Heliyon 9 (2023) e21724

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Comprehensive safety risk evaluation of fireworks production enterprises using the frequency-based ANP and BPNN

Feiyue Wang [a], Xinyu Wang [a], Dingli Liu [b], Hui Liu [a,*]

[a] Institute of Disaster Prevention Science and Safety Technology, School of Civil Engineering, Central South University, Changsha, 410075, Hunan, China

[b] School of Traffic & Transportation Engineering, Changsha University of Science and Technology, Changsha, 410114, Hunan, China

ARTICLE INFO

Keywords:
Fireworks
Safety risk evaluation
Frequency-based ANP
BPNN
Grounded theory

ABSTRACT

The fireworks industry has long struggled with the problem of safety. Scientific, reasonable, and operable evaluation models are prerequisites of reducing risk. Based on the data from over 100 fireworks production safety accidents in China from 2010 to 2022, two evaluation models were established from the perspective of safety risk definition. Firstly, a weight calculation derivative method, the frequency-based analytic network process (ANP), was proposed creatively. This method optimized the importance ranking index calculation process in the ANP by considering the causal frequency of risk factors in the historical accident samples, thus determining how much each indicator affects the likelihood of accidents. Secondly, utilizing the historical accident samples as the dataset, a back propagation neural network (BPNN) model was developed to extract the mathematical relationship between each risk factor and the severity of accident consequence. Finally, the frequency-based ANP and BPNN models were combined to determine the safety risk level of the fireworks production enterprises. Meanwhile, the safety evaluation research samples were used as the comparison set for empirical study with historical accident samples, involving 100 fireworks production enterprises in China evaluated from 2017 to 2020. The significance result of zero shows that there is a statistically significant difference between the likelihood evaluation results of the accident and non-accident companies. Additionally, the severity evaluation model exhibits an excellent result, revealing a classification accuracy of 98.21 %, a mean square error of 8.97×10^{-4}, a percent bias of 1.24 %, and a correlation coefficient and Nash-Sutcliffe efficiency coefficient both of 0.96. The frequency-based ANP and BPNN models integrate self-learning, self-adaptive, and fuzzy information processing, obtaining more accurate and objective evaluation results. This work provides a new strategy for the promotion and application of artificial intelligence in the field of safety risk evaluation, thus offering real-time safety risk evaluation and decision support of the safety management for the enterprises.

1. Introduction

China is the largest producer, distributor, and exporter of fireworks in the world [1–3]. There are many risks and hidden dangers behind the vast market and production scale. Fireworks are primarily made from flammable and explosive pyrotechnic powder [4], which are extremely sensitive to any mechanical process, leading that modernizing the manufacturing facility difficult [5]. Meanwhile,

* Corresponding author.
E-mail address: lhui0421@163.com (H. Liu).

https://doi.org/10.1016/j.heliyon.2023.e21724
Received 23 June 2023; Received in revised form 19 October 2023; Accepted 26 October 2023
Available online 3 November 2023

most employees are from less developed areas with an older age per capita and a lower education level, and a need for more awareness and skills in safety production [6]. Therefore, compared with other manufacturing industries, the production process of the fireworks industry involves more couplings of risky factors, thus resulting in a greater accident rate [7,8].

Safety risk evaluation is the basis and premise of risk control. A scientific evaluation model can not only help the safety supervision department accurately review the safety production conditions of companies, but also enable the graded control measures based on it to accomplish the effect of outlining. After nearly a century of research, safety risk evaluation has successfully shifted from qualitative to quantitative, and some corresponding software have been created based on computer technology. However, the manufacture of fireworks is mainly manual [9,10], unlike other industries with standard production lines and instrumentation diagrams [11]. Apart from that, evaluation indicators are mostly qualitative that are difficult to quantify. Therefore, the study on comprehensive safety risk evaluation of fireworks production enterprises is still mainly qualitative methods, only reaching conformity conclusions, such as what-if analysis [12] and job safety analysis [13], semi-quantitative or quantitative methods with subjective judgment, such as fuzzy approach [14], risk assessment for safety and health and chemical health risk assessment [6], and hazard identification and risk assessment [15], and quantitative methods that evaluate only from the severity like fire & explosion index [16], or from the likelihood like prediction human error analysis technique [17]. Previous studies have favorably explored the issue of uncertainty in the evaluation process from different perspectives. However, there is still a lack of evaluation models that can comprehensively, efficiently, and objectively quantify the safety risk of fireworks production enterprises.

Safety risk evaluation has become more effective and informed with advancements in artificial intelligence, machine learning, and deep learning [18,19]. The back propagation neural network (BPNN) is a mathematical model that intelligently processes data by simulating the human brain, including learning, recognition, and self-adaptation. The BPNN model trains the network repeatedly to find patterns between sample inputs and outputs using an error back propagation algorithm. The neural network model has been widely used in evaluation [20], prediction [21], classification [22], and other fields with good results. In particular, Indumathi et al. [23] developed an artificial neural network model to predict occupational accidents, which took values from historical accident data due to the atmospheric conditions for Sivakasi (2009–2021). The proposed model by Indumathi et al. gave the highest accuracy compared with other models, but needed more comprehensive consideration of risk factors.

In light of the above considerations, the BPNN model was introduced into the severity modeling process. However, due to the lack of quantitative data on the likelihood and taking into account the numerous factors affecting the safety of fireworks production and their interactions, the analytic network process (ANP) was optimized and used to determine the corresponding weights of the evaluation indicators. The key idea of the ANP is to construct a comparison matrix using the nine-scale method, utilizing each factor as the criterion for a two-by-two comparison of the factors influenced by that factor [24]. However, the subjective character of expert scoring and the vagueness of the judgment boundary make it difficult to draw a clear line between the relative importance of two factors in practical applications. The causal frequency may indirectly indicate the importance of each risk factor in the chain of accidents [25], which is useful to improve the risk assessment quality and prevent accidents [26]. Therefore, the frequency-based ANP was proposed to objectively determine the importance ranking index by substituting the causal frequency of the indicators for the subjective judgment of experts. According to the historical accident data, this work aims to eliminate the overlap and subjectivity of evaluation indicators information in multiple links. Moreover, the comprehensive safety risk evaluation results, regarded as R, can be used to classify enterprises or risk points within the same enterprise and provide decision-making support for safety management.

2. Design of the indicator system

The evaluation indicator system was constructed from the perspective of accident causation. The grounded theory (GT) [27] was

Table 1
Part of the open coding process and results.

Examples of original sentences	Concepts	Subcategories
After the victim returned from playing cards outdoors, the security managers saw him in a bad mood and tried to persuade him not to work that day, but victim insisted on continuing to work.	Poor mental state of personnel	Physical and mental state of the personnel
When recruiting and arranging work types, the company did not carefully examine the status of the employees' age and so on. The company arranged for victim, who had exceeded the legal retirement age, to engage in the heavy labor production work of filling the filling room with drugs and collecting the cake room with sealing powder.	Poor age and physical condition of personnel	
The emergency plan has not been well practiced or trained, so when the accident first occurred, the staff members on duty panicked and struggled to cope with it.	Inadequate training and rehearsal of emergency plans	Preparation and exercise of the emergency plan
The production safety emergency plan was not produced in line with requirements, arranged for expert assessment, or lodged with the appropriate agency. The plan's relevance, viability, and convergence are weak.	Unqualified emergency plan preparation	
The average daily relative humidity was 29 % on the accident day, while the minimum daily relative humidity was 10 %. The dry weather made it simple to create electrostatic buildup.	Low humidity in the workplace	Humidity
The factory did not strengthen the management of raw materials and drugs according to weather changes, resulting in the explosion of drugs after spontaneous combustion due to moisture.	Raw materials, finished products or machinery and equipment are damped	

applied to systematically abstract the indicator system from historical accident data without preconceptions to meet the requirements of comprehensiveness, purposefulness, and salience. The 112 fireworks production safety accidents collected on the information disclosure platform of the Chinese government were used as root materials. One hundred samples were randomly chosen for coding and analysis, and the rest were set aside as test samples. The indicator system was established using the proceduralised GT [28].

2.1. Open coding, axial coding, and selective coding

After dividing the coded materials into semantically distinct sentences, the similarity and dissimilarity of sentences were analyzed. The 76 conceptualized causal factors of fireworks production safety accidents were identified and grouped into 27 subcategories with related traits and definitions. Table 1 only shows a portion of the open coding process and results due to space limitations. For each concept, only one original statement is excerpted. The concepts and subcategories resulted from open coding were investigated for their potential logical relationships using the paradigm model [29]. The five main categories governed the subcategories were then refined, as shown in Table 2. An example of the analysis process of the paradigm model is shown in Fig. 1, where the phenomenon is the main category. Subcategories and main categories were again gathered and refined based on the principal goal of evaluation. Finally, the "Evaluation indicators system of safety risk for fireworks production enterprises" was identified as the core category of the rooted material.

2.2. Test of coding results (significance level $\alpha = 0.05$)

When the 12 reserved materials were coded at three levels in order, no new concepts, categories, or links could be made, indicating that the refining study of assessment system has hit theoretical saturation. Using SPSS software, the significance of differences in the frequency of the five main categories in each rooting material was tested in order to further confirm the extraction effect of evaluation indicators. Among them, when the main category belonging to the same concept appeared repeatedly, it was only counted once. Firstly, a distribution test was performed using the Kolmogorov-Smirnov (K–S) test [30], which was appropriate for sample size ($n = 112$) greater than 50. As illustrated in Table 3, the results of the significance test were less than α, indicating that the frequency of the 5 main categories did not follow the normal distribution. Therefore, the Friedman test [30] in nonparametric tests was chosen to perform the significance of differences test. The results revealed that there were significant differences among the 5 main categories (significance $p = 0 < \alpha$), and the evaluation indicator system was well constructed. The main categories and subcategories were utilized as primary and secondary indicators, respectively, constructing the evaluation indicator system based on the coding results.

Table 2
Axial coding results.

Subcategories	Main categories
Awareness level of responsibility among safety managers (A_1)	The safety risk level of personnel (A)
Literacy level of safety among workers (A_2)	
Quota situation of workers (A_3)	
Physical and mental state of the personnel (A_4)	
Setting condition of safety facilities and equipment (B_1)	The safety risk level of equipment (B)
Working condition of machinery and equipment (B_2)	
Qualified status of tools (B_3)	
Electrostatic (C_1)	The safety risk level of environment (C)
Temperature (C_2)	
Humidity (C_3)	
Arrangement of production processes (C_4)	
Situation of overall layout (C_5)	
Qualified compliance of raw and auxiliary materials (D_1)	The safety risk level of material (D)
Drug residue situation (D_2)	
Quantitative production, storage and transportation situation (D_3)	
Safety education and training situation (E_1)	The safety risk level of management (E)
Qualification status of security managers (E_2)	
Qualification status of special workers (E_3)	
Preparation and exercise of the emergency plan (E_4)	
Implementation status of raw material access system (E_5)	
Implementation status of hazardous materials storage and transportation system (E_6)	
Implementation status of the full production safety responsibility system (E_7)	
Construction status of regulations (E_8)	
Construction status of the organization of safety production (E_9)	
Situation of hidden danger investigation and rectification (E_{10})	
Acquisition and maintenance status of equipment and facilities (E_{11})	
Management of the safety production site (E_{12})	

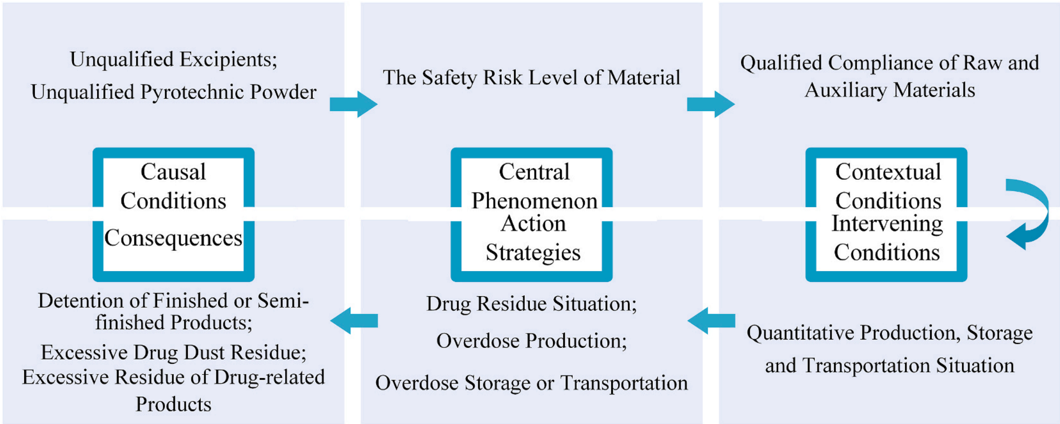


Fig. 1. Example of the analysis process of the paradigm model.

Table 3
K–S test results.

Main categories	*A*	*B*	*C*	*D*	*E*
p	3.5380×10^{-10}	6.6501×10^{-18}	1.3419×10^{-15}	2.2159×10^{-17}	1.4641×10^{-18}

3. Methods

3.1. Evaluation scale of indicators

Given that the causal frequency of each indicator indicates its importance in the accident chain, the evaluation scale proposed in this paper is characterized by the conceptual frequencies under each indicator. Therefore, in the evaluation process of likelihood (L) and severity (S), the normalized dimensionless evaluation values $V_L\_I_k$ and $V_S\_I_k$ of secondary indicator I_k ($k = 1, 2, \ldots, m_I$) under the primary indicator I ($I = A, B, \ldots, E$) are set as equations (1) and (2), respectively:

$$V_L\_I_k = f'_L\_I_k \Big/ f_L\_I_k \tag{1}$$

$$V_S\_I_k = f'_S\_I_k \Big/ f_S\_I_k \tag{2}$$

where m_I is the number of secondary indicators that I covers; $f_L\_I_k$ is the number of concept categories that I_k covers; $f'_L\_I_k$ and $f'_S\_I_k$ are the frequency of concepts that belong to I_k in one evaluation; $f_S\_I_k$ is the highest frequency of concepts that belong to I_k in a historical accident research data. The impact of each risk factor on likelihood is primarily driven by its quality, while the impact on severity also includes its quantity. Therefore, when the same concept is repeated, it is only recorded once in $V_L\_I_k$, while accumulated in $V_S\_I_k$. If $f'_S\_I_k > f_S\_I_k$, then $V_S\_I_k$ takes 1. For instance, three issues were found when the overall layout (C_5) of an enterprise was examined, including the insufficient number of workplaces, insufficient safe distance from the workplace (two locations), and non-compliant workplace protection level and protective barrier. C_5 covers 6 concepts with $f_S\_C_5$ of 5. Therefore, $V_L\_C_5$ is 0.5, and $V_S\_C_5$ is 0.8.

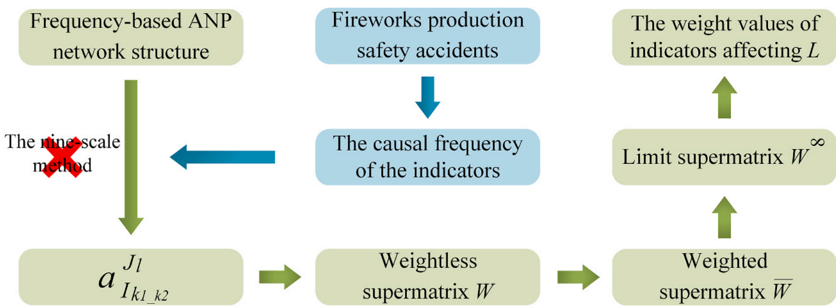


Fig. 2. Flow and architecture diagram of the proposed frequency-based ANP model.

3.2. Frequency-based ANP

Fig. 2 illustrates the flow and architecture diagram of the proposed Frequency-based ANP model. Firstly, the frequency-based ANP network structure of the evaluation system was provided after analyzing the relationship among the risk factors, as shown in Fig. 3. The control layer only contains the decision objective, which is the likelihood of fireworks production safety accidents. Among them, the secondary and primary indicators are also called risk impact factors and factor groups, respectively. The connecting lines denote relationships between factors, and factors in the arrow-tailed factor group influence factors in the factor group pointed by the arrow. Secondly, in historical accident research data, the duplication concepts are eliminated, and the frequency of I_k and J_l together as the accident causal factors, is recorded as $c_{I_k}^{J_l}$ ($c_{I_k}^{J_l} = c_{J_l}^{I_k}$). Then the importance ranking index of I_{k_1} compared to I_{k_2} under the criterion J_l is optimized as equation (3):

$$a_{I_{k1}-I_{k2}}^{J_l} = c_{I_{k1}}^{J_l} / c_{I_{k2}}^{J_l} \tag{3}$$

In this way, 27 factors are progressively utilized as a criterion for a two-by-two comparison of all factors in the same factor group to construct the judgment matrix for each of the 5 factor groups. Following that, the normalized eigenvectors of each judgment matrix are aggregated to produce the weightless supermatrix W. Similarly, using each of the 5 factor groups as the criterion in turn, two-by-two comparisons are made for all factor groups to construct 5 judgment matrices. The normalized eigenvectors of each judgment matrix are combined to produce the weighted matrix $\overline{A}$ ($\overline{A} \stackrel{\text{def}}{=} (\overline{a}_{ij})$) that reflects the relationship between factor groups. As shown in equation (4) [24], The elements of W are weighted to create the weighted supermatrix $\overline{W}$ based on $\overline{A}$.

$$\overline{W} \stackrel{\text{def}}{=} \left(\overline{W}_{ij}\right) \stackrel{\text{def}}{=} \left(\overline{a}_{ij} W_{ij}\right) \tag{4}$$

Finally, the limit supermatrix W^∞ is created by self-multiplication of $\overline{W}$, until the values in each row are stable and constant. The values of W^∞ in each row represent the weight values of relevant risk factors affecting likelihood.

3.3. BPNN with AdamW optimizer

The settings of the learning rate (*LR*) and gradient algorithm significantly impact on the training of a network. As a result, the adaptive moment estimation (Adam) [preprint] [31] with decoupled weight decay (AdamW) [preprint] [32] was introduced to determine the appropriate *LR* and gradient algorithm. The parameters in Adam were updated using the experience gained from previous iterations, which dampened the tendency for oscillations. Based on Adam, AdamW introduced a weight decay (*WD)* term decoupled from gradient descent to regularize larger weights and avoid overfitting the model.

Fig. 4 illustrates the flow and architecture diagram of the proposed BPNN model. According to the above evaluation scale in 3.1, the $V_S\_I_k$ of each indicator I_k in samples was quantified and used as the input data of the BPNN model. The imbalance in the number of samples from different categories may affect the classification function of the model due to insufficient data acquisition. Preprocessing of the data is necessary to avoid model overfitting and enhance the generalizability of the algorithm. As a result, the SMOTE (Synthetic Minority Over-sampling Technique) data enhancement method [33] was applied to acquire new sample data by interpolating between

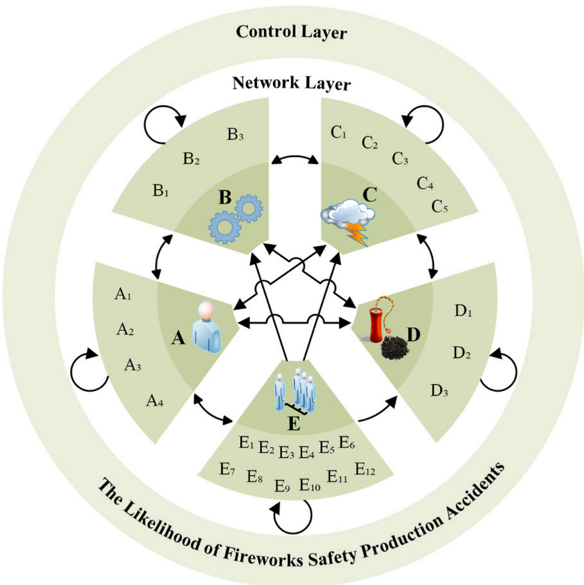


Fig. 3. Frequency-based ANP network structure of safety risk evaluation system for fireworks production enterprises.

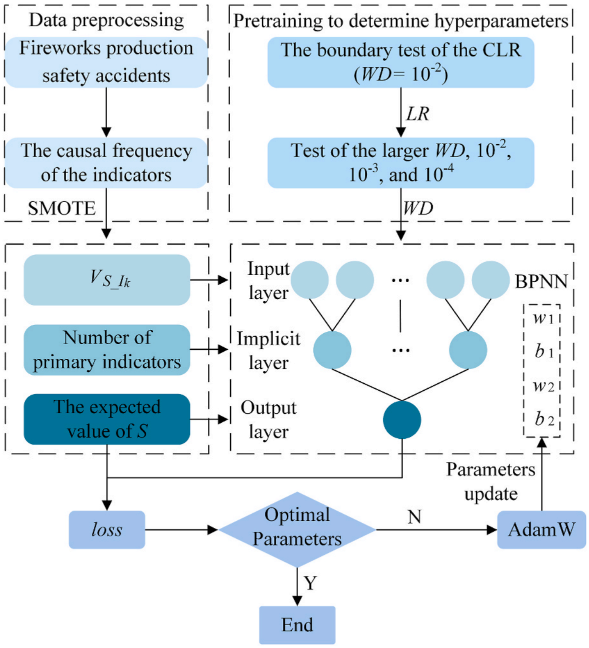


Fig. 4. Flow and architecture diagram of the proposed BPNN model.

samples of small sample classes. Due to the significant ambiguity in the quantification process of severity, the evaluation of severity was converted into a classification problem. The evaluation result level of severity was divided into five levels according to the classification standard of production safety accident level, as shown in Table 4. The acceptable value of the sample output expectation is set to the group median of the corresponding value range of the evaluation result level, and the acceptable mean square error (AMSE) is 0.01. Among these, the mean square error (MSE [34]) is given as equation (5), y_i and $\hat{y}_i$ are the expected value and output value of sample i, respectively.

$$MSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \Big/ n \tag{5}$$

It was found that AdamW increased the separation of the hyperparameters search space [32]. As a result, the value of *WD* was set to the default value of 0.01 in pretraining, looking for a better *LR*, thus modifying the *WD* by using the better *LR*. The *LR* was typically empirically set at a lower value ($10^{-3} \sim 10^{-2}$), which was inefficient for training. To save training time and ensure network convergence, the boundary test of the cyclical learning rate method (CLR) was used to objectively determine the maximum bound of the *LR* [35]. In order to create a graph to show how the loss changes with the *LR*, the *LR* was initially set to a low number and then gradually increased after each iteration. The maximum bound value corresponded to the *LR* when the loss grew inversely. The *LR* was set to the empirical value, the maximum bound value with the CLR method, and a larger value, respectively, and pretraining was performed to determine the better value. In addition, the larger *WD*, 10^{-2}, 10^{-3}, and 10^{-4}, should be tested because the shallow architecture of BPNN needed more regularization [36].

The purpose of sensitivity analysis for the improved BPNN model established by this paper is to identify the key risk factors in control severity. Although there are many sensitivity analysis methods, the fundamental concept is similar. The impact of input neurons on the output was analyzed using the mean influence value (MIV) approach [37], which is a method often used in neural networks. Equation (6) [37] sets $AMIV_{I_k}$, is the absolute value of MIV for I_k.

$$AMIV_{I_k} = \left| \sum_{i=1}^{n} \left(\hat{y}_{i1\_I_k} - \hat{y}_{i2\_I_k} \right) \Big/ n \right| \tag{6}$$

where, two new input samples are created by a 10 % increase and decrease in the value of the input variable corresponding to I_k in

Table 4
Correspondence table of evaluation value and level of severity.

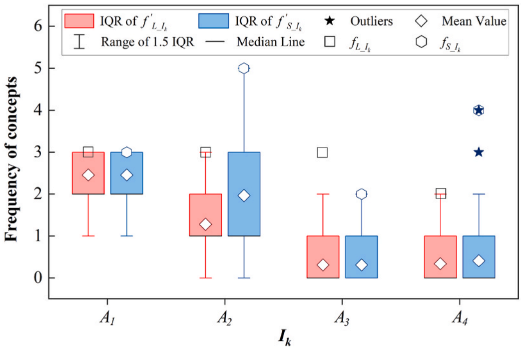
Evaluation result level of *S*	Extremely high	High	Medium	Low	Extremely low
Number of deaths	≥30	[10,30)	[3,10)	[1,3)	0
Evaluation result value of *S*	[0.8,1]	[0.6,0.8)	[0.4,0.6)	[0.2,0.4)	[0,0.2)

sample i, respectively. These two new input samples are then imported into the trained BPNN model to produce new output values of $\hat{y}_{i1\_I_k}$ and $\hat{y}_{i2\_I_k}$, respectively.

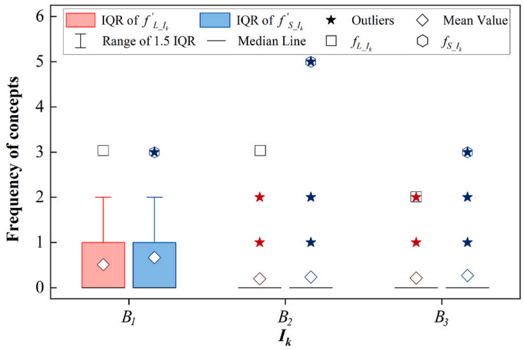
4. Results and discussion

4.1. Frequency results for each indicator

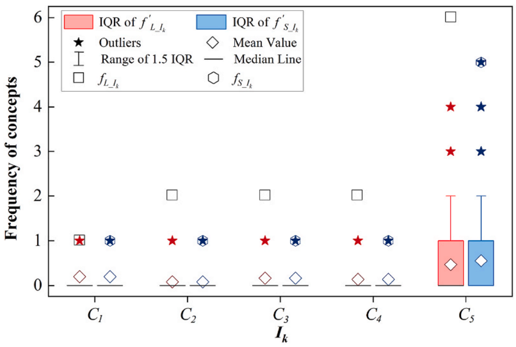
The frequency of each indicator including and excluding the same concepts in each fireworks production safety accident research sample is shown in Fig. 5(a–e), where IQR is the inter quartile range. According to the evaluation scale in 3.1, each $V_L\_I_k$ and $V_S\_I_k$ in samples was quantified.



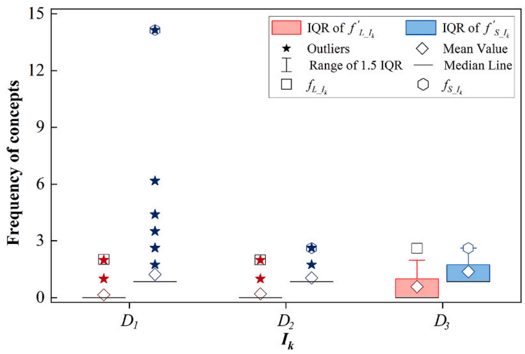
(a) Frequency results for each indicator under *A*



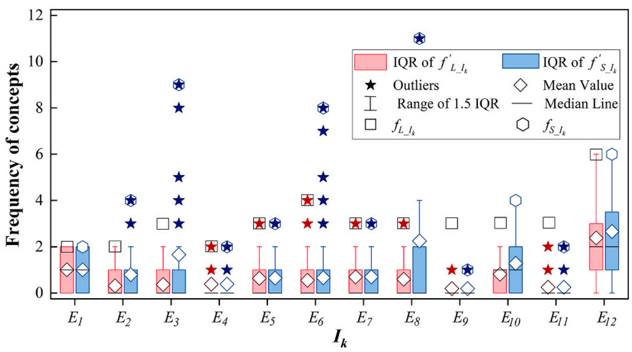
(b) Frequency results for each indicator under *B*



(c) Frequency results for each indicator under *C*



(d) Frequency results for each indicator under *D*



(e) Frequency results for each indicator under *E*

Fig. 5. Frequency results for each indicator.

4.2. Results and empirical study of the likelihood evaluation

It is not feasible to use the current software, such as Super Decision and yaanp, which only provide importance ranking index options in nine-scale, five-scale, three-scale, or two-scale. Furthermore, it is challenging to verify the effectiveness and correctness of manual computation. Therefore, the frequency-based ANP model was built in the Python environment, and its calculation results are shown in Table 5.

To make a comparative empirical study with the 112 accident enterprises, this exploration collected the safety evaluation report of 100 fireworks production enterprises in China. Among them, the samples of accidental and non-accidental enterprises were noted as group N_1 and N_2, respectively, among which n were 112 and 100; group N_2 used data on the production safety conditions before the rectification, and no production safety accidents occurred in the period before and after the evaluation. Fig. 6 shows the evaluation results L of two sample groups. As shown in Fig. 6, the likelihood assessment results of group N_1 are all smaller and more concentrated than those for group N_2. To further verify the significance of the difference between the assessment results of group N_1 and N_2, a K–S test using SPSS software was performed. The Friedman method was also chosen for the significance of differences test because the p in the K–S test were determined as 0.0240 and 6.0893×10^{-11}, respectively. The result of the Friedman test demonstrates a significant difference between the evaluation results of N_1 and N_2 ($p = 0 < \alpha$), which is compatible with the objective fact.

4.3. Results and validation of the severity evaluation

Based on the established evaluation indicator system, the BPNN model adopted a three-layer structure. The input layer was set to the evaluation value of each secondary evaluation indicator with 27 neurons, which was the $V_{S\_}I_k$ obtained from the evaluation scale in 3.1. The number of neurons in the implicit layer was regarded as the number of primary evaluation indicators, which was 5. One neuron was set in the output layer, which was the algebraic value of the severity evaluation result. The loss function adopted the MSE loss function, where $loss = \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{2n} = MSE/2$. Given that AMSE/2 $= 5 \times 10^{-2}$, the goal loss was considered to be 10^{-3}, resulting in MSE being smaller than AMSE. The sigmoid function was used as the activation function since the input and output layers ranged from 0 to 1. The maximum number of iterations was taken as 10^5.340 groups of data were obtained after data preprocessing. One group of each type was then randomly chosen to serve as the test dataset, while the remaining 335 groups served as the training data set. Due to the moderate number of the study for training samples, training was carried out using a full data set.

Fig. 7 displays the results of the LR range test for this dataset. According to the CLR method, the maximum bound value of LR is 9×10^{-2}. Accordingly, the pretraining result is displayed in Fig. 8 when the LR was set to the empirical value of 10^{-2}, the maximum bound value of 9×10^{-2}, and a larger value of 10^{-1}, respectively. The LR determined by the CLR method has a faster training speed compared to smaller LR, and prevents the model from exhibiting overfitting before they reach a predetermined accuracy compared to larger LR. According to the pretraining result of test $loss$ under different WD shown in Fig. 9, the model can demonstrate better generalization performance when the WD is taken as 10^{-2}.

Table 5
Normalized weight values and ranking of evaluation indicators.

Primary indicators	Weighting and ranking of primary indicators	Secondary indicators	Weighting and ranking of secondary indicators
A	0.3531 (2)	A_1	0.0453 (8)
		A_2	0.1955 (1)
		A_3	0.0559 (4)
		A_4	0.0564 (3)
B	0.0781 (5)	B_1	0.0464 (7)
		B_2	0.0132 (20)
		B_3	0.0184 (14)
C	0.0918 (3)	C_1	0.0190 (13)
		C_2	0.0066 (24)
		C_3	0.0140 (17)
		C_4	0.0121 (21)
		C_5	0.0402 (9)
D	0.0798 (4)	D_1	0.0051 (26)
		D_2	0.0196 (12)
		D_3	0.0552 (5)
E	0.3972 (1)	E_1	0.1830 (2)
		E_2	0.0060 (25)
		E_3	0.0137 (19)
		E_4	0.0172 (16)
		E_5	0.0359 (10)
		E_6	0.0279 (11)
		E_7	0.0138 (18)
		E_8	0.0112 (22)
		E_9	0.0040 (27)
		E_{10}	0.0182 (15)
		E_{11}	0.0111 (23)
		E_{12}	0.0552 (5)

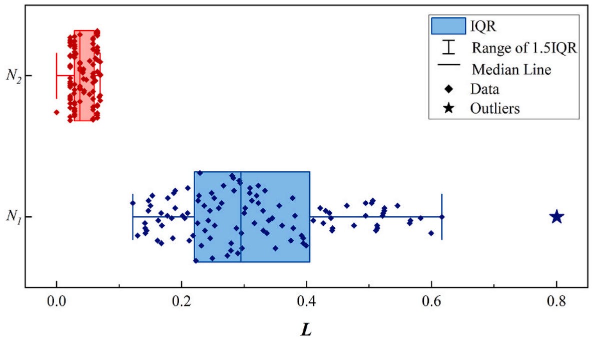


Fig. 6. Empirical study results of the frequency-based ANP model.

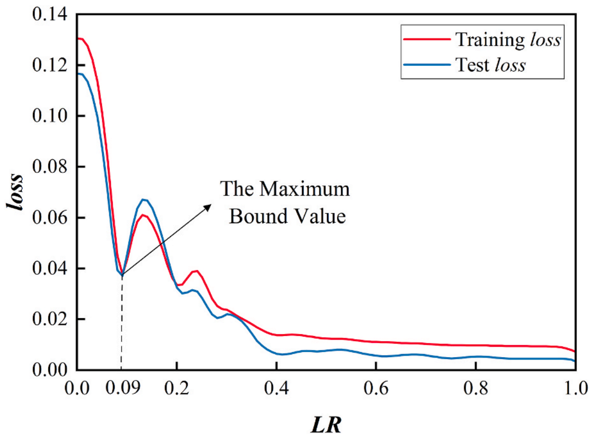


Fig. 7. LR range test results.

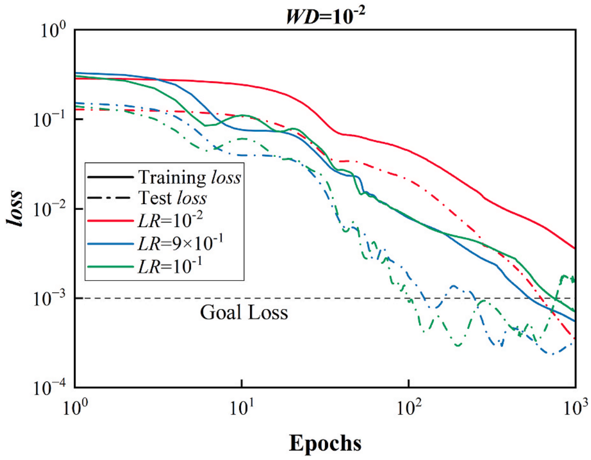


Fig. 8. Pretraining results under different *LR*.

In summary, the BPNN model was trained with the hyperparameters shown in Table 6, and the training results are shown in Fig. 10. Based on that, the optimal parameters of the BPNN severity assessment model were derived. Fig. 11 illustrates the evaluation results *S* of accident enterprises. The following metrics were employed to evaluate the performance of model: CA (classification accuracy, equation (7) [38]), MSE, R^2 (correlation coefficient, equation (8) [39]), NSE (Nash-Sutcliffe efficiency coefficient, equation (9) [39]), and PBIAS (percent bias, equation (10) [40]). Among these, Table 7 shows the meanings of TP, TN, FN and FP in the binary categorization problem, and the categorization results include both P (positive) and N (negative) categories; $\overline{y}$ and $\overline{\hat{y}}$ are the mean expected

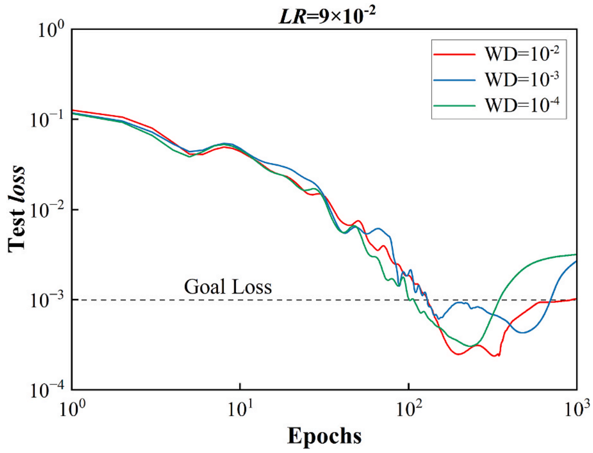


Fig. 9. Pretraining results under different *WD*.

Table 6
Hyperparameters for training BPNN model.

Hyperparameter	Value	Hyperparameter	Value	Hyperparameter	Value
Layers	3	Number of neurons in the input layer	27	Number of neurons in the implicit layer	5
Number of neurons in the output layer	1	Goal loss	10^{-3}	The maximum number of iterations	10^{-5}
Batch Size	335	*LR*	9×10^{-2}	*WD*	10^{-2}

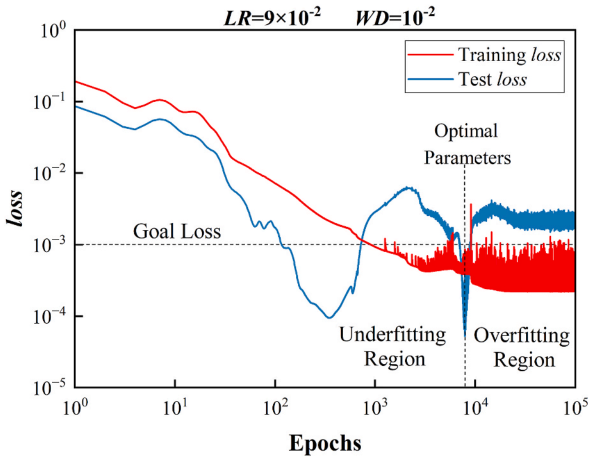


Fig. 10. BPNN model training results.

value and mean output value for all samples, respectively. As seen in Table 8 and Fig. 11, CA is close to 100 %. Besides, MSE and PBIAS are both close to 0, and R^2 and NSE are both close to 1, indicating a good match between the evaluation and expected values.

$$\mathrm{CA} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FN}+\mathrm{FP}+\mathrm{TN}} \tag{7}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n} (y_i - \overline{y})(\widehat{y}_i - \overline{\widehat{y}})}{\sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2 \sum_{i=1}^{n} (\widehat{y}_i - \overline{\widehat{y}})^2}} \right)^2 \tag{8}$$

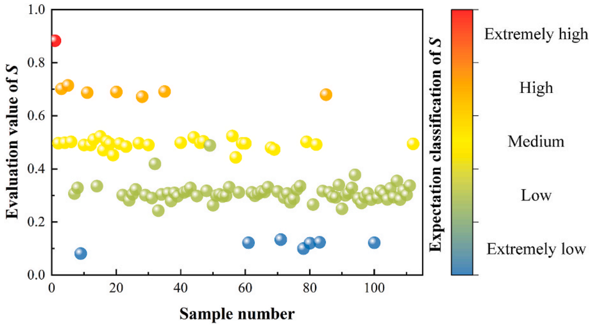


Fig. 11. Evaluation results for historical accident samples under the BPNN model.

Table 7
The meanings of TP, TN, FN and FP.

Classification of expectations	Classification of outputs	
	P	N
P	TP	FN
N	FP	TN

Table 8
The evaluation effect of the model.

CA	MSE	R^2	NSE	PBIAS
98.21 %	8.97×10^{-4}	0.96	0.96	1.24 %

$$NSE = 1 - \frac{\sum_{i=1}^{n} (\widehat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \tag{9}$$

$$PBIAS = \sum_{i=1}^{n} \frac{\widehat{y}_i - y_i}{y_i} \times 100\% \tag{10}$$

4.4. Results of the comprehensive safety risk evaluation

According to the evaluation models of likelihood and severity developed by frequency-based ANP and BPNN, respectively, the comprehensive evaluation results of safety risk were obtained for 212 enterprises. Among group N_2, the enterprises judged to be qualified after rectification were recorded as group N_{2-1}, and the rest were recorded as group N_{2-2}.

The evaluation level of safety risk is divided into four categories based on the safety risk classification and control, including significant risk, higher risk, general risk, and low risk. To be consistent with the actual evaluation results of the enterprises, the upper limit value of low risk was initially set to the minimum value of R for the group N_{2-1}, which was 0.0023. The comprehensive evaluation results of R for groups N_1 and N_{2-1} were then subjected to K-means [41] cluster analysis using SPSS, which provided a scientific and theoretical basis for the division of the remaining three levels. The clustering outcomes are displayed in Table 9, which provide a scientific and theoretical basis for the establishment of the safety risk level assessment scale. The range of values for each evaluation level is modified downward under the strict and high principle.

Fig. 12 displays the results of the comprehensive safety risk evaluation for samples and the level evaluation scale of that. As shown in Fig. 12, most of the samples with general risk and below are from group N_2 with no production safety accidents. Although there are occasional accidents, the severity is not higher, which did not result in many fatalities. The frequency-based ANP and BPNN evaluation

Table 9
Correspondence table of comprehensive evaluation value and evaluation level of R.

Clusters	1	2	3
Number of clustered cases	115	72	24
Clustering center	0.0128	0.0976	0.2548
Corresponding sample value range	0.0023 ~ 0.0564	0.0574 ~ 0.1763	0.1802 ~ 0.3907

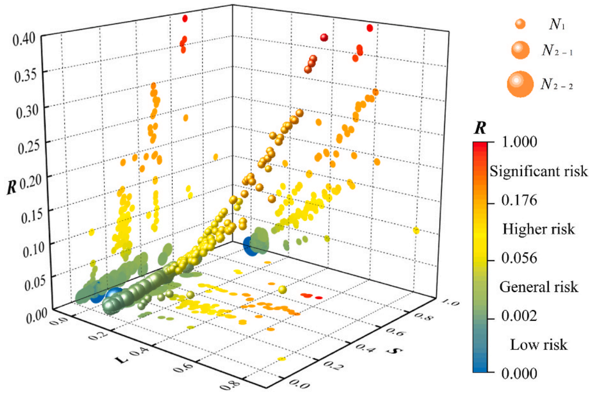


Fig. 12. Comprehensive evaluation results and level evaluation scale of safety risk for fireworks production enterprises.

models can better achieve the goal of reviewing the safety production conditions.

4.5. Discussion of key factors in risk control

The frequency-based ANP and BPNN models are not only useful for the safety risk evaluation, but also can understand the importance of each indicator among the safety risk. Therefore, it is crucial to identify the indicators that can be modified to reduce the risk after considering factors such as cost, feasibility, and effectiveness.

As shown in Table 5, the top five risk factors, including A_2,E_1,A_4,A_3, and E_{12}, accounted for 54.6% of the total weight and the other 22 risk factors accounted for the rest weight. This indicates that eliminating the probability of occurrence for only these five critical risk factors can significantly reduce the occurrence likelihood of fireworks production safety accidents. The importance of the key factors that have not been stressed enough in previous research and practical applications of safety management, such as the physical and mental state of the personnel (A_4), should also be considered. The *AMIV* for risk factors in BPNN model were calculated respectively, and the sequence of the risk factors was sorted according to the *AMIV* value. From Table 10, it is illustrated that C_5, E_{12}, A_1, E_1, and B_1 are the 5 most important factors for severity evaluation, although the safety risk level of materials (D) is directly related to severity. Environment factors like C_5 and equipment factors like B_1 are the key measures to limit D, and the failure of measures largely contributes to the expansion of severity. Besides, management factors like E_{12} and personnel factors like A_1 increase severity by raising the safety risk level of materials, environment, and equipment. Therefore, factors in D are less sensitive in the severity assessment model.

5. Conclusion

The goal of this study is to increase the objectivity and accuracy of the safety risk evaluation of enterprises that produce fireworks from multiple perspectives. To remove the interference of redundant information, GT was used for the first time in the fireworks field to thoroughly and methodically refine the evaluation indicators system from historical accident causation data. Then, from the perspective of the definition of safety risk, frequency-based ANP and BPNN models have been established and improved based on the evaluation characteristics of likelihood and severity indicators, which are of pioneering significance.

Compared with traditional methods like the fuzzy approach, the frequency of evaluation indicators was innovatively used as the basis for quantifying their weights and values, which simultaneously improved objectivity while maintaining a high degree of operationalization. Unlike previous attempts of artificial intelligence in the field of safety risk evaluation of fireworks production enterprises, the two models proposed in this study consider the risk-influencing factors in a comprehensive way. In addition, the likelihood was not simply regarded as two limit states, but the evaluation of their specific values were explored. The significance of the

Table 10
Sensitivity analysis results of BPNN model.

Risk Factor	*AMIV*	Ranking of the risk factors	Risk Factor	*AMIV*	Ranking of the risk factors	Risk Factor	*AMIV*	Ranking of the risk factors
A_1	0.0092	3	A_2	0.0006	21	A_3	0.0012	17
A_4	0.0020	12	B_1	0.0072	5	B_2	0.0030	9
B_3	0.0015	14	C_1	0.0021	11	C_2	0.0017	13
C_3	0.0007	20	C_4	0.0008	19	C_5	0.0200	1
D_1	0.0001	26	D_2	0.0005	22	D_3	0.0001	26
E_1	0.0082	4	E_2	0.0014	15	E_3	0.0036	7
E_4	0.0002	24	E_5	0.0002	24	E_6	0.0013	16
E_7	0.0032	8	E_8	0.0030	9	E_9	0.0003	23
E_{10}	0.0038	6	E_{11}	0.0009	18	E_{12}	0.0114	2

difference test and the results of the five performance metrics verify that the proposed evaluation indicator system and models provide a realistic way for companies to determine their safety risk level in an objective manner. Taking into account the specific circumstances of the company and the results of a sensitivity analysis, it is also effective and practicable to make decisions for safety management.

Due to the limited sample size, there are some gaps between the five metrics characterizing the performance of the BPNN model and the ideal values, which means that there are some systematic biases in the evaluation results. Thus, it is necessary to increase the size of the data set used to build the models in the future. And specialists and academics in the fields of computer science and safety must work together to further improve the accuracy and speed of intelligent safety risk evaluation.

Funding statement

This work was supported by the National Natural Science Foundation of China (No. 51974362) and the Government Procurement Project of Department of Emergency Management of Hunan Province (H201912280480001). This work was supported in part by the High Performance Computing Center of Central South University.

Data availability statement

Data will be made available on request.

CRediT authorship contribution statement

Feiyue Wang: Conceptualization, Funding acquisition, Project administration, Resources, Writing – review & editing. **Xinyu Wang:** Data curation, Investigation, Methodology, Software, Writing – original draft. **Dingli Liu:** Formal analysis, Supervision, Validation, Writing – review & editing. **Hui Liu:** Formal analysis, Supervision, Validation, Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C. Cao, W. Chen, C. Shu, Prediction and assessment of fly-up type of fireworks by thermokinetics model, J. Therm. Anal. Calorim. 142 (2) (2020) 927–936, https://doi.org/10.1007/s10973-020-09840-3.
[2] Q. Li, Y. Yu, F. Wang, S. Chen, Y. Yin, H. Lin, F. Che, P. Sun, J. Qin, J. Liu, H. Wang, Urinary perchlorate exposure and risk in women of reproductive age in a fireworks production area of China, Arch. Environ. Contam. Toxicol. 67 (1) (2014) 42–49, https://doi.org/10.1007/s00244-014-0042-6.
[3] S. Nandhagopal, K. Praveen Kumar, B.P. Yadav, A. Nandan, N.A. Siddiqui, Occupational hazards and risk exposures in firework industry: a review, in: HSFEA: International Conference on Advances in the Field of Health, safety, fire, environment, allied sciences and engineering, Singapore, 2020, pp. 243–252, https://doi.org/10.1007/978-981-15-6852-7_22.
[4] W. Chen, Y. Lin, S. Ouyang, Y. Lee, Effect of relative humidity on the emission height and reaction force of single-tube fireworks, J. Therm. Anal. Calorim. 148 (11) (2023) 4997–5007, https://doi.org/10.1007/s10973-023-12033-3.
[5] V.P. Sridhar, M. Surianarayanan, S.P. Sivapirakasam, A.B. Mandal, Accelerating rate calorimeter studies of water-induced thermal hazards of fireworks tip mixture, J. Therm. Anal. Calorim. 112 (3) (2013) 1335–1341, https://doi.org/10.1007/s10973-012-2716-5.
[6] S. Ajith, C. Sivapragasam, V. Arumugaprabu, A review on hazards and their consequences in firework industries, SN Appl. Sci. 1 (1) (2018) 120, https://doi.org/10.1007/s42452-018-0129-1.
[7] R.M. Billock, T. Chounthirath, G.A. Smith, Pediatric firework-related injuries presenting to United States emergency departments, 1990-2014, Clin. Pediatr. 56 (6) (2016) 535–544, https://doi.org/10.1177/0009922816664063.
[8] A. Arumugachamy, P.L. Pandian, An experimental study of mitigation of fire hazards of flash powder by adding mono ammonium phosphate as a solid inertant for explosive waste disposal in firework industries, Environ. Sci. Pollut. Res. (2023), https://doi.org/10.1007/s11356-023-29126-z.
[9] A. Arumugachamy, P. Pandian, K. Jeyasubramanian, Experimentation on minimum ignition temperature for boron blended fireworks flash powder dust cloud at normal and inert environmental conditions, Fire Mater. 46 (2) (2021) 410–419, https://doi.org/10.1002/fam.2977.
[10] N. Indumathi, R. Ramalakshmi, N. Selvapalam, Application of machine learning techniques for an evaluation of hazardous chemical assessment in firework industry, AIP Conf. Proc. 2520 (1) (2022), 020008, https://doi.org/10.1063/5.0103004.
[11] S. Rajaram, G. Sivakumar, Assessment of safety culture in the fireworks industry, Int. J. Occup. Saf. Ergon. 29 (2) (2022) 466–473, https://doi.org/10.1080/10803548.2022.2079830.
[12] S.P. Sivapirakasam, M. Surianarayanan, G. Swaminathan, Hazard assessment for the safe storage, manufacturing and handling of flash compositions, J. Loss Prev. Process. Ind. 22 (2) (2009) 254–256, https://doi.org/10.1016/j.jlp.2008.09.001.
[13] N.K. Singh, V. Banodha, P. Patel, Job safety analysis in fire crackers industry, Int. J. Sci. Adv. Res. Technol. 5 (10) (2019) 67–69.
[14] P. Tamizhchelvi, Projecting safety measures in fireworks factories in Sivakasi using fuzzy based approach, Int. J. Comput. Appl. Technol. 147 (12) (2016) 22–27, https://doi.org/10.5120/ijca2016911272.
[15] S. Ajith, V. Arumugaprabu, V. Ajith, K. Naresh, P.S.R. Sreekanth, Hazard identification and risk assessment in firework industry, Mater, Today Off.: SAVE Proc. 56 (3) (2022) 1083–1085, https://doi.org/10.1016/j.matpr.2021.10.102.
[16] M. Anandhan, T. Prabaharan, M. Muhaidheen, S. Balamurugan, Fire safety assessment in manufacturing of fireworks products, Int. J. Mech. Eng. Technol. 9 (10) (2018) 1336–1345.
[17] N. Indumathi, R. Ramalakshmi, A.O. Salau, T.U. Badrudeen, C.A. Mmonyi, Predictive analytics of human errors in the fireworks industry, in: 2022 International Conference on Decision Aid Sciences and Applications (DASA), Chiangrai, Thailand, 2022, pp. 234–238, https://doi.org/10.1109/DASA54658.2022.9765179.
[18] M. Drakaki, P. Tzionas, Investigating the impact of site management on distress in refugee sites using Fuzzy Cognitive Maps, Int. J. Disaster Risk Reduc. 60 (2021), 102282, https://doi.org/10.1016/j.ijdrr.2021.102282.
[19] N. Indumathi, P. Savaram, S. Sengan, M. Alharbi, S. Alshathri, M. Bajaj, M.H. Aly, W. El-Shafai, Impact of fireworks industry safety measures and prevention management system on human error mitigation using a machine learning approach, Sensors 23 (9) (2023) 4365, https://doi.org/10.3390/s23094365.

[20] A. Bakshi, M. Esraz-Ul-Zannat, Application of urban growth boundary delineation based on a neural network approach and landscape metrics for Khulna City, Bangladesh, Heliyon 9 (6) (2023), e16272, https://doi.org/10.1016/j.heliyon.2023.e16272.
[21] M. Li, L. Zhao, S. Jin, D. Li, J. Huang, J. Liu, Process schemes of ethanol coupling to C4 olefins based on a genetic algorithm for back propagation neural network optimization, Heliyon 8 (12) (2022), e12301, https://doi.org/10.1016/j.heliyon.2022.e12301.
[22] M.C. Guerrero, J.S. Parada, H.E. Espitia, EEG signal analysis using classification techniques: logistic regression, artificial neural networks, support vector machines, and convolutional neural networks, Heliyon 7 (6) (2021), e07258, https://doi.org/10.1016/j.heliyon.2021.e07258.
[23] N. Indumathi, R. Ramalakshmi, D.A. Pustokhin, I.V. Pustokhina, D.K. Sharma, S. Sengan, Prediction of influencing atmospheric conditions for explosion Avoidance in fireworks manufacturing Industry-A network approach, Environ. Pollut. 304 (2022), 119182, https://doi.org/10.1016/j.envpol.2022.119182.
[24] G. Ozkaya, C. Erdin, Evaluation of smart and sustainable cities through a hybrid MCDM approach based on ANP and TOPSIS technique, Heliyon 6 (10) (2020), e05052, https://doi.org/10.1016/j.heliyon.2020.e05052.
[25] S. Li, M. You, D. Li, J. Liu, Identifying coal mine safety production risk factors by employing text mining and Bayesian network techniques, Process Saf. Environ. Protect. 162 (2022) 1067–1081, https://doi.org/10.1016/j.psep.2022.04.054.
[26] B. Wang, J. Zhao, Automatic frequency estimation of contributory factors for confined space accidents, Process Saf. Environ. Protect. 157 (2022) 193–207, https://doi.org/10.1016/j.psep.2021.11.004.
[27] J. He, T. Wang, The community proactive health management model based on the grounded theory: the case of Beijing, China, Heliyon 9 (4) (2023), e14992, https://doi.org/10.1016/j.heliyon.2023.e14992.
[28] X. Han, J. Zhang, Business model innovation paths of manufacturing oriented towards green development in digital economy, Int. J. Environ. Res. Publ. Health 19 (24) (2022), 16454, https://doi.org/10.3390/ijerph192416454.
[29] J. Huang, D. Yoon, Paradigm model of online learning experience during COVID-19 crisis in higher education, Front. Educ. 8 (2023), https://doi.org/10.3389/feduc.2023.1101160.
[30] G.W. Corder, D.I. Foreman, Nonparametric Statistics: A Step-by-step Approach, Wiley, Hoboken, US, 2014.
[31] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv 1412 (6980) (2014), https://doi.org/10.48550/arXiv.1412.6980.
[32] I. Loshchilov, F. Hutter, Decoupled weight decay regularization, in: International Conference on Learning Representations, Toulon, France, 2017, https://doi.org/10.48550/arXiv.1711.05101 [preprint].
[33] X. Zhen, Y. Ning, W. Du, Y. Huang, J.E. Vinnem, An interpretable and augmented machine-learning approach for causation analysis of major accident risk indicators in the offshore petroleum industry, Process Saf. Environ. Protect. 173 (2023) 922–933, https://doi.org/10.1016/j.psep.2023.03.063 [preprint].
[34] K.C. Igwe, O.D. Oyedum, A.M. Aibinu, M.O. Ajewole, A.S. Moses, Application of artificial neural network modeling techniques to signal strength computation, Heliyon 7 (3) (2021), e06047, https://doi.org/10.1016/j.heliyon.2021.e06047.
[35] L. N.S, Cyclical learning rates for training neural networks, in: 2017 IEEE Winter Conference on Applications of Computer Vision, WACV), Santa Rosa, CA, 2017, pp. 464–472, https://doi.org/10.1109/WACV.2017.58.
[36] Y. Bi, S. Wang, C. Zhang, H. Cong, B. Qu, J. Li, W. Gao, Safety and reliability analysis of the solid propellant casting molding process based on FFTA and PSO-BPNN, Process Saf. Environ. Protect. 164 (2022) 528–538, https://doi.org/10.1016/j.psep.2022.06.032.
[37] Y. Bi, S. Wang, C. Zhang, H. Cong, B. Qu, J. Li, W. Gao, Safety and reliability analysis of the solid propellant casting molding process based on FFTA and PSO-BPNN, Process Saf. Environ. Protect. 164 (2022) 528–538, https://doi.org/10.1016/j.psep.2022.06.032.
[38] J. Lu, X. Tong, H. Wu, Y. Liu, H. Ouyang, Q. Zeng, Image classification and auxiliary diagnosis system for hyperpigmented skin diseases based on deep learning, Heliyon 9 (9) (2023), e20186, https://doi.org/10.1016/j.heliyon.2023.e20186.
[39] S. Samantaray, P. Sahoo, A. Sahoo, D.P. Satapathy, Flood discharge prediction using improved ANFIS model combined with hybrid particle swarm optimisation and slime mould algorithm, Environ. Sci. Pollut. Res. 30 (35) (2023) 83845–83872, https://doi.org/10.1007/s11356-023-27844-y.
[40] E. Tesfaye, B. Abate, T. Alemayehu, Y. Dile, Evaluating the skills of the CMIP5 global climate models using multicriteria decision-making analysis in Ethiopia, Heliyon 9 (10) (2023), e20320, https://doi.org/10.1016/j.heliyon.2023.e20320.
[41] Y. Zhao, X. Huang, Y. Zhao, X. Liu, R. Zhou, The application of landscape character classification for spatial zoning management in mountainous protected areas – a case study of Laoshan national park, China, Heliyon 9 (3) (2023), e13996, https://doi.org/10.1016/j.heliyon.2023.e13996.