Social stability indicator system and its early warning model construction

Jinrun Xu, Menglu Liu, Tiantai Shao, Chan Shu*

Huanggang Normal University, HuBei, China

ABSTRACT

We are now in the midst of a rapidly developing information age, and it is not easy to maintain social stability and overcome social crises in an increasingly competitive and crisis-prone modern society. A social stability mechanism system is an organic whole consisting of multiple types of social stability mechanisms combined in a certain way, which can play a "synergistic" role in stabilising society. How to build a scientific, rational and accurate social stability assessment system and its early warning model is an important step to change the emergency response at the end to maintaining stability at the source, which is a common concern in the world, and is also the main object of this modelling study.

To address question one, we collected data reflecting various aspects of social stability in China from official public databases and statistical yearbooks such as the National Bureau of Statistics of China, the World Bank and the World Trade Organisation, and supplemented and standardised these data with missing values. Taking China as an example, a first-level indicator assessment system was constructed for four dimensions of policy making, social environment, public safety and natural environment, containing a total of 20 sub-dimensions. The rationality of the social stability indicator system is analysed from both quantitative and qualitative perspectives, and the correlations and causal relationships between the four dimensions are explored.

For problem two, the collected data were normalised and standardised, and the prediction targets and feature indicators were selected to build an early warning model of social stability. Based on the processed data, a support vector machine (SVM) model was built, and the warning level was selected as the prediction target and the corresponding sub-dimensions as the feature indicators. The parameters of the SVM model are optimised separately and the model loss is calculated. The parameter corresponding to the smaller loss is the final parameter setting, and the internal parameter settings are adjusted using the grid search method, and the results show that the optimal parameters c=0.25 and g=0.7011.

For question three, Belarus was used as an example to process data related to the country of Belarus by referring to the method used in question one to collect and process data from China. Based on the model already established in question two, the processed data are input into the SVM model as characteristic indicators, thus realising the evaluation of social stability in Belarus. Finally, a grey evaluation system is constructed and the results are obtained using grey correlation analysis and effective suggestions are made for social stability based on the analysis results.

In response to question four, the data relating to the country of Ukraine was approached with reference to the methodology used to collect and process the Chinese data in question one. Firstly, the causes of the colour revolutions leading to regime change in the country of Ukraine were explored. Secondly, the gap between each sub-dimension and the stability warning level was analysed through grey correlation analysis. Finally, the main causes of regime change are identified in relation to the relevant analytical documents and the results of this question.

In response to question five, based on the results of the analysis of the above four questions, and taking into account the relevant literature, practical suggestions are made to prevent colour revolutions and maintain social stability from both the state and individual perspectives.

Keywords: Social Stability Indicator System; Early Warning Model; SVM; Grey Correlation Analysis; Prevention Recommendations

1 INTRODUCTION

1.1 Problem Background

Humans, like all animals, have the instinct to avoid harm. The key to human beings being the masters of all things is that they are better at avoiding harm than other animals. Crisis is always lurking in the future, and the history of human development is a history of constant attempts to transcend it. With the rapid development of science and technology, the world has entered the age of information technology. In a modern society where competition is increasingly fierce and crises are frequent, overcoming them is no easy task. We must rely on modern social measurement science and modern social early warning science to accurately warn of social crises in order to facilitate the sustainable and healthy development of society.

Social stability is a relatively broad concept, covering many disciplinary fields such as political science, sociology, management, economics, law and psychology. In fact, any factor that may cause the poor functioning of society is the object of social stability research, such as social control, economic support, social distribution, external environment, etc. Therefore, the study of social stability indicator systems is an important step to change the emergency response at the end into stability maintenance at the source, and an important prerequisite for early warning and prevention of unstable problems, which is conducive to comprehensively and accurately grasping the operation status of social stability, making more scientific predictions on the stability situation, and preventing and reducing the occurrence of unstable problems at the source [1]. At present, studies on social stability indicator systems are springing up and playing an important role in maintaining stability.

1.2Restatement of the Problem

Based on the problem context, develop a mathematical model and address the following guiding questions:

The indicator system of social stability is an important prerequisite for social stability early warning. Selecting indicators that are representative of each can reflect all aspects of social stability comprehensively [2]. Please establish a system of indicators that are likely to affect social stability from both qualitative and quantitative perspectives, and discuss the correlation and causality between them.

A stable society requires synergy, checks, and balances among indicators. For example, a society in economic hardship may use the human spirit to compensate for economic deficiencies, which can also make society desperately in need of stability, but only if it can survive [3]. Please consider similar ideas, develop an early warning model of social stability, and discuss.

Select a country or region where a color revolution attempted to overthrow the regime, and evaluate its social stability by using the established social stability early warning model. This paper points out the main reasons for the failed color revolution, judges the trend of social stability in the future, and puts forward some suggestions.

Please select a country or region where color revolution leads to regime change and point out the main reasons for regime change by using the established social stability early warning model.

In order to prevent color revolution and maintain social stability, please put forward relevant suggestions.

2 PROBLEM ANALYSIS

To address problem one, a system of social stability indicators is constructed by collecting data that can explain China's social stability from public data sets such as the National Bureau of Statistics of China, the World Bank and the United Nations Statistics Division, describing the rationality of the construction of the system from both qualitative and quantitative perspectives, and using SPSS to analyse the correlations and causal relationships between them.

To address question two, a model of social stability early warning was established. Firstly, the collected data were normalised and standardised, and the prediction targets and characteristic indicators were selected. Secondly, considering the existence of complex functional relationships between the indicators, a simple linear regression model could not meet the requirements. Therefore, the data is processed by building an SVM model that relies on the complexity of its system and adjusts the settings of the internal parameters using the grid search method.

For problem three, the prediction is essentially made using the model already established for problem two. The data set for the social stability system was first collected for Belarus as an example, as required by the question. The processed data is then fed into the SVM model as feature indicators, thus enabling the evaluation of social stability. Finally, based on the constructed grey evaluation system, the results of the grey correlation analysis are used to make recommendations on social stability [4].

In response to question four, by looking for the successful cases of the colour revolution, taking Ukraine as an example, the early warning level is Later, through the grey correlation analysis, the correlation between each in- Later, through the gray correlation analysis, the correlation between each index and the warning level was analyzed, so that the main reasons for the change in Ukraine were obtained combined with the relevant literature.

In response to question five, the results of the analysis of the first four questions and the published research literature were combined to provide practical suggestions for preventing colour revolutions and maintaining social development.

3 BASIC ASSUMPTIONS AND NOTATION

3.1 Assumptions

It is assumed that there will be no world war in the future.

It is assumed that the eigenvectors collected are strongly representative of the level 1 indicators.

Assume that the data collected for the two representative countries, Belarus and Ukraine, are correct.

3.2 Notations

Table 1: Notations

Symbols	Description
μ	Average value

z	zScore
σ	Standard deviation
x_i	Variables
ξ^*	Slack variables
a_i	Lagrange multiplier
$K(x_i, x_j)$	Kernel functions
D^*	Ideal Solution Distance
ρ	Resolution factor

4 MODEL BUILDING AND SOLUTION

4.1 Solution of question 1

4.1.1 Source of data

In the process of mathematical modelling, the selection and collection of data sets is important. A high quality dataset can often improve the quality of model training and the accuracy of predictions, and mathematical models built on the correctness and veracity of the data are more relevant.

China is a country with 9.6 million square kilometres of land, a population of 1.4 billion, one language, a unified culture, the same religion and is already the second largest economy in the world according to data published by countries around the world in 2021 [5]. Unlike the multi-party rotation in the West, where the centre and localities are in a loose alliance, the system is led by the Chinese Communist Party and is highly centralised from the centre to the localities, ensuring the concentration of power on major issues, a political system more conducive to measuring and maintaining social stability. In terms of labour resources, there is a high level of quality among students and community members. In terms of defence and diplomacy, there is active military cooperation, making friends, enhancing military mutual trust and establishing friendly relations [6]. Overall, Chinese society is relatively stable, with no social unrest or collapse. Using China as an example to construct an indicator system for social stability can provide reference value for other countries. Therefore, the data source for this question is China as the main target. To ensure the reliability of the data, the data sources are mainly drawn from official and publicly available databases and statistical yearbooks such as the National Bureau of Statistics of China, the World Bank, the United Nations Statistics Division and the World Trade Organisation (see the table in the Annex for details). It is also the source of data for Q.3 and Q.4, and the data is treated with reference to Q.1.

4.1.2 Data pre-processing

In the real world, data comes from a variety of sources of varying quality, so raw data is generally flawed, incomplete, repetitive and highly susceptible to contamination. This data is not at all suitable for direct data mining, so in order to produce more accurate results, we have to pre-process the raw data. From the perspective of the whole big data processing process, the level of data pre-processing technology determines the authenticity and integrity of the data, and plays a very critical role in the subsequent data analysis [7]. As this data mainly comes from the data sets published by major authoritative officials, the data integrity is good, but there are data missing such as national fixed asset investment, number of people receiving education, international murder crime rate (per 100,000 people), CO_2 emissions, etc. such data. There are missing data for only a few years; investment in assets, the number of educated people is certainly increasing over time, and the international murder rate is decreasing over time. Therefore, the decision to fill in the missing years was made to use regression interpolation.

The regression interpolation method involves constructing a regression equation based on the missing and obtained variables in the sample, i.e. estimating the missing values of the target variables in the survey based on the available sample data [8]. The relationship between the independent variable $X_i(i = 1,2,\cdots,m)$, and the target variable Y is constructed and the interpolated estimate of the kth missing value can be expressed as:

$$y_k = \alpha_0 + \sum_{i=1}^{m} \alpha_i X_{ik} \tag{1}$$

It follows from the above equation that for the same $X_i(i = 1,2,\cdots,m)$, the regression transformation yields the same estimate, which is consistent with the mean interpolation. So a random factor needs to be added to the regression process as a way to fill in the gaps in this interpolation method. At this point the regression equation is expressed as:

$$y_k = \alpha_0 + \sum_{i=1}^{m} \alpha_i X_{ik} + \xi_k \tag{2}$$

Where ξ denotes the construction of the dataset.

This process was implemented using the regression interpolation function of SPSS software to obtain the data set required for this question (see annex qusetion1.xlsx).

4.1.3 The construction of a social stability indicator system

In order to develop a scientific and reasonable social stability indicator system, this question combines the political system stability index, the public policy social stability risk assessment indicator system, the Chinese social security and stability early warning level model and the theoretical framework and indicator system of social stability risk assessment in China from the perspective of social ecosystem governance, and constructs a social stability indicator system from a combination of qualitative and quantitative analysis [9].

As a framework for the quantitative analysis of politics, the Political System Stability Index was co-proposed by Dan Haendel, Gerald T. West and Robert G. Meadow. It consists of three main components: an index of the socio-economic characteristics of the country, an index of social conflict and an index of the role of government (as shown in the figure below).

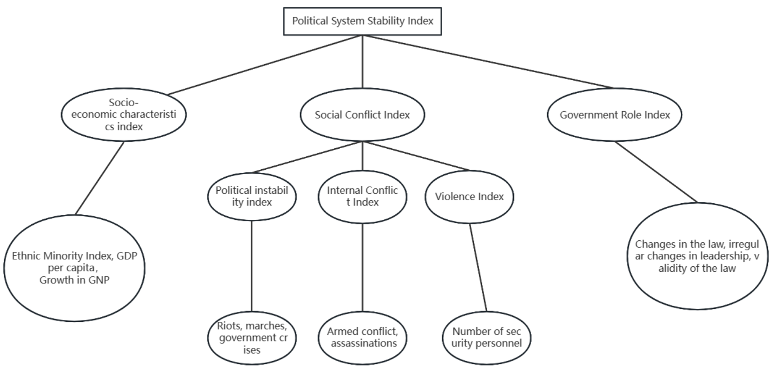


Figure 1: Political System Stability Index Framework

The public policy social stability risk assessment index system includes internal policy factors and external environmental factors. The early warning level model of social security and stability in China contains public safety, social economy, natural disasters, public health, major accidents, major events, and the surrounding environment. The theoretical framework and indicator system of social stability risk assessment in China from the perspective of social-ecological system governance contains environment, mentality, behaviour and governance. This paper combines the above models, measurement systems and the collected representative authoritative data sets to construct the social stability assessment index system as shown in the following table.

Table 2: Social stability indicator system

Level 1 indicators	Level 2 indicators
Political system	GDP per capita
	Investment in fixed assets
	GNP
	CPI
	PPI
Social environment	Population
	Net increase in population
	Population density
	Number of people educated
	Jobless rate
	Engel's coefficient
	Happiness index
Public safety	Administrative Litigation
	International murder crime rate
Natural environment	Land area
	Forest area
	Area of agricultural land
	Agricultural added value
	CO_2 emissions

From the table, we can learn that this study divides social stability into four dimensions: political system, social environment, public safety and natural environment, where political system contains five secondary indicators: GDP per capita, Investment in fixed assets, GNP, CPI, PPI. The social environment consists of seven secondary indicators: Population, Net increase in population, Population density, Number of people educated, Jobless rate, Engel's coefficient and Happiness index [9]. The secondary indicators are Public safety includes the Administrative Litiga-tion and International murder crime rate. Natural environment includes five secondary indicators: Land area, Forest area, Area of agricultural land, Agricultural added value and CO2 emissions.

Quantitatively, studies have been conducted to assign values to social stability indicators through questionnaires, using a combination of expert and substantive leader consultation and fieldwork methods. Nature is generally considered to be the environment on which human beings depend, with a mean score of 96% for necessity, 80% for importance and 76% for feasibility. In terms of the social environment, the mean score was 91.7% for necessity, 62.5% for importance and 48% for feasibility. In terms of the political system, 96% of the scores were for necessity, 79.2% for importance and 70.8% for feasibility. In terms of public security, 91.3% on necessity, 85.7% on importance and 80.3% on feasibility [10]. The four level indicators have a mean value of 93.75% in terms of necessity, 76.85% in terms of importance and 68.775% in terms of feasibility. Overall, there was greater agreement and higher scores for these four dimensions. Based on the existing research results, the rationality and scientificity of this constructed social stability indicator system is fully affirmed [11].

From a qualitative point of view, the rational formulation of a political system is a decisive factor in the stability of society. Long-term political stability is the basic prerequisite for achieving social stability, and with a stable domestic political environment as a political prerequisite, the cause of modernisation can continue to move forward. Human beings are social animals, and the development of society cannot be achieved without human beings, therefore social stability cannot be achieved without human participation. In today's rapidly developing economy, economic development is an important factor affecting social security and stability, and when people are well-off and happy, society is naturally safe. Factors that affect social security and stability include the situation of unemployment, population density and people's happiness index. Public safety involves the surrounding environment of people, which directly determines the safety of residents' lives and property as well as their inner sense of security, and directly determines whether society is stable or not. Factors affecting social security and stability include the number of criminal cases filed, terrorist attacks, etc. The natural environment is the environment on which human beings depend, and therefore a harsh natural environment is a sudden factor that affects social security and stability. When destructive natural disasters occur, they can cause irreparable and huge losses to people's lives and properties, generate instability among the people and epidemic epidemics, and affect social security and stability [12].

In summary, combining the available research results and data documents, the analysis in terms of quantification and implementation, the system of social stability indicators studied therefore contains four aspects: political system, social environment, public safety and natural environment.

4.1.4 Correlation analysis of indicator systems

To analyse the correlation between political system, social environment, public safety and natural environment, the total scores of political system, social environment, public safety and natural environment and the total scores of each dimension were first calculated, and then Pearson correlation analysis was used to explore the correlation between the four indicators. As the data collected has different units, bins and scales, it is necessary to first standardise the data, which is commonly done by normalisation, averaging and intervalisation, etc. In this case, Z-Score standardisation is a common method of data processing, through which data of different scales can be converted into a uniform measure of Z-Score scores for The formula is as follows It is calculated using the following formula:

$$z = \frac{x - \mu}{\sigma} \tag{3}$$

Where x is the sample value, μ is the sample mean and σ is the standard deviation.

After processing the data results, a matrix scatter plot was drawn using SPSS to test whether there was a downward inverse correlation between the indicators and the results are shown in the figure:

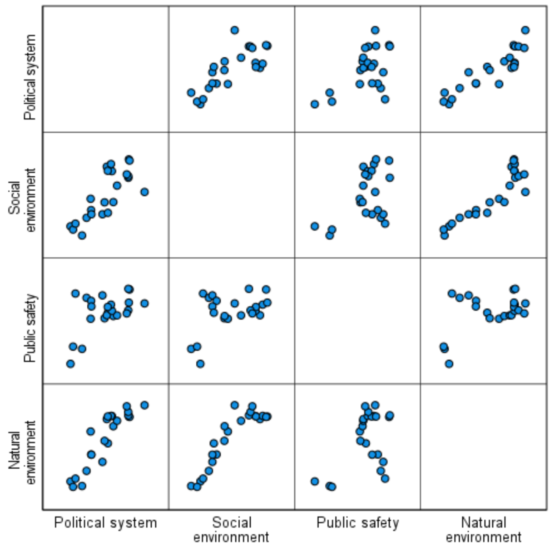


Figure 2: Matrix scatter plot

As can be seen from Figure 2, there is a significant linear correlation between the four indicators, between the political system and public safety and the natural environment, between the social environment and the natural environment and public safety, between public safety and the political system, and between the natural environment and the social environment, for which a Pearson correlation analysis was carried out to obtain the results shown in the table below:

Table 3: Results of correlation analysis

	Public safety	Social environment	Political system	Natural environment
Public safety	1(0.000***)	0.477(0.021**)	0.528(0.010***)	0.476(0.022**)
Social environment	0.477(0.021**)	1(0.000***)	0.779(0.000***)	0.906(0.000***)
Political system	0.528(0.010***)	0.779(0.000***)	1(0.000***)	0.902(0.000***)
Natural environment	0.476(0.022**)	0.906(0.000***)	0.902(0.000***)	1(0.000***)

Note: ***, **, * represent 1%, 5%, 10% level of significance respectively

As can be seen from Table 3, the relationship between the four indicators is significant and has a strong positive correlation, with the overall correlation ranging from 0.476 to 0.906. With the exception of public safety and social environment and public safety and natural environment, the significance level between all other indicators is above 1%. This means that the four indicators influence each other, and fluctuations in one indicator will drive fluctuations in the other three indicators.

4.1.5 Cause and effect analysis of the indicator system

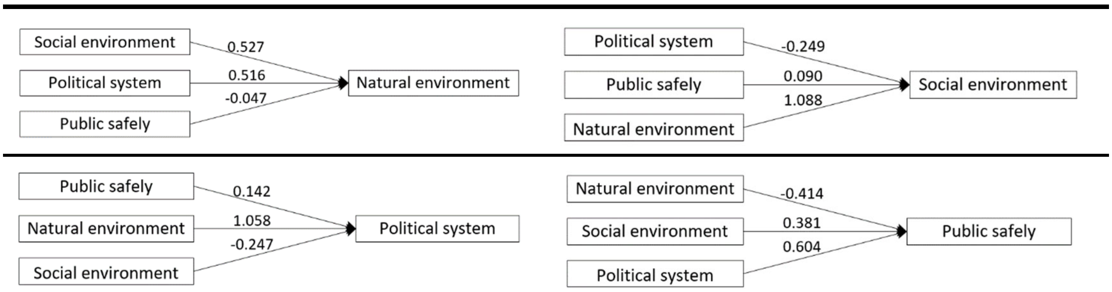
To further understand the causal relationship between the four indicators of political system, social environment, public safety and natural environment, a stepwise regression analysis was conducted based on correlation analysis with natural environment as the dependent variable and political system, social environment and public safety, which are significantly and positively correlated with natural environment, as independent variables, and the results are shown in Table 4. Due to the problem of space in the thesis, the natural environment is used as an example for detailed explanatory notes, and only the model paths are otherwise plotted.

Table 4: Results of regression analysis

	Non-normalized coefficients		Normalization factor	t	P	VIF	R^2	Adjust R^2	F
	B	Standard error	Beta						
Constant	0	0.025	-	0	1.000	-	0.92	0.908	F=72.988 P=0.000***
Social environment	0.623	0.123	0.527	5.05	0.000***	2.587			
Political system	0.293	0.061	0.516	4.783	0.000***	2.771			
Public safety	-0.035	0.057	-0.047	-0.617	0.545	1.408			
Dependent variable：Natural environment									
Note: ***, **, * represent 1%, 5%, 10% level of significance respectively									

As can be seen from the table above, the analysis of the results of the F-test yields a significance p-value of 0.000***, presenting significance at the level and rejecting the original hypothesis of a regression coefficient of 0. The model therefore basically meets the requirements. For the performance of variable co-linearity, the VIF is all less than 10, so the model has no problem of multiple co-linearity and the model is well constructed. R^2 is 0.92 close to 1, the fit is good. In terms of the regression coefficients for each dimension, the P-value for political system in the scale is less than 0.05, indicating an explanatory effect on the natural environment. The social environment p-value is less than 0.05, indicating an explanatory effect on the natural environment. The p-value for public safety is insignificant and does not have a strong effect on the interpretation of the social environment. The formula for the model is as follows: Natural environment = 0.0 + 0.623*Social environment + 0.293*Political system - 0.035*Public safety. the other three indicators in turn are used as dependent variables to plot the summary table of the model path diagram as shown in the table below.

Table 5: Summary of model path diagrams



From the above table, it can be seen that, taking the natural environment as an example, the social environment, political system and public security have a significant impact on the natural environment and there is a causal relationship. β coefficients of 0.527 and 0.516 are positive, indicating that the social environment and political system have a significant positive impact on the natural environment, because changes in the social environment and political system will positively affect the natural environment. The beta coefficient of public safety has fewer negative coefficients and the causal inertia relationship is not strong enough. In addition, there is a causal relationship between the natural environment and the social environment, a causal relationship between the natural environment and the political system, and a causal relationship between the social environment and the political system and public safety, a result that is consistent with our real life. The above results provide a basis for improving instability and enhancing social stability.

4.2 Solution of question 2

In order to improve the generalization ability of the model and the reliability of the results, this paper establishes a support vector machine (SVM) based on the processed data attachment, selects the warning level as the prediction target and the corresponding other indicators as the feature indicators. Firstly, 85% of the total data are randomly selected as the training set and 15% as the test set. Using the grid search method and the built-in parameter optimisation command of MATLAB, the parameters of the SVM model are optimised and the model loss is calculated respectively, and the parameter corresponding to the smaller loss is the final parameter setting; finally, the model results are analysed and discussed based on the evaluation index system.

4.2.1 Early Warning Level

Early Warning Level is Social Security and Stability Indicator Levels. Based on the current situation and its development trend, the city's social security and stability is divided into three levels: light police normal, medium police alert and heavy police dangerous, and is expressed in orange, grey and yellow, as shown in the table：

Table6: Classification of the police situation

Alert Level	Description
Orange level social security and stability situation has seen a light alert	Low likelihood of recent incidents affecting social security and stability
Moderate alert for grey level social security and stability conditions	There is a general sense of insecurity but confidence in government, factors in society that threaten security and stability, and the possibility of recent events affecting security and stability
Yellow level social security and stability is clearly bad, heavy special police situation risk level	People feel threatened by social instability, doubts about the government's ability to control them, rumours of instability are beginning to emerge, discontent is already building in the community, and society is in a state of mild unrest that could lead to large scale mass incidents if the situation continues

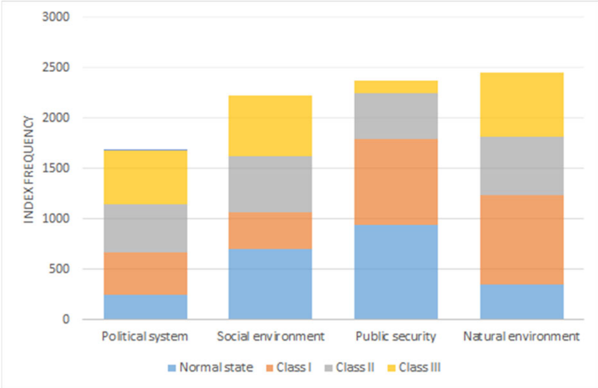


Figure 3: Frequency diagram of specific class divisions

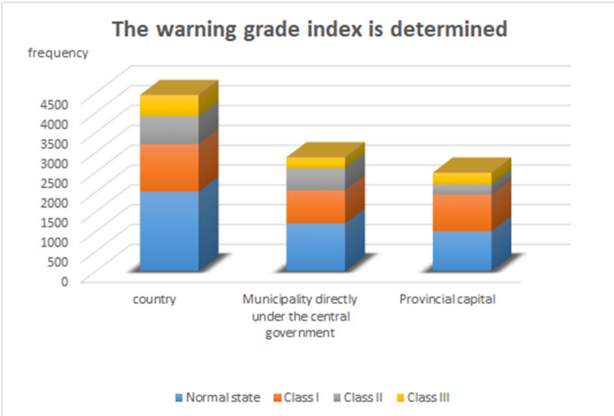


Figure 4: Specific grade determination diagram

The principle of early warning level determination. The warning level is actually a reflection of the relationship between the total population of the city, the relevant indicators, the state of the police and the number of occurrences, and the cyclical pattern of occurrence. The formula is:

$$FWD = \sum_{j=1}^{n} I_j W_j \tag{4}$$

Where FWD stands for warning level, I for indicator and W for weighting.

4.2.2 Introduction to SVM

Support Vector Machine (SVM) is a supervised learning algorithm. It maps low-dimensional sample data to a high-dimensional space by means of a kernel function, which in turn converts non-linearity into linearity, and maximises the distance between samples of different classes by finding the optimal hyperplane that separates as much data as possible correctly.

There are two types of support vector machines, linear and non-linear. In real life, for the classification of large amounts of industrial data, non-linear support vector machines are often used. A non-linear SVM is obtained by applying a linear SVM after mapping the input data to a high-dimensional space using a non-linear function. Usually SVMs which have the following characteristics.

(1) It can be represented as a convex optimisation problem, so that the global minimum of the objective function can be found using known efficient algorithms.

(2) The ability to control the model by maximising the edges of the decision boundary.

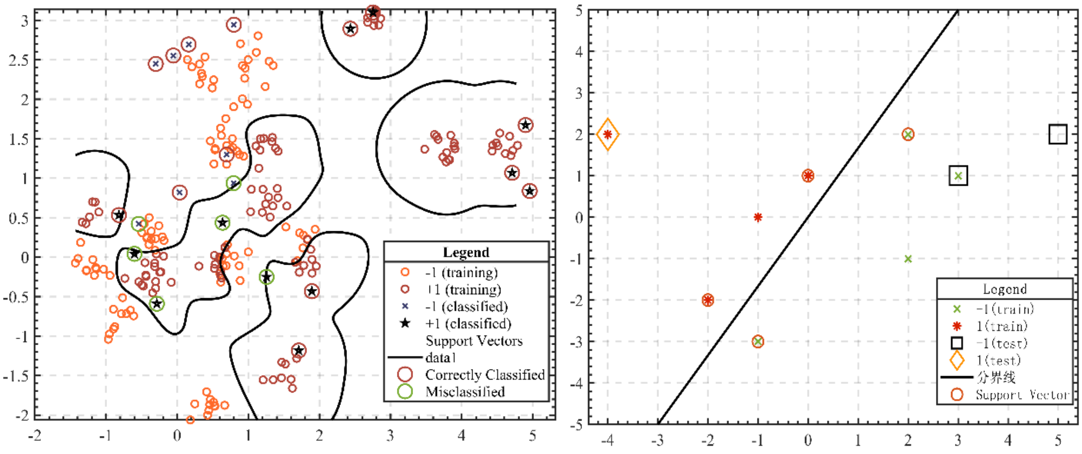


Figure 5: Left: Nonlinear SVM Right: linear SVM

4.2.3 SVM principle

$$min(w,b) = \frac{1}{2}\left\|w\right\|^2 + C\sum_{i=1}^{l}|y_i - f(x_i)| \tag{5}$$

Where: w is the weight vector of the sample; x_i is the ith input vector of the sample; b is the fit bias of the sample; y_i is the actual vertical coordinate value corresponding to sample x_i; and C denotes the penalty factor of the optimisation problem, which keeps the error of the whole training process within the normal range.

After adding the relaxation variables ξ and ξ*, the Lagrange multipliers a_i and a_i^* are introduced subject to the constraints, β_i and β_i^*≥0, i =1,2..., l. The objective function is:

$$L = \frac{1}{2}\left\|w\right\|^2 + C\sum_{i=1}^{l}(\xi + \xi^*) - \sum_{i=1}^{l} a_i((w\,x_i + b) - y_i - \varepsilon - \xi_i) -$$

$$\sum_{i=1}^{l} a_i^*(y_i(w\,x_i + b) - \varepsilon - \xi_i^*) - \sum_{i=1}^{l}\beta_i\xi_i - \sum_{i=1}^{l}\beta_i^*\xi_i^* \tag{6}$$

Derivation of equation (6) leads to the decoupling of the solution problem as follows.

$$\begin{cases} max\sum_{i=1}^{n} y_i(a - a_i^*) - \varepsilon\sum_{i=1}^{l}(a_i \text{ 和 } a_i^*) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a - a_i^*)(a - a_j^*)K(x - x_i^*) \\ s.t\sum_{i=1}^{l}(a - a_i^*) \le a_i\text{，} a_i^* \le c \end{cases} \tag{7}$$

Where: a_i and a_i^* are used to represent the Lagrange multipliers added during training. β_i and β_i^* From the above equations, the SVM model function.

$$f(x) = \sum_{i=1}^{n} y_i(a - a_i^*)K(x - x_i^*) + b^* \tag{8}$$

Where: b^* is the intercept of the separation hyperplane; $K(x_i, x_j)$ denotes the kernel function of the SVM model.

4.2.4 Model building process

Step1.Data pre-processing, remove missing values, abnormal values, etc. and do normalization.

Step2. use the warning level as output features and the rest of the features as input.
Step3. set the ratio of training set to test set.
Step4. select and function.
Step5. use grid search method to find the best parameters.
Step6. test the model using the test set.

4.2.5 Introduction to model evaluation indicators

(1) Mean Square Error

MSE is the sum of the squared distances between the target variable and the predicted value:

$$MSE(\theta) = E\left(\theta - \hat{\theta}\right)^2 \tag{9}$$

where $\hat{\theta}$ is the estimated value and θ is the true value.

(2) Mean absolute error

MAE is the sum of the absolute values of the differences between the target and predictor variables

$$AE = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{10}$$

where x_i is the actual value, μ is the mean value and n is the sample size.

(3) Model loss values

The loss is calculated through training, the loss is not a percentage but an accuracy, it is the sum of the errors for each example in training.

4.2.6 Parameter settings

The classification effect of the SVM model on the samples mainly depends on the sample capacity, data quality and the reasonableness of the parameter settings. The model parameters in this paper are set as follows.

Selection of forecasting objectives and characteristic indicators

Table 7: Prediction target and characteristic index

Target	Characteristic index
Warning level	Indicators beyond the warning level

Sample data division

In fields such as machine learning and pattern recognition, samples are generally divided into three separate training, validation and test sets. The training set is used to estimate the model, the validation set is used to determine the network structure or parameters that control the complexity of the model, and the test set examines how well the final selection of the optimal model performs. The ratio column between the data sets in this paper is set as follows.

Table 8: Ratio Settings for test and training sets

SVM model	Proportion	Quantity
Training set	80%	387
Test set	20%	73

(3) Introduction to SVM parameters

Table 9:Partial parameters of SVM model

Parameter name	Description	Set

Kernel function type	0—linearity 1—polynomial 2—RBF function 3—sigmoid	Default is 2
Change range of penalty parameter c	Penalty parameter c	Range of parameter c $[2^{-8}, 2^{8}]$
-g r(gama)	The gamma function setting in the kernel function	The reciprocal of the number of default categories
-r coef0	The coef0 setting in the kernel	Default 0
-s svm	SVM setting type	Default 0

(4) Hyperparameter optimisation

For this paper, using the grid search method and the built-in optimization commands of MATLAB, the parameters c and g are searched for respectively, and finally the optimal parameter is selected according to the evaluation index of the model.

4.2.7 Model training results

(1) Grid search method

The grid search was implemented through MATLAB programming and the results are shown in Figure.

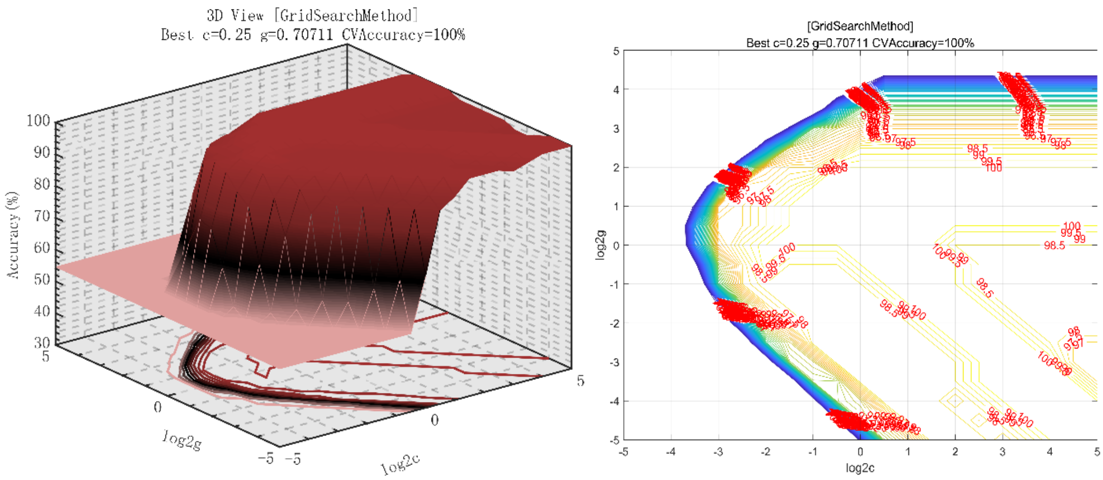


Figure 6: Grid search results

From the above graph, the optimal parameters c = 0.25 and g = 0.7011. In this case, the model is 100% accurate, with a model mean square error of 0.3978 and a mean absolute error of 0.5180.

(2) MATLAB built-in function optimization

Using automatic hyperparameter optimisation via the built-in MATLAB function, hyperparameters were found that minimised the 5-fold cross-validation loss. For reproducibility, random seeds were set and the "expectedimprovement -plus" acquisition function was used. The results are shown in the figure.

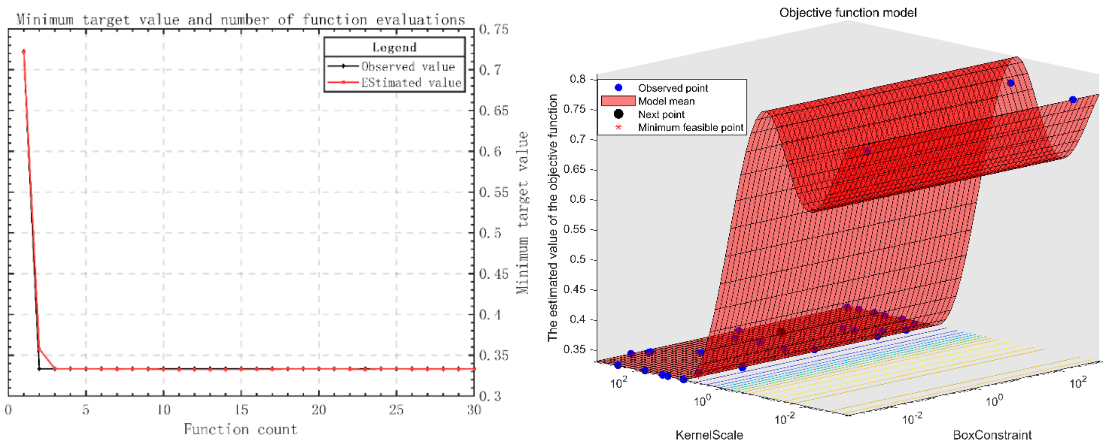


Figure 7: Optimization process diagram

After optimisation by MATLAB's built-in optimiser, the model was 100% accurate, with a model mean square error of 0.648 and an average absolute error of 0.4918.

With the above analysis, we build the SVM model based on the grid search method.

4.3 Solution of question 3

In response to question three, essentially a forecast of national or regional stability is made using the model that has been developed. Based on the data from Belarus, the results are obtained by entering the corresponding indicator data for that country into the model. A grey evaluation system is then established, with the early warning system as the parent series and the rest of the indicators as sub-series, and a grey correlation analysis is carried out to obtain the correlation between each sub-series and the parent series, which leads to recommendations for social stability.

4.3.1 Analysis of predicted results

The data from Belarus was selected and entered into the model, and the result was an average level of social stability.

4.3.2 Grey correlation analysis

(1) Selection of evaluation systems

A measure of the magnitude of the correlation between two systems of factors that change over time or with different objects is called the degree of correlation. In the development of a system, if the two factors change in tandem to a greater extent, the degree of association is higher; conversely, it is lower. The basic idea of grey correlation analysis is therefore to determine the precision of the linkage of a series based on the similarity of its geometry; the more similar the curves, the greater the correlation between the corresponding series.

For this question, as the data is time-series data, we can establish a grey correlation analysis evaluation system to calculate the degree of correlation between each indicator and GDP, so as to compare the most influential indicators.

(2) Model building

Step.1 Determine the parent and child sequences.

Step.2 Pre-process the data.

Step.3 Calculate the grey correlation coefficient.

Step.4 Calculate the grey correlation, that is, calculate the average of grey correlation coefficient of each series.

Step.5 Compare the grey correlation, the higher the grey correlation the deeper the effect, and vice versa, the shallower the effect.

Model solving

The correlation coefficients were calculated by Matlab and the results are as follows:

Table 10: Calculation results

Correlation coefficient						
Indicators in columns 1 through 7						
0.9843	0.6205	0.9997	1.0000	1.0000	0.9998	0.9979
Indicators in columns 8 through 14						
0.9996	0.9984	1.0000	0.9999	1.0000	0.9844	0.7624
Indicators in columns 15 through 20						
1.0000	0.9973	0.9994	0.9985	1.0000	0.9879	

From the above table, it can be seen that the cpi, PPI, unemployment rate, population growth rate, international murder crime rate and agricultural value added have a strong link with the colour revolution. The cpi index reflects the spending power of the population, while the ppi index captures the price of industrial production.

4.3.3 Recommendations for maintaining stability

Based on the data collected, we propose the following recommendations to stabilise society:

(1) The government should always pay attention to the trend of CPI and PPI, and keep the CPI index up and the PPI index down as much as possible, which means that industrial enterprises' production profits will increase and the economy will enter the next expansion period

(2) Control the speed and scale of population movement between urban and rural areas. With too much movement of population between urban and rural areas, there will be too many workers and the supply will exceed the demand, so the natural unemployment rate will increase

(3) Control the population growth rate and improve the quality of national education

(4) In terms of agricultural output, strengthening public infrastructure, accelerating the intelligent transformation of infrastructure such as farmland water conservancy, cold chain logistics and agricultural production and processing in rural areas, and mobilising agricultural enterprises to invest in agricultural production

(5) Building an agricultural big data service platform so that more agricultural enterprises can share the dividends of big data development, understand the market, monitor and warn through big data, and improve the competitiveness of the agricultural industry.

4.4 Solution of question 4

4.4.1Ukrainian background of the Orange Revolution

The Orange Revolution refers to a series of protests and political events in the 2004-2005 period surrounding the 2004 Ukrainian presidential election due to serious corruption, voter influence and direct electoral fraud. After these protests, the Supreme Court of Ukraine declared the results of the reelection invalid and provided for repeated reelections on December 26 of the same year. This second reelection was closely observed. The second re-selection was basically no problem. Yushchenko clearly won the reelection with 52% results.

Mr.Yanukovych won 44%. On January 23,2005, he marked the final victory of the Orange Revolution, which led to the decline of the regime.

4.4.2 Good and inferior solution distance method (TOPSIS) model

(1) Analytical procedure

Prepare the data, and conduct the same trend processing and dimensional problems.

To confirm the weight of each index, the entropy weight method and custom weight can be used (for self-processing, quantitative-AHP can be used).

Find out the optimal and worst matrix vectors (system processing).

Calculate the distance between the evaluation object and the positive ideal solution D + or the negative ideal solution D-respectively.

The comprehensive degree score C value is calculated by combining with the distance value, and the conclusion is sorted.

(2) Indicator weight calculation

Table 11: Indicator weight

Entropy Weight Method			
Nape	Information entropy value	Information utility valued	Weight (%)
Population (100 million)	0.007	0.993	4.979
CPI(%)	0.005	0.995	4.987
Population density (person/km2)	0.001	0.999	5.011
GNP (trillion dollars)	0	1	5.014
GDP per capita (USD)	0.001	0.999	5.006
Investment in fixed assets (100 million yuan)	0	1	5.014

Conclusion: Population density (person / km2), GNP (trillion dollars), Administrative litigation, Housing price (yuan), Unemployment rate (%). 1, Engel coefficient (%), the proportion and weight of the orange level of Ukrainian social warning is relatively large, and there is a certain influence between them.

(3) Calculation results of TOPSIS evaluation method

Table 12: TOPSIS evaluation results

Index Value	Positive Ideal Solution Distance (D_+)	Negative Ideal Solution Distance (D_-)	Comprehensive Score Index	Sort
Class I	0.54737339	0.83636735	0.60442489	1
Class II	0.83636735	0.54737339	0.39557511	2

Comprehensive degree score C value, $C = \frac{D_-}{D_+ + D_-}$, in the calculation formula, the molecule is D-value and the denominator is the sum of D_+ and D_-; the larger the D_-value, the farther from the worst solution, the better the study object; the greater the C value, the better the study object.

4.4.3 Gray association analysis model

Table 13: Correlation results

Evaluation item	Correlation degree	Ranking
Population growth rate	1	1
International murder rate (per 100,000 people)	0.987	2

Agricultural land area (10,000 square kilometers)	0.934	3
Administrative litigation	0.902	4
CO_2 emissions (tonnes per capita)	0.813	5
year	0.766	6
GDP per capita (USD)	0.766	7
CPI(%)	0.75	8
Population density (person/km2)	0.74	9
Agricultural value added (%)	0.736	10

Firstly, by the distance method (TOPSIS) model, the better the research index; the greater the C value, the better the research index. Based on the above correlation coefficient results, the correlation degree value is obtained. The correlation degree value is used to rank the 21 evaluation objects. The correlation degree value is between 0 and 1, the stronger the correlation between the "reference value" (parent sequence), the higher the evaluation. As can be seen from the above table: for the 21 evaluation items, Population growth rate has the highest evaluation (correlation:: 1.0), followed by International murder rate (per 100,000 people) (correlation: 0.987), and the correlation between these indicators and the warning level is stronger.

4.4.4 The main reasons for regime change

The color revolution was a street movement caused by a series of internal and external factors, due to the weak national strength of Ukraine. The internal factor is that the Ukrainian people, especially those in the central and western regions, have always wanted to integrate into Europe as soon as possible. The external factor is that some countries take advantage of this desire to push a wedge between Ukraine and Russia in an attempt to break their separate ways. This situation between the two mountains, the opposition has released the air before the event, Yushchenko lost the presidential election, they will organize national protests and purposeful activities.

4.5 Solution of question 5

Taking into account the findings of this study, the following recommendations are made in order to prevent colour revolutions and maintain social stability.

From the national level, everything should be people centred. Safeguarding the fundamental interests of the general public, protecting people's lives and property, building a harmonious social environment, reducing social cost expenditures and comprehensively promoting the process of social and economic construction are the starting and ending points of early warning work. Social development should adhere to the people-oriented principle and be guided by the scientific concept of development. First, a correct view of purpose should be established. In social stability risk assessment, the most important criterion is public opinion, and the interests of the masses should always be put in the highest position, so as to build a three-dimensional interactive bridge between the decision-making level and the general public by collecting public opinion, pooling the wisdom of the people and gathering the hearts of the people, so as to promote the democratization and scientification of decision-making, enhance policy transparency and boost the credibility of the government. Taking the study of Ukraine in section 4.4 as an example, the Chestnut Orange Revolution in Ukraine showed that social stability relies on the people, and observers in Ukraine and abroad reported that the officially declared result of Yanukovych's victory was the result of fraud, which caused the government to lose its credibility degree and thus also led to the defeat of the election, marking the final victory of the Orange Revolution, which led to the fall of the regime. Secondly, a scientific political system should be established, suggesting that the state should be governed by law, administered, established, built and implemented in accordance with the law, with administrative accountability to monitor policy implementation. The study of Belarus in section 4.3 found that a cohesive country with no tolerance for foreign interference in politics is also more socially stable compared to Ukraine. Third, we must actively innovate mechanisms to build norms and keep up with the trend of the information age. We should set up a pool of experts, listen to the opinions and suggestions of all parties, make scientific analyses and forecasts, put forward practical proposals and come up with assessments that will stand up to scrutiny. Establish a social stability evaluation system that keeps up with the times, and the Office of Stability Maintenance is specifically responsible for the organisation, command, coordination, implementation, supervision and assessment of the evaluation work, ensuring that the evaluation work is built with strong organisational safeguards from top to bottom. Establish a mechanism for resolving conflicts, actively participate in diplomacy and establish friendly and harmonious diplomatic relations.

For individuals, the hearts and minds of the people will be united and the joint efforts of the masses will make society more stable. We must build up a sense of national identity, national pride and self-confidence, study science and culture, improve our own quality, strive to maintain national security and social stability, and fight against all phenomena that harm national interests. We must do what is conducive to the common progress of all mankind, and not do what is done under the guise of doing it for the sake of all mankind. Individuals should always develop the idea that the national interest is above all else, a necessity for the national interest and a need for personal security, and a unanimous demand of all countries in the world. At the same time, we should actively care for the natural environment, which is the home on which mankind depends for its survival. As individuals, protecting the environment is mainly in our daily lives, the main thing is to actively practise the concept of low carbon and green living. We should develop the habit of separating waste, actively promote the importance of protecting the environment and raise citizens' awareness of environmental protection.

5 MODEL EVALUATION

5.1 Strengths

Support vector machine (SVM) Advantages:

1) High classification accuracy by introducing maximum interval.

2) Accurate classification even when the sample size is small and good generalization ability.

3) The introduction of kernel functions makes it easy to solve non-linear problems.

4) It can solve the classification and regression problems of high-dimensional features, even if the feature dimension is larger than the sample data, it can also have good performance.

5) The final decision function is determined by only a small number of support vectors, the complexity of the calculation depends on the number of support vectors, not the dimensionality of the sample space, avoiding the "dimensional disaster".

6) A small number of support vectors determines the final result, which not only helps us to catch the key samples, but also to "eliminate" a large number of redundant samples.

7) Good "robustness", adding or removing non-support vector samples has no effect on the model.

Advantages of grey correlation analysis:

It is equally applicable to the size of the sample, or the sample size with or without a pattern, and is computationally small and convenient, and there is no discrepancy between the quantitative and qualitative analysis results.

5.2 Weaknesses

SVMs are difficult to implement for large training samples.
SVM is sensitive to missing data.

5.3 Improvements

Build a random forest model solution, which balances errors for unbalanced classification datasets.
Find more suitable algorithms to optimize SVMs, such as particle swarm algorithms.
Build more complex non-linear regression models that better represent the relationships between indicators.

REFERENCES

[1] Haendel, D., Meadow, R. G., & West, G. T. (1975). Overseas investment and political risk. Foreign Policy Research Institute.
[2] Wu, X. (2012). Research on political risk of Chinese enterprises' investment in developed countries in the post-financial crisis era (Master's thesis). Hubei University of Technology.
[3] Zhao, C. (2022). Understanding the causes of neighborhood avoidance incidents and suggestions for solutions from a socialization perspective. China Emergency Management Science, 2022(10), 83–92.
[4] Zhong, Z., Zhang, H., & Kong, X. (2021). How to make social stability risk assessment of major decisions more scientific: An empirical analysis based on the use of social survey methods. Journal of Zhongnan University for Nationalities (Humanities and Social Sciences Edition), 41(07), 128–140.
[5] Sun, X. (2022). Research on social stability risk assessment system for major resource and environmental projects. Sichuan Environment, 41(06), 299–304.
[6] Hou, A. (2022, April 28). Economic indicators continue to recover, social situation harmonious and stable. Zhengzhou Daily, p. 010.
[7] Qu, Q., & Du, J. (2022). Cognition and prevention of social risks: A typological theoretical investigation. Journal of the Party School of the Hangzhou Municipal Committee of the Communist Party of China, 2022(05), 67–77.
[8] Tang, Y. (2021). Study on social stability risk assessment of waste separation complex project (Master's thesis). Zhejiang University.
[9] Feng, X., & Jane, Z. (2023). An analysis of long-term social stability experience in Chinese modernization: A united front perspective. United Front Studies, 7(03), 25–35.
[10] Yu, Z. (2023). Playing a model and leading role of people's insurance in maintaining national security and social stability. China Insurance, 2023(05), 5–7.
[11] Liu, S., & Chen, X. (2023). Research on the evaluation of students' scientific reasoning ability based on the hierarchical analysis method: A case study of high school biology. Education and Teaching Research, 37(07), 44–58.
[12] Yan, J. (2023). From economic development to social governance: A demand-response-based analysis of social stability in China. Governance Studies, 39(02), 4–19, 2, 157.