not\ni x_j}} \frac{|s|!(p - |s| - 1)!}{p!} \times (v(s \cup \{x_j\}) - v(s)) \tag{9}$$

Here, denotes the Shapley value of the feature and it indicates how much this feature contributes to the prediction of the model, is the collection of all features in the model, is the overall number of features, stands as the feature vector of the sample that needs to be explained, and is the model output based on the subset of features. Weightage guarantees the equality of the Shapley values so that every combination of features sets has the same importance in the computation.

In addition, assuming that is the model explanation, is the number of variables, and indicates the presence or absence feature (values of 0 or 1), we can express the predicted value of the model as:

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z'_i \tag{10}$$

In this case, denotes the Shapley value of each attribute, where indicates the average prediction of all cases. By doing so, the SHAP approach does not only measure the effect of separate attributes, but also combines the efforts of all the attributes to provide an overall picture of the model prediction. Our research paper focuses on a full-scale analysis of how important characteristics impact accident casualty prediction models based on the SHAP method. We find out the main factors that have a significant influence on the prediction outcomes by computing the SHAP values of the training samples. The analysis offers a solid ground on which the accident casualty prediction can be enhanced.

D. MODEL CONSTRUCTION

Prior to the empirical analysis, this paper made the appropriate encoding transformations of some of the variables which are categorical variables and turned them into binary numerical variables to enable modeling. In particular, accident severity was divided into two categories depending on the number of fatalities: those with three or more fatalities were classified as positive samples (coded as 1) and those with less than three fatalities were defined as negative samples (coded as 0). Figure 6 shows that 39.53 percent of the accident cases have caused at least three deaths and 60.47 percent has caused less than three deaths. This binary classification system does not just make it easier to construct the models but also enhance their predictive quality.

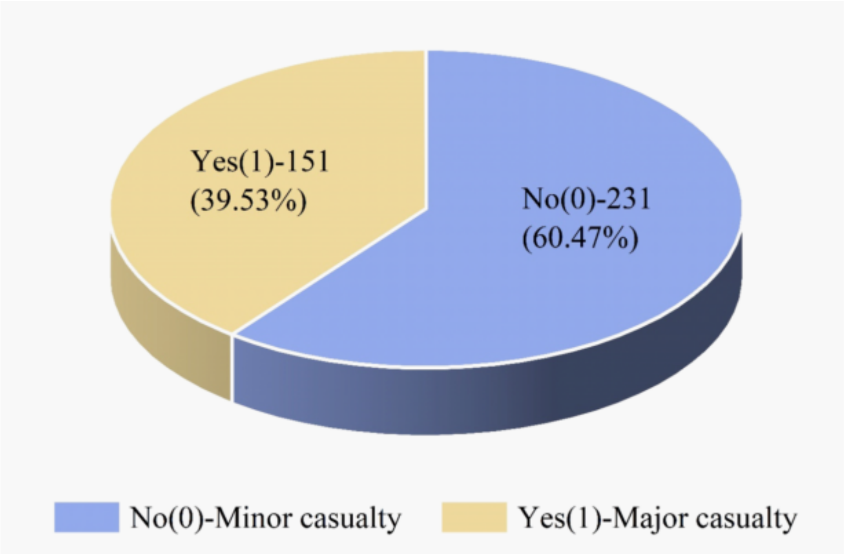

FIGURE 6. Dependent Variable Data Categorization

Moreover, in case of mistakes associated with the regulation, this research applied binary code and coded accidents with these errors as 1 and no errors as 0. Having performed the data preprocessing described above, this paper undertook a detailed statistical analysis of 382 coal mine accidents affected by regulatory factors and formulated a classification model. The model is capable of accurately calculating the casualty results of coal mine accidents caused by regulatory problems and thus offers an opportunity to provide a scientific foundation to the managers when coming up with the mitigation strategies of risks. The training set was used to train the classification model and the test set was used to evaluate its performance. The dataset was divided strictly, with 80 percent of the observations used for training and the remaining 20 percent reserved for testing. Although the binary outcome distribution was moderate, severe cases were concentrated more heavily in earlier years. Therefore, year of occurrence was retained as a contextual temporal variable to capture historical heterogeneity across the study window rather than as a direct regulatory factor. Due to the large dimensionality of the accident data, this paper employed a filter-based feature selection algorithm (Liu and Yu, 2005) to select and preserve features strongly related to coal mine accidents. The importance-related coefficient was set at 0.667 to eliminate features with weaker correlations, thereby simplifying model complexity and improving predictive accuracy [32]. The present study built machine learning models with five algorithms which are logistic regression, C4.5 decision tree, CART decision tree, CHAID decision tree and random forest based on the chosen features. The reason behind building these models is to investigate the complex nonlinear relationship in the data by means of various algorithmic structures and optimizing predictive performance. In order to enhance model accuracy and stability even more, this study considered adding boosting and bagging methods to the CART and CHAID algorithms. This method will be used to remove features of zero importance and to decrease the noise in the models and enhance their capability to determine the important features.

III. RESULTS AND DISCUSSION

A. MODEL COMPARISON

When researching the categorization of the severity of coal mine accidents, this paper has used Logistic Regression, C4.5 Decision Tree, CART Decision Tree, CHAID Decision Tree, and Random Forest models to act as classifiers. The number of casualties on coal mine accidents was used as the main classification indicator and the severity of accidents was classified into large and small, which is a binary classification issue. In order to perform an exhaustive evaluation of various classifiers, this paper performed a complete evaluation in two aspects, namely, model performance and factor screening, based on quantitative classification metrics and qualitative factor analysis.

1) Comparison of modeling effects

Initially, based on model performance, this paper quantifies the comparative performance of the five mentioned models in the task of classifying the severity level of coal mine accidents. The performance of each model was measured using a set of important classification measures, which consisted of accuracy, precision, recall, F1-score, geometric mean (G-mean) and area under the ROC curve (AUC). All these measures fully represent the balance and accuracy of every model to manage positive and negative cases.

Of these models, random forest model (RF) demonstrated the best performance with an accuracy of 0.7895 compared to the C4.5 and CHAID models with the worst accuracy of 0.7263. It means that random forest model is the most precise one to predict the severity of the coal mine accident. F1 score, as the harmonic mean of precision and recall,

TABLE 3. Comparison of Model Effectiveness

Models	Accuracy	Precision	Recall (TPR)	TNR	FPR	FNR	F1-Score	G-Mean	AUC
Logistics	0.7474	0.9412	0.4103	0.9821	0.0179	0.5897	0.5714	0.1624	0.7270
C4.5	0.7263	0.7407	0.5128	0.8750	0.1250	0.4872	0.6061	0.2013	0.7610
CART	0.7368	0.7333	0.5641	0.8571	0.1429	0.4359	0.6377	0.2338	0.7770
CHAID	0.7263	0.8824	0.3846	0.9643	0.0357	0.6154	0.5357	0.1376	0.8130
RT	0.7895	0.7568	0.7179	0.8393	0.1786	0.3077	0.7368	0.3631	0.8310

represents the overall balance of these two measures in the model. The highest F1 score was obtained by the random forest model, at 0.7368, which means that this model has a better tradeoff between precision and recall compared to other models. Also, the logistic regression model showed the greatest precision, 0.9412, which means that it is very accurate in predicting large-scale accidents. The random forest model has the highest recall, 0.7179, which indicates that it can be used to identify large-scale accidents and minimize the probability of missing detection. Also, the random forest model recorded the highest Gmean, which is especially useful when working with imbalanced datasets and indicates good performance even when the positive and negative examples are not evenly distributed.

The given table shows the particular values of the classification performance measures of each model. Of these models, the random forest model (RF) was the most accurate, with an accuracy of 0.7895, and the C4.5 and CHAID models were the least accurate, with an accuracy of 0.7263. This suggests that the random forest model is the most effective when it comes to the prediction of the severity of coal mine accidents. The F1 score being the harmonic mean of both precision and recall also indicates how balanced these two metrics are in the model. The random forest model gave the highest value of F1 score, i.e., 0.7368 which means that the balance between the precision and recall is greater than in any of the other models. Moreover, the logistic regression model was most precise with 0.9412, which means that the model is highly accurate in forecasting large-scale accidents. The random forest model recorded the highest recall of 0.7179 implying that it can be used effectively both to identify large-scale accidents and minimize the possibility of missing detections. Besides, the random forest model performed the highest G-mean, especially in the case of unbalanced datasets, and it performs significantly well with unequal distribution of positive and negative samples.

The ROC curve is drawn, depending on the relationship between the true positive rate (TPR) and the false positive rate (FPR), as illustrated in Figure 7. Every point on the curve corresponds to the model performance at a given classification threshold. The AUC is a measure of the average performance of the model over the entire range of classification thresholds, so an increased AUC value indicates a better overall performance of the model. Figure 7 demonstrates that the ROC curve of the random forest model is the nearest to the upperleft corner and is higher than the rest of the models in most areas. It also has the largest AUC value, which is 0.8310, which confirms once again the best average performance of the random forest model over the whole range of classification thresholds. According to the given analysis, the random forest model has better performance in predicting the severity level of a coal mine accident particularly in the case of unbalanced data. This benefit might arise due to the combination of many decision trees in the random forest model, which is likely to minimize overfitting, but still have good robustness to outliers and noise. Hence, along with the available literature (Rhodes et al., 2023), and the findings of this paper, the random forest model can be considered as an excellent option to explore the determinants of coal mine safety regulation and accident fatalities due to its high precision and overall highperformance level [33]. Because severe accidents were concentrated in earlier years, these metrics should be interpreted as performance within the 2012-2022 study window rather than as time-invariant evidence of present-day risk prediction.

2) Comparison of factor screening results

Despite the fact that the random forest model has proven to be very efficient in terms of classification in this research, every model has some weaknesses and is unable to consider all the influencing factors. The aim of this study is thus to find out which of the identified key factors of the different models have a significant impact on the severity of the coal mine accident classification through an extensive comparison of these factors, which will help give scientific background to the process of preventing accidents and managing safety. This paper ranked the factors that were influential as determined by each algorithm based on their level of importance and provided a summary of these factors in Table 4 as a result of the comparative analysis of the screening findings across various models.

As shown in Table 4, various models demonstrate some consistency in factor selection. All models ranked year of occurrence highly; however, in this study year is interpreted as a temporal contextual variable rather than as a direct regulatory indicator. It likely absorbs broader historical changes across 2012-2022, including improvements in regulation, mechanization, monitoring technology, rescue capability, and the industrial environment. Therefore, the substantive regulatory interpretation of the model focuses on the direct regulatory variables rather than on year itself. Among the direct regulatory variables, the effectiveness of organizational measures, the completeness and accuracy of technical documents, the legality and rationality of process methods, and the presence of management systems remain the most important findings. Besides the year variable, there are other variables that have been found to be significant in different models, including the scientific effectiveness of measures, completeness of document content, legality of the process methods, presence of corresponding management systems. The following factors are the most important in the outcome of accidents casualties. Of special note is that the factor of

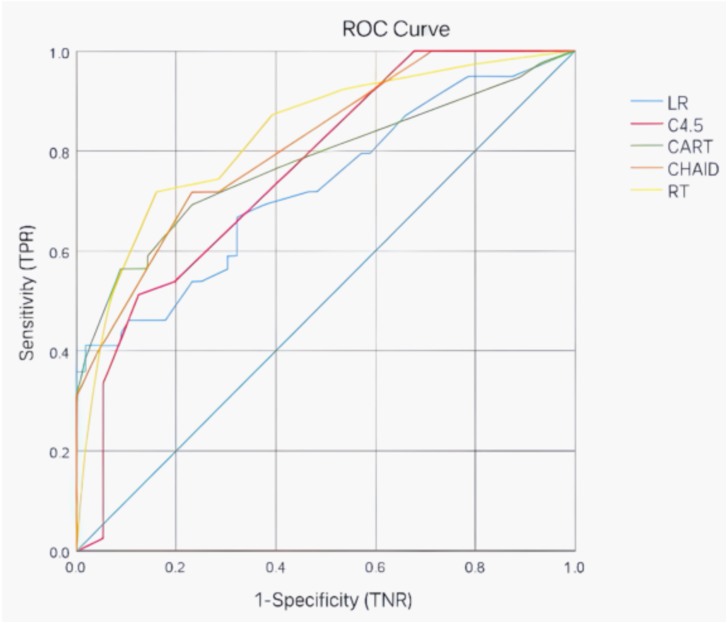

FIGURE 7. Comparison of ROC curves of different models

scientific effectiveness of measures is relevant in all models, highlighting the need to implement scientifically viable prevention and control measures prior to an accident, as well as efficient emergency measures during an accident that can minimize the number of casualties. The significance of both the concepts of complete technical documents and rational and legal process methods in all the models is that they play a positive role by providing a clear technical guidance and legal procedure to prevent accidents and minimize their effects. Also, the focus on the concept of the corresponding management systems in different models means that the creation and optimization of management systems are critical to the prevention and response to accidents[34]. To sum up, by conducting an extensive examination of the screening outcomes of different models, this paper has determined the main determinants of coal mine accident severity and highlighted the importance of these factors in accident prevention and safety management. These results do not only offer a scientific ground on the safety management of coal mines but also give insights into the further study and application. The key factors can be addressed more effectively with a better understanding of such factors which will contribute to reducing the number of accidents and casualties caused by them if the accident prevention and response capacities are significantly enhanced in coal mines.

B. SHAP VISUALIZATION AND ANALYSIS

In comparison to simple decision trees, more sophisticated machine learning algorithms like random forests are capable of providing better predictive results, but can also be devoid of any intuitive prediction procedure and cannot necessarily represent interactions between features. The limitation can decrease the transparency and interpretability of the prediction process of the model. In order to overcome this problem, SHAP value analysis offers a quantitative mechanism to attribute the contributions of each feature to the prediction result. Not only does it measure the influence of a single feature, but also explains how other features behave together to affect the predictions made by the model, thus improving the model credibility and its usefulness.

The global feature importance ranking based on the random forest model is shown in Figure 8. The high ranking of year of occurrence again indicates substantial temporal heterogeneity across the study period rather than suggesting that year itself constitutes a direct regulatory failure. Among the direct regulatory variables, technical documentation, process methods, design schemes, management systems, and organizational measures contributed most strongly to the prediction of severe casualties.

Figure 9 provides both the magnitude and direction of each predictor's effect on the model output. Positive SHAP values indicate feature states that increase the predicted probability of severe casualties, whereas negative SHAP values indicate states associated with lower predicted severity. For the direct regulatory variables, the direction of impact is also consistent with accident causation logic. Unreasonable or illegal process methods can weaken critical safety barriers by bypassing required procedures for ventilation, support, drainage, blasting, or hazard isolation, thereby increasing worker exposure when abnormal events occur. Incomplete or inaccurate technical documents can delay hazard recognition and reduce operational consistency because workers may lack reliable design parameters, operating instructions, or emergency guidance. Likewise, insufficient or ineffective organizational measures may reflect inadequate supervision, training, coordination, and emergency preparedness, all of

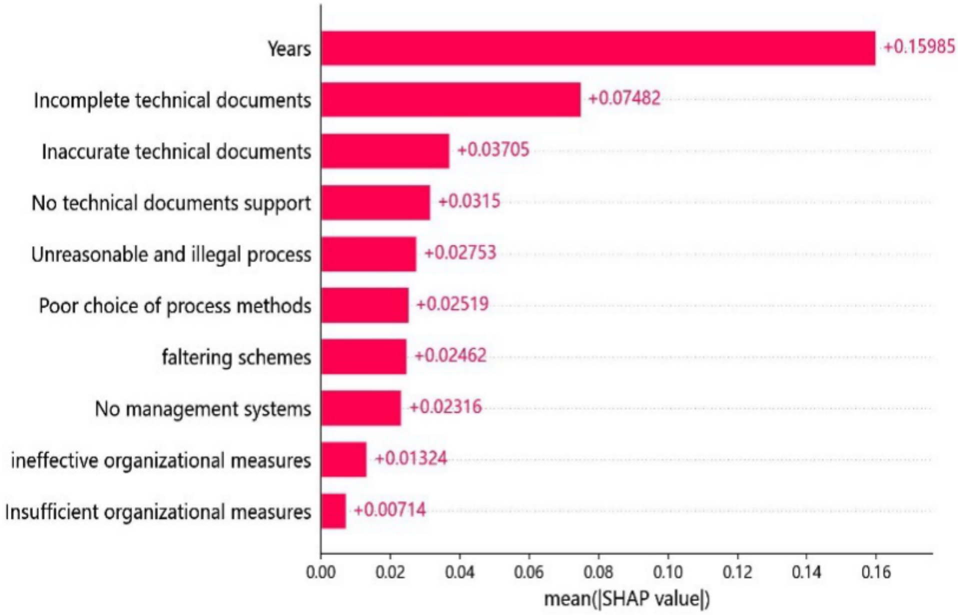

FIGURE 8. RT model predicts global feature importance for larger casualty impact features

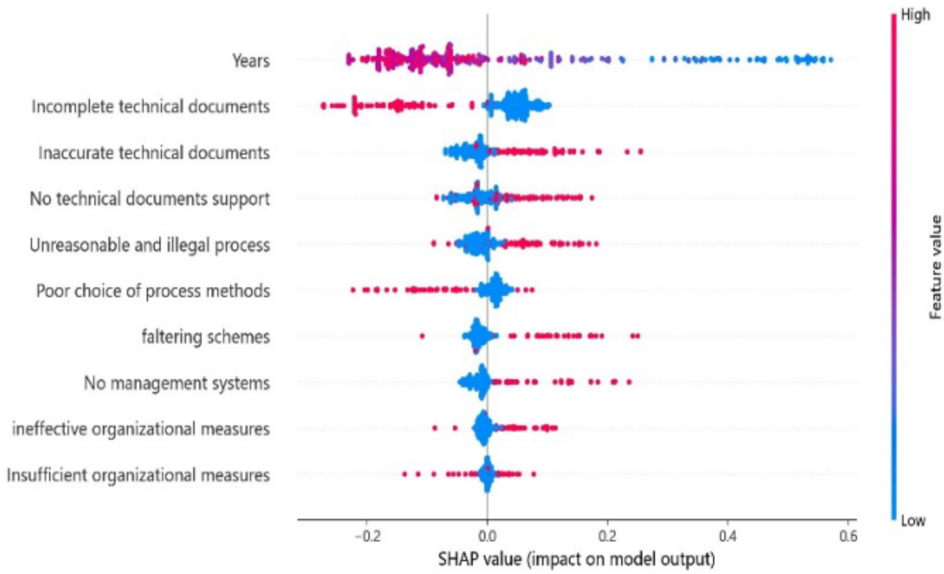

FIGURE 9. SHAP summary plot of RT model

which can enlarge the consequences of an accident. Accordingly, SHAP is used here not only to rank variables but also to interpret how specific regulatory deficiencies may escalate accident severity. The results obtained are very significant to the practical management of safety in coal mines. In case these factors are not present or are defective, they can cause more serious outcomes. On the other hand, in case these factors are not perfect or improperly applied, they may be associated with only trivial injuries or incidents. It highlights the significance of setting up and putting into practice rules and procedures. Although some regulations are not optimal, their presence will also help reduce the severity of accidents. Moreover, the constant enhancement of safety regulations is an essential step towards the reduction of accidents in coal mines [35].

Lastly, in order to improve the comprehension of individual case forecasts, the present paper describes the process of predicting of the model on the basis of case analysis of the 4. 26 massive water inrush incident at Xinsheng Coal Mine in Qiaojia Town, Yanhe County, Tongren City, Guizhou Province with SHAP feature contribution waterfall plot (Figure10). The analysis shows what factors were important to the occurrence of the accident and how they have impacted the prediction result. The predicted

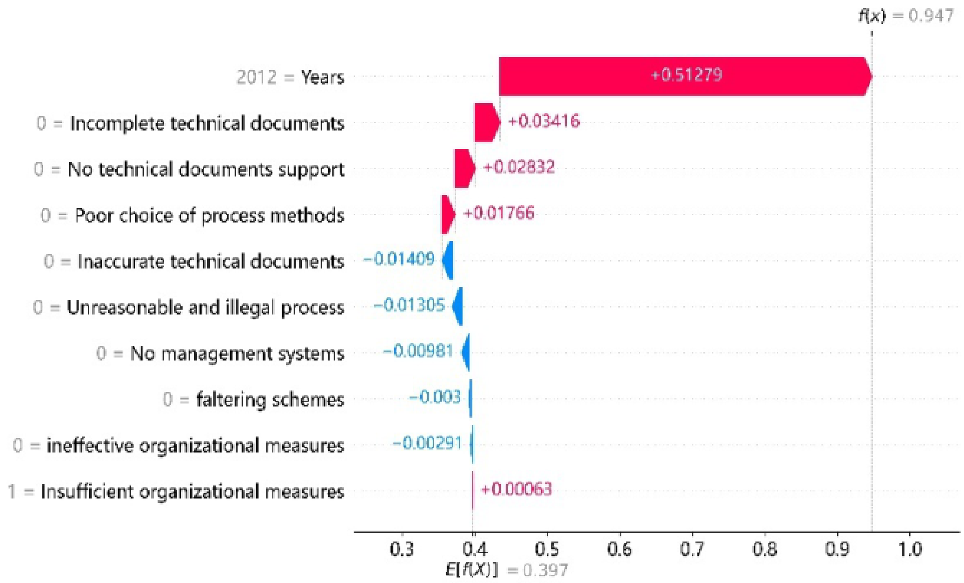

FIGURE 10. Example waterfall diagram of SHAP interpretation of the "4.26" accident

probability of the model is 0.947 which is much greater than the base value of the mean prediction of the training dataset (0.397) suggesting that the combined features in this catastrophe greatly raised the chances of serious casualties. Figure 10 further illustrates the impact of each feature on the prediction result in the selected '4.26' case. Positive SHAP contributions were mainly associated with early-year occurrence, inadequate technical documentation, lack of technical document support, improper process selection, and insufficient organizational measures. From an operational perspective, these factors indicate that the accident-control system lacked both upstream technical guidance and downstream organizational response capacity, thereby increasing the likelihood of severe casualties. The case-level explanation is consistent with the global SHAP results and supports the practical usefulness of the model for accident analysis. Given these findings, policymakers and practitioners should recognize that coal mine safety management is a continuous and adaptive process. Safety rules, technical documents, and process measures should be periodically reviewed and updated in response to changing production conditions and emerging hazards. Standardized procedures are needed to ensure that technical documents are regularly revised and that organizational and process measures remain effective under new operating conditions.

IV. CONCLUSION

The Random Forest model has shown better results in coal mine accident severity classification with a 0.7895 accuracy, 0.7368 F1 score, and 0.8310 AUC which is much higher than the rest of the models. This good result of these metrics suggests that the Random Forest model could be used to identify the underlying complex trends that determine the severity of coal mine accidents and thus can be used as a dependable predictive indicator in formulating preventive measures against accidents. The screening and analysis of different machine learning models helped the study to identify five major groups of influencing factors which are temporal trends, effectiveness of organizational measures, completeness of technical documentation, compliance with process methods and establishment of management systems. These factors interact to influence the prevention of accidents and the reduction of casualties to a significant extent, which is why it is crucial to consider various factors in coal mine safety management. This paper builds on the SHAP method to investigate how much each feature contributes in the model. SHAP value analysis does not only explain the exact contribution of each of the factors to predict the severity of accidents but also gives grounds to make priorities in the area of accident prevention. Moreover, the analysis highlights the significance of the presence, completeness, and accuracy of regulatory systems in avoiding severe outcomes. The paper is useful in contributing to the prevention and control of coal mine accidents. Nevertheless, there is space to conduct additional research. The future research might include exploring the integration of more appropriate indicators that would enhance and perfect the accident prediction index model regarding regulatory factors of coal mines. Other methods of machine learning could be introduced to boost predictive capabilities. Moreover, the combination of the main factors discussed in the current research with the factors of other areas, like the geological situation, and the severity of the accidents can bring to the more profound knowledge of the reasons behind coal mine accidents. Also, the approaches and models that have been developed in this study can be used in predicting and

TABLE 4. Comparison of factor screening results

Models	Logistics	C4.5	CART	CHAID	RT	Synthesize
Important variable	Years	Years	Years	Years	Years	Years
	Ineffective organizational measures	Ineffective organizational measures	Ineffective organizational measures	Incomplete technical documents	Incomplete technical documents	Ineffective organizational measures
	No management systems	Poor choice of process methods	Unreasonable and illegal process	No management systems	Inaccurate technical documents	Incomplete technical documents
	Unreasonable and illegal process	Incomplete technical documents	No management systems	Unreasonable and illegal process	No technical documents support	Unreasonable and illegal process
	No technical documents support	Faltering schemes	Unreasonable and illegal process	No management systems		
	Incomplete technical documents	Poor choice of process methods				
	No technical documents support	Faltering schemes				
	Insufficient organizational measures	No management systems				
	Poor choice of process methods	Ineffective organizational measures				
	Insufficient organizational measures					

managing safety problems in other industrial areas.

AUTHOR CONTRIBUTIONS

Mengxiu Li: Conceptualization, methodology, data curation, formal analysis, software, visualization, and writing - original draft. Xihua Zhou: Supervision, project administration, methodology, and writing - review and editing. Hongyang Li: Resources, validation, investigation, and writing - review and editing. All authors have read and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

REFERENCES

[1] B. Wu, J. Wang, M. Zhong, C. Xu, and B. Qu, "Multidimensional Coal Mine Safety Accidents in China - Review of 70 Years," Mining, Metallurgy and Exploration, vol. 40, no. 1, pp. 253-262, 2023, doi: 10.1007/s42461-022-00722-w.

[2] A. R. Hale and P. Swuste, "Procedural freedom or constraining action safety rules?" Safety Science, vol. 29, no. 3, pp. 163-177, 1998, doi: 10.1016/S0925-7535(98)00020-4.

[3] J. Reason, D. Parker, and R. Lawton, "Organizational controls and safety: The varieties of rule-related behaviour," Journal of Occupational and Organizational Psychology, vol. 71, no. 4, pp. 289-304, 1998, doi: 10.1111/j.2044-8325.1998.tb00678.x.

[4] L. Wang, Q. Cao, and L. Zhou, "An investigation into the factors that affect the safety of coal mine production combining the use of DEMATEL and ISM," Safety Science, vol. 103, pp. 51-61, 2018, doi: 10.1016/j.ssci.2017.11.007.

[5] Y. Zhang, W. Shao, M. Zhang, H. Li, S. Yin, and Y. Xu, "Structural equation modeling analysis of 320 coal mine accidents with unsafe conditions of the rules and regulations as exogenous variables," Accident Analysis and Prevention, vol. 92, pp. 189-201, 2016, doi: 10.1016/j.aap.2016.02.021.

[6] X. Shi, Y. Liu, K. Ma, Z. Gu, Y. Qiao, G. Ni, C. Ojum, A. Opoku, and Y. Liu, "Evaluation of risk factors affecting the safety of coal mine construction projects using an integrated DEMATEL-ISM approach," Engineering, Construction and Architectural Management, 2024, doi: 10.1108/ECAM-02-2023-0103.

[7] D. Laurence, "Safety rules and regulations in mine sites: An issue and an answer," Journal of Safety Research, vol. 36, no. 1, pp. 39-50, 2005, doi: 10.1016/j.jsr.2004.11.004.

[8] Q. Liu, X. Li, and M. Hassall, "Regulation system of coal mine safety in China and Australia: A comparative study and general results," Resources Policy, vol. 74, Art. no. 101454, 2021, doi: 10.1016/j.resourpol.2019.101454.

[9] A. Hopkins, "Risk-management and rule-compliance: Decision-making in hazardous industries," Safety Science, vol. 49, no. 2, pp. 110-120, 2011, doi: 10.1016/j.ssci.2010.07.014.

[10] B. Tan, Y. Song, and W. Shi, "The research on innovation mechanism of the coal mine safety production supervision," in Proceedings - 2014 IEEE Workshop on Advanced Research and Technology in Industry Applications, WARTIA 2014, 2014, pp. 524-527, doi: 10.1109/WARTIA.2014.6976312.

[11] B. Wu, J. Wang, B. Qu, and P. Qi, "Development, effectiveness, and deficiency of China's Coal Mine Safety Superivision System," Resources Policy, vol. 82, Art. no. 103524, 2023, doi: 10.1016/j.resourpol.2023.103524.

[12] M. Dagdeviren and I. Yuksel, "Developing a fuzzy analytic hierarchy process (AHP) model for behavior-based safety management," Information Sciences, vol. 178, no. 6, pp. 1717-1733, 2008, doi: 10.1016/j.ins.2007.10.016.

[13] X. Li, D. Chen, J. Fu, S. Liu, and X. Geng, "Construction and application of fuzzy comprehensive evaluation model of rockburst based on microseismic monitoring," Applied Sciences, vol. 13, no. 21, Art. no. 12013, 2023, doi: 10.3390/app132112013.

[14] S. Yang et al., "Prediction of mine water inflow by robust deep learning models with respect to the management of groundwater environment," Science of the Total Environment, vol. 871, 2023, doi: 10.1016/j.scitotenv.2023.162056.

[15] A. U. Khan, S. Salman, K. Muhammad, and M. Habib, "Modelling coal dust explosibility of Khyber Pakhtunkhwa coal with the help of random forest algorithm," Energies, vol. 15, no. 9, 2022, doi: 10.3390/en15093169.

[16] X. Xuecai, S. Xueming, F. Gui, S. Shifei, J. Qingsong, H. Jun, and W. Zhirong, "Data-driven accident causes based on coal and gas outburst accidents prevention: Implementation of data mining and machine learning in accident path mining and accident case-based deduction," Process Safety and Environmental Protection, vol. 162, pp. 891-913, 2022, doi: 10.1016/j.psep.2022.04.059.

[17] J. Zhang, Q. Wang, W. Guo, and L. Li, "A prediction using support vector machine of sensitivity to coal ash blast at different levels of deterioration," Journal of Sensors, 2022, doi: 10.1155/2022/7604338.

[18] C. Xu, K. Zhou, X. Xiong, F. Gao, and Y. Lu, “Predictions of subsidence due to mining using sparrow search algorithm with extreme gradient boosting and TOPSIS method,” Acta Geotechnica, vol. 18, no. 9, pp. 4993-5009, 2023, doi: 10.1007/s11440-023-01830-7.
[19] D. Miao, K. Yao, W. Wang, L. Liu, and X. Sui, “Prediction of risk of rock burst in a coal mine using machine learning and feature selection algorithm,” Georisk, 2024, doi: 10.1080/17499518.2024.2346653.
[20] S. Bi, L. Shao, Z. Qi, Y. Wang, and W. Lai, “Prediction of coal mine gas emission using hybrid machine learning model,” Earth Science Informatics, vol. 16, no. 1, pp. 501-513, 2023, doi: 10.1007/s12145-022-00894-5.
[21] B. Ibrahim, I. Ahenkorah, and A. Ewusi, “Explainable risk assessment of rockbolts failure in underground coal mines based on categorical gradient boosting and SHapley additive explanations,” Sustainability, vol. 14, no. 19, 2022, doi: 10.3390/su141911843.
[22] Y. Qiu and J. Zhou, “The short term rockburst prediction of underground projects: Lessons of an explainable and interpretable ensemble learning model,” Acta Geotechnica, vol. 18, no. 12, pp. 6655-6685, 2023, doi: 10.1007/s11440-023-01988-0.
[23] Z. Xiaoliang, L. Wenhao, Z. Lei, and X. Sheng, “Quantitative analysis of contribution of indexes in predicting coal and gas outburst using machine learning,” Fuel, vol. 338, 2023, doi: 10.1016/j.fuel.2023.127389.
[24] Y. Xiong, C. Zhang, H. Qi, R. Zhang, and Y. Zhang, “What is the safety cognition capability of urban citizens? A measurement model using the theoretical framework of cognitive progression theory,” Frontiers in Psychology, vol. 13, Art. no. 707172, 2022, doi: 10.3389/fpsyg.2022.707172.
[25] G. Goldkuhl and S. Cronholm, “Putting theoretical base into grounded theory: On the way to multi-grounded theory,” International Journal of Qualitative Methods, vol. 9, no. 2, pp. 187-205, 2010, doi: 10.1177/160940691000900205.
[26] X. Wang, C. Wang, Y. Huang, H. Zhang, and Q. Wang, “Paper on the correlation between work accidents and safety policies in China,” Processes, vol. 9, no. 5, 2021, doi: 10.3390/pr9050805.
[27] Y. Tian, X. Yang, J. Yang, K. Mao, Y. Yao, and H. Liang, “Dynamics of evolution of intelligent construction strategy of coal mine enterprises in China,” Heliyon, vol. 8, no. 10, 2022, doi: 10.1016/j.heliyon.2022.e10933.
[28] N. Srimaneekarn, A. Hayter, W. Liu, and C. Tantipoj, “Analysis of binary response by logistic regression in dentistry,” International Journal of Dentistry, 2022, doi: 10.1155/2022/5358602.
[29] M. Wang, G. Li, L. Dong, Z. Hou, J. Zhang, and D. Li, “Severity classification of Graves orbitopathy based on random forest algorithm,” Hormone and Metabolic Research, 2023, doi: 10.1055/a-2287-3734.
[30] S. M. Lundberg, P. G. Allen, and S.-I. Lee, “A unified approach to interpreting model predictions.”
[31] T. Mo, S. Li, and G. Li, “A readable machine learning model of predicting cavity water depth and cavity length using XGBoost SHAP,” Journal of Hydroinformatics, vol. 25, no. 4, pp. 1488-1500, 2023, doi: 10.2166/hydro.2023.050.
[32] H. Liu and L. Yu, “The way to integrate feature selection algorithms in classification and clustering,” IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 4, pp. 491-502, 2005, doi: 10.1109/TKDE.2005.66.
[33] J. S. Rhodes, A. Cutler, and K. R. Moon, “Geometry- and accuracy-preserving random forest proximities,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 9, pp. 10947-10959, 2023, doi: 10.1109/TPAMI.2023.3263774.
[34] Z. Fa and X. Li, “Aggregation to part: A research on the shift of coal mine safety records in China as a result of risk pre-control management system,” Resources Policy, vol. 73, Art. no. 102159, 2021, doi: 10.1016/j.resourpol.2021.102159.
[35] Y. Xiong, H. Qi, Z. Li, and Q. Zhang, “Where is risk and where is capability? Construction of the emergency management capability structure of coal mining companies under the idea of risk matching,” Resources Policy, vol. 83, Art. no. 103695, 2023, doi: 10.1016/j.resourpol.2023.103695.