International Journal for Housing Science and Its Applications
Publish August 3, 2025. Volume 46, Issue 3 Pages 855-867

https://doi.org/10.70517/ijhsa46364

An intelligent machine learning-based approach: Enhancing the sustainability of rural cultural pension systems

Qinjian Yang[1,2,*] and Zepu Wang[2]
[1] College of Arts and Design, Ningbo University of Finance & Economics, Ningbo, Zhejiang, 315000, China
[2] Ningbo University of Finance & Economics, Ningbo, Zhejiang, 315000, China
Corresponding authors: (e-mail: 15168582465@163.com).

Abstract The rural social pension replacement rate reflects the level of social pension protection for individuals by the rural social pension insurance pillar, and is a relative indicator capable of measuring the level of pension protection by inter-period comparison under different levels of economic development. Aiming at enhancing the sustainable development strategy of rural cultural pension system, this paper establishes a sustainable development index system, uses Random Forest, Logistic Regression and Bayes as the base classifiers, proposes an integrated learning model based on the Stacking integration strategy, and Support Vector Machines as the meta-classifiers for parameter optimization, to assess the sustainable development of the pension system. Secondly, SHAP is introduced to explain the influencing factors of pension system sustainability found by the model. According to the empirical analysis and visualization analysis, the Stacking integration method is better than other single learner models in terms of accuracy, checking rate, AUC value, etc., except for the true instance rate which is lower than random forest, with values of 0.953, 0.877, 0.972, 0.949, respectively, indicating that Stacking integration can be better applied to the pension system sustainability in the assessment. Integrating the influencing factors, it is found that the development speed of the system, the economic affordability at all levels, the reasonableness of the institutional setup, and the efficiency of the management service are the main influencing factors for the sustainable development of the pension system. Finally, recommendations are made to optimize the demographic structure, improve the quality of the population, and active aging in order to cope with the risk of population aging and unsustainability of rural pensions.

Index Terms integrated learning, SHAP values, sustainability, rural cultural pensions

I. Introduction

With the deepening of China's aging, the number of elderly population continues to increase, the problem of old age has become a social focus of the problem [1]. 2020 seventh national census data show that: China's population aged 60 years old and above there are 260 million people, accounting for 18.7%, and the aging of the rural areas is even more serious, the elderly population accounted for the proportion of the total population has reached 23.8% [2]. From the point of view of the incidence of poverty in old age, the elderly entrance in rural areas due to unstable sources of income lack of livelihood security, the incidence of poverty, close to 10%, more than three times higher than the urban [3]. Compared with urban areas, the elderly population in rural China has more serious problems. From the reality, in rural areas of China, many elderly people, even if they are over the legal retirement age, are still doing heavy physical labor to meet their basic needs [4]-[6]. Moreover, due to the limitations of knowledge and ability, the rural elderly population is mostly engaged in agricultural labor with high physical exertion, or mechanical repetitive non-agricultural labor with a lower level of technology, which will adversely affect their physical health and life well-being index [7], [8]. Therefore, the labor supply is related to the overall welfare level of the rural elderly population, and the high intensity of labor supply implies a heavy labor burden and low welfare level, and reducing the labor burden of the rural elderly population is an important aspect of improving their overall welfare level and effectively reducing the welfare gap between urban and rural elderly.

In response to the serious problem of old-age security in rural areas, the Government has also been exploring ways to socialize the issue of old-age security in rural areas. After continuous exploration and practice, the Chinese government began to implement the pilot work of the new rural social pension insurance (referred to as “new rural insurance”) in 2009, and the implementation of the new rural insurance has been very smooth, and by the end of 2012, full coverage was realized throughout China [9]. In 2014, the government decided to merge the implementation of new rural social pension insurance and urban residents' social pension insurance, establishing a unified social pension insurance system for urban and rural residents, breaking the urban-rural dichotomy to a certain extent [10]. As a result, the rural elderly population has been incorporated into a relatively complete social pension system, and there is basic protection for their old age.

The new rural old-age insurance has rapidly reached a participation rate of more than 90% in just a few years after its initial implementation and has maintained a high level of participation, filling the gap in the rural old-age insurance system and providing basic livelihood protection for a large number of rural elderly population [11], [12]. Developing to today, a large number of studies related to the NPS have emerged, exploring the impacts of the NPS since its implementation on rural households' income, consumption, poverty, labor supply, health status, subjective welfare, etc., as well as the spillover effects of the NPS through old-age pension models, political trust, etc. [13]-[15].

The study establishes the measurement indexes of the sustainable development of the new rural cultural pension system from the aspects of finance, management, and effect, and uses random forest, logistic regression, and Bayes as the base classifiers. An integrated learning model based on Stacking integration strategy is constructed to evaluate the sustainable development of the pension system. Comparative evaluation is made from the metrics of accuracy rate, true case rate, checking accuracy rate and AUC value to comprehensively compare the performance of the model on all the metrics. By introducing the SHAP explanatory model, the identification of factors affecting the sustainability of the pension system is carried out by using SHAP values and feature importance ranking, and then the sustainable development strategy to enhance the rural cultural pension system is proposed.

II. Sustainability assessment model for rural cultural pension systems

II. A. Indicators for assessing the sustainability of pension systems

The indicators start from the main body of the NPS, i.e., the system itself, and scientifically analyze the specific data of the system itself, including finance, management, effects and other aspects, so as to judge whether it has the ability of sustainable development based on the specific results of the data analysis, and then find out its constraints as shown in Table 1. The method of data analysis can more accurately and scientifically reflect the problems that exist in the sustainable development of the pension system.

Table 1: Measures for sustainable development of pension systems

Primary indicator	Secondary indicator
Stability of system	System development speed
The guarantee of money raising	Economic affordability at all levels
Normative of organizational management	Rationality of institutional setting
	The efficiency of management services
	The integrity of the supervisory mechanism
Significance of effect	Insurance beneficiary
	The level of satisfaction of life
The fairness of the system	Intergenerational fairness

II. B. Sustainability assessment model based on Stacking integration

II. B. 1) A single machine learning model

(1) Logistic regression

The logistic regression model uses an intermediate conversion function in mapping the output values of the linear regression model to probability values. The function of this function is to convert continuous output values to probability values between 0 and 1. These probability values are used as the basis for classification, and when the probability value is greater than a preset threshold, the model output is 1, otherwise the output is 0. The process is as follows:

For the function:

$$y = \theta x \tag{1}$$

where $\theta = (\theta_0, \theta_1, \theta_2, \cdots, \theta_n)^T$ and $x = (x_0, x_1, x_2, \cdots, x_n)$. A Sigmoid function is introduced:

$$\sigma(t) = \frac{1}{1+e^{-t}} \tag{2}$$

The Sigmoid function is utilized to map the output y value of the linear function to a value between 0 and 1, i.e.:

$$p = \sigma(x_b\theta) = \frac{1}{1+e^{-x_b\theta}} \tag{3}$$

After transforming the output of the linear model into probability values between 0 and 1, the probability values p are then categorized 0-1 according to the determined threshold.

The problem to be considered after determining the mathematical form of the logistic regression model is to find the optimal model parameters. In statistics, the method of great likelihood estimation is a commonly used method for solving the parameters, and its basic idea is to find a set of parameters that maximizes the probability of occurrence of the observed data under this set of parameters.

$$\begin{aligned} P(Y=1 \mid x) &= p(x) \\ P(Y=0 \mid x) &= 1-p(x) \end{aligned} \tag{4}$$

The likelihood function is:

$$L(w)=\prod\left[p(x_i)\right]^{y_i}\left[1-p(x_i)\right]^{1-y_i} \tag{5}$$

Rewrite the equation in log-likelihood functional form by taking the same logarithm on both sides:

$$\begin{aligned} L(w) &= \sum\left[y_i \ln p(x_i)+(1-y_i)\ln\left(1-p(x_i)\right)\right] \\ &= \sum\left[y_i \ln\frac{p(x_i)}{1-p(x_i)}+\ln\left(1-p(x_i)\right)\right] \\ &= \sum\left[y_i(w\cdot x_i)-\ln\left(1+e^{w\cdot x_i}\right)\right] \end{aligned} \tag{6}$$

(2) Support Vector Machine

The core idea of SVM is to use a hyperplane in the feature space to distinguish between different classes of samples. This hyperplane is defined by mapping data points into a high-dimensional feature space. In this high-dimensional space, we can find an optimal hyperplane such that it separates data points of different categories and maximizes the minimum distance to get that hyperplane, i.e., maximizes the interval.

The assumption that the hyperplane is able to perfectly classify all points of both categories correctly should be satisfied:

$$\begin{cases} X_i^T W+b \geq +1, y_i=+1 \\ X_i^T W+b \leq -1, y_i=-1 \end{cases} \tag{7}$$

Calculate the distance between the two support planes there:

$$m\arg in=\rho=\frac{2}{\|W\|} \tag{8}$$

We aim to maximize the intervals, and thus minimize $\|W\|$, and by constructing a Lagrangian function, the problem of maximizing the intervals can be expressed as:

$$L(\omega,b,a)=\frac{1}{2}\|\omega\|^2-\sum_{i=1}^{n} a_i\left(y_i\left(\omega^T x_i+b\right)-1\right) \tag{9}$$

where a_i is a Lagrange multiplier and $a_i \geq 0$. The classification function is obtained by solving its dual problem:

$$f(x)=\operatorname{sgn}\left(W^T x+b\right)=\operatorname{sgn}\left(\sum_{i=1}^{n}\alpha_i y_i (x_i)^T x+b_i\right) \tag{10}$$

This also proves that only the support vectors play a role in determining the optimal hyperplane, while other data points do not.

In real classification problems, it is often not possible to find a hyperplane that can perfectly separate two categories, so we introduce the concept of “soft spacing”, i.e., we introduce a slack variable ξ_i, and solve the problem as:

$$\begin{aligned} &\min_{W,b,\xi}\frac{1}{2}\|W\|^2+C\sum_{i=1}^{n}\xi_i \\ &s.t.\ y_i\left(X_i^T W+b\right)\geq 1-\xi_i \\ &\xi_i \geq 0, i=1,2,\cdots,n \end{aligned} \tag{11}$$

where C denotes the degree of penalization of erroneous samples, $C>0$, when C is infinite that is, hard spacing.

(3) Simple Bayes

Plain Bayes algorithm as a classification algorithm main theoretical basis is the Bayesian theory and feature independence test assumptions, the conditional independence of Plain Bayes makes the model sacrifice a certain degree of accuracy, but makes the Plain Bayes method simple [16]. The specific theoretical algorithm of plain Bayes is as follows:

For the training dataset $\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$ and the output categories $c_k, k=1,2,\cdots,K$, $k=1,2,\cdots,K$. Eqs. (12) and (13) can be obtained based on Bayes' formula and conditional independence test.

$$P(X=x\mid Y=c_k)=P(X_1=x_1,X_2=x_2,\cdots,X_n=x_n\mid Y=c_k)=\prod_{j=1}^{n}P(X_j=x_j\mid Y=c_k) \tag{12}$$

$$P(Y=c_k\mid X=x)=\frac{P(X=x\mid Y=c_k)P(Y=c_k)}{\sum_{k=1}^{K}P(X=x\mid Y=c_k)P(Y=c_k)} \tag{13}$$

Substituting Eq. (12) into Eq. (13) yields the plain Bayesian classifier expression (14).

$$y=\arg\max_{c_k}\frac{P(Y=c_k)\prod_{j=1}^{n}P(X_j=x_j\mid Y=c_k)}{\sum_{k=1}^{K}P(Y=c_k)\prod_{j=1}^{n}P(X_j=x_j\mid Y=c_k)} \tag{14}$$

Again, because Eq. (14) is the same for every category in the denominator, Eq. (14) can be organized into the form of Eq. (15):

$$y=\arg\max_{c_k}P(Y=c_k)\prod_{j=1}^{n}P(X_j=x_j\mid Y=c_k) \tag{15}$$

(4) Decision Tree

Decision tree classification model is a tree structure model consisting of nodes and directed edges. Nodes can be divided into two categories: internal nodes and leaf nodes. Internal nodes often represent features or attributes and leaf nodes represent categories [17].

Decision trees begin with feature selection, finding features with classification ability improves the efficiency of the decision tree, in general using information gain or information gain ratio as a criterion for feature selection. Information gain refers to the extent to which the uncertainty of the class information is reduced when the feature information is known. Assuming that the feature is A, the information gain information expression for feature A over set D is:

$$g(D,A)=H(D)-H(D\mid A) \tag{16}$$

where $H(D)$ is the empirical entropy and $H(D)$ is the empirical conditional entropy. However, the division of features by information gain has the problem of tendency to select features with more values. At this point the choice to use the information gain ratio can compensate for this deficiency. The expression of information gain ratio is shown in the following equation:

$$g_R(D,A)=\frac{g(D,A)}{H_A(D)} \tag{17}$$

Among them:

$$H_A(D)=-\sum_{i=1}^{n}\frac{|D_i|}{|D|}\log_2\frac{|D_i|}{|D|} \tag{18}$$

There are three more commonly used algorithms for decision tree modeling, which are ID3 algorithm, C4.5 algorithm and CART algorithm, where C4.5 algorithm is an improvement on ID3 algorithm.ID3 algorithm classifies based on the criterion of information gain, while C4.5 changes the criterion of classification to the rate of information gain, and CART can be used for regression problems, as compared to the ID3 and C4.5 models which can only classify.

II. B. 2) Stacking Integrated Learning Modeling Framework

StackTADB is an integrated learning model based on the Stacking integration strategy, which integrates three base classifiers, RF, LR, and even prime Bayes.

The input data of the model is trained by the first level classifier, and then the output of the first level classifier is used for the second level classifier. In order to avoid overfitting, the stacked cross-validation algorithm provided in the "mlxtend" package is used to implement the stacked integrated learning classification prediction model. The training and test sets are labeled as D and T, respectively. In the stacked integrated learning model, five-fold cross-validation is used to split D into five subsets of equal size: $D=\{D_1,D_2,D_3,D_4,D_5\}$, and for each $D_k(k=1,2,3,4,5)$, in this paper, one of them is rotated as the test set and the remaining four subsets are used as the training set. Then train and output the prediction results of $D_k(k=1,2,3,4,5)$ using each base classifier and output the prediction results of T. In this way, this paper finally obtains five prediction results for $D_k(k=1,2,3,4,5)$ and five prediction results for T. The five prediction results of $D_k(k=1,2,3,4,5)$ are then merged into D' and the average of the five prediction results in T is calculated as T'. After performing the above operations on the three base classifiers, this paper obtains three matrices $D_k'(k=1,2,3)$ and three matrices $T_k'(k=1,2,3)$. Then the matrix $D_k'(k=1,2,3)$ and labels are used as the training set of the second-level classifier, and the matrix $T_k'(k=1,2,3)$ and labels are used as the test set of the second-level classifier.

Then, this paper utilizes the new training set to train the secondary classifier and predicts the new test set using the trained model. Finally, the prediction result of the new test set is considered as the final result of the model.

II. B. 3) Parameter optimization

Parameter optimization is the process of tuning the parameters in an algorithm to optimize the performance of the algorithm in the field of machine learning and optimization. In the process of parameter optimization, trial-and-error methods or optimization algorithms are usually used to search for optimal parameters. Trial-and-error methods usually include manual parameter tuning, grid search, and random search, which often encounter the problem of high spatial dimensionality of parameters, resulting in time-consuming search and excessive consumption of computational resources. Optimization algorithms find the optimal parameters by defining an optimization objective function and a search algorithm. The algorithm usually needs to define the loss function or objective function, and iteratively optimize the parameters according to the change of function gradient or function value to finally reach the optimal value.

The optimal hyperparameter combination is selected among the performance of all hyperparameter combinations, and the final model is trained using the selected optimal hyperparameter combination and tested on a test set. Through this process, a model with optimal performance on the test set is obtained.

Based on the computational results of parameter optimization, the StackTADB integrated model uses three base classifiers as the first-stage learners, and their parameters are adjusted by tuning the even-vegetated Bayesian classifier's $n\_neighbors=1$, $leaf\_size=30$, $P=2$, the Random Forest classifier's $n\_estimators=300$, and the logistic regression classifier's $penalty='l2'$, and $C=1.0$, and $\max\_iter=50000$ for optimization. In the second level of the learner, the model of this study uses SVM and is integrated by adjusting the parameters $probability=True$, $gamma='scale'$, $C=1.0$, and $kernel='rbf'$.

II. C. Interpretable models for the sustainability of pension systems

II. C. 1) Characteristic importance

Although it is difficult to use feature importance for model interpretation of integrated learning algorithms, it is possible to measure the influencing factors and explore which factors are critical to model prediction, and the feature importance size represents the degree of its contribution to the model's ability to generalize prediction.

For tree-based integration algorithms, the importance of feature j is measured by the average of the importance of feature j across all trees, i.e., the lift of the loss function at the cut-off node of that feature.

$$impor\tan ce_j=\frac{\sum_{i=1}^{n}T_i^j}{n} \tag{19}$$

In practice, feature importance is widely used, and most people use feature importance for feature screening, i.e., the average gain brought by the feature in the whole algorithm process to score the feature importance, and then filter the important features after sorting. Generally speaking, the larger the value of feature importance indicates that the feature has a greater impact on the prediction results of the model, but it is not known whether it has a positive or negative effect on the target variable.

II. C. 2) SHAP Interpretation Framework

SHAP- denotes SHapley Additive ExPlanations, which originated in cooperative game theory22 and is suitable for calculating the importance of features for both black-box23 and non-black-box models. Considering that for many nonlinear models, features interact with each other, it is necessary to evaluate the importance of individual features in combination with other features and calculate the marginal effects of features, which is the central idea of the SHAP method. The basic idea is to calculate the marginal contribution of a feature when it is added to the model, and then calculate the marginal contribution of the feature when it is added to the model.

marginal contribution, and then calculate the different marginal contributions of the feature in the case of all feature sequences, and finally take the average value as the SHA Pbaseline value of the feature.The algorithm can view the contribution of each feature to the model prediction, which helps to explain the model.

For example, if the prediction model has four features X_1,X_2,X_3,X_4, there are $4!=24$ permutations, and then the marginal returns for each permutation need to be calculated. For example, the marginal returns after feature X_1 is introduced into model S is calculated as:

$$\delta(S) = v\left(S \cup \{X_1\}\right) - v(S) \tag{20}$$

Based on this formula, the marginal returns for each possible value of each feature are calculated.

The Shapley value for each eigenvalue, which is the contribution of that eigenvalue to the prediction, is obtained by weighted summation of all possible combinations of eigenvalues, and the specific formula for the SHAP value φ_i is then:

$$\varphi_i = \sum_{S \subseteq F\{i\}} \frac{|S|!\left(|F| - |S| - 1\right)!}{|F|!} \left[f_{S\cup\{i\}}(x_{S\cup\{i\}}) - f_S(x_S) \right] \tag{21}$$

Assume that x_i denotes the i th sample, $x_{i,j}$ denotes the j th feature of the i th sample, and y_i denotes the predicted value of the model for the i th sample, and set the baseline of the whole model to be y_{base}, so that the SHAP value satisfies Equation (22):

$$y_i = y_{base} + f(x_{i,1}) + f(x_{i,2}) + \cdots + f(x_{i,k}) \tag{22}$$

where $f(x_{1,2})$ represents the SHAP value, i.e. contribution value, of the 2nd feature in the 1st sample to y_1, if $f(x_{1,2})$ is positive, it means that the feature has a positive influence on the predicted value; conversely, it has a negative influence.

III. Empirical tests for sustainability assessment of rural cultural pension systems

III. A. Sustainability assessment of rural cultural pension systems

III. A. 1) Data pre-processing

Among the many indicators there will be many indicators that do not contribute to the assessment, the existence of these indicators will cause redundancy in the data, which will increase the prediction time, and the data will exist F in the outliers, nulls, etc., so the data will first be processed accordingly.

Based on the availability of data, the sustainable development of the pension system in a village in 2023 was finally selected for research, and the database credit rating AA score or above was the non-default group, and the others were the default group, because the model could not recognize the text, the "default" was marked as 1 and the "non-default" was marked as 0. Since the number of non-defaulting groups is small and the number of defaulting groups only accounts for about a quarter of the non-defaulting groups, the data is balanced using the SMOT oversampling technique, before and after balancing as shown in Table 2.

Table 2: The SMOOT balance is distributed before and after

	Default group	Non-default group
Pre balance	55	231
After balance	231	231

In the calculation of the importance score of the variables, this paper adopts the Random Forest for screening, the Random Forest measures the average accuracy through the out-of-bag error rate, and the more a certain indicator decreases, the more important it means that the indicator is more important, and the results of the calculation show that the Random Forest classification accuracy rate reaches 0.863, and the AUC reaches 0.942, which is a better assessment, and the specific importance score of each indicator is shown in Table 3.

According to the importance score, the indicator of system development speed has the highest importance score (0.232), which indicates that the Stacking integration model is very important for the effectiveness of pension system assessment, the economic burden capacity at all levels and the reasonableness of the institutional setup, the efficiency of the management service, and the soundness of the supervision mechanism.

Table 3: Index importance score

Index	Score
System development speed	0.23174
Economic affordability at all levels	0.10578
Rationality of institutional setting	0.07992
The efficiency of management services	0.06814
The integrity of the supervisory mechanism	0.05451
Insurance beneficiary	0.05244
The level of satisfaction of life	0.04775
Intergenerational fairness	0.03602

III. A. 2) Evaluation of impact analysis

Using Python to construct the Stacking model, after the data is pre-processed, the appropriate parameters are selected by constantly tuning the parameters, and then the Stacking integrated classification effect is compared with the classification effect of each base learner to compare the applicability of different machine learning models in the degree of sustainability of the pension system. Parameter selection, Random Forest the most important parameters when the number of classifiers and depth, usually the more classifiers, the higher the accuracy, but to a certain number began to stabilize, by debugging the number of classifiers 10, 30, 50, 100, 200 classifiers, found that classifiers to 100 when the stabilization, and then select the maximum depth, after continuous adjustment found that, when the Random Forest selects the classifiers 100, maximum depth 4, the model classification effect is the best. Support vector machine parameter selection, because this paper studies the nonlinear problem, the use of Gaussian kernel function, the penalty factor to take 10000, Bayesian parameters are set to the default value, lightgbm when the number of base classifiers is 30, the maximum depth of 4 when the prediction effect is the best.

Figure 1-Figure 5 shows the ROC curves of each classifier, the AUC values of Random Forest, Support Vector Machine, Bayesian, lightgbm and Stacking integration, respectively, are: 0.953, 0.670, 0.73, 0.91, 0.97, the AUC value of Random Forest, lightgbm and Stacking integration is greater than 0.85 in terms of the performance of the classifier, indicating that the model classifies better. It shows that the model classification effect is better, indicating that the method of Stacking integration can be applied to the risk assessment of listed companies.

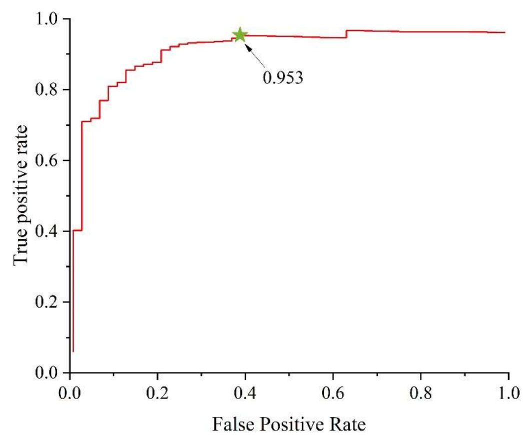

Figure 1: Random forest Roc curve

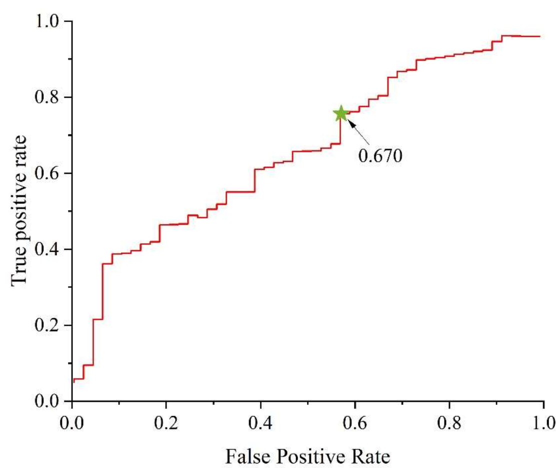

Figure 2: Support vector machine Roc curve

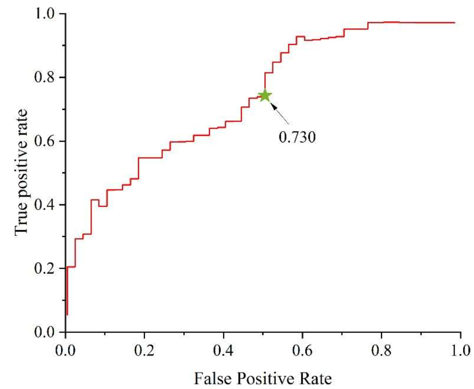

Figure 3: Beels Roc curve

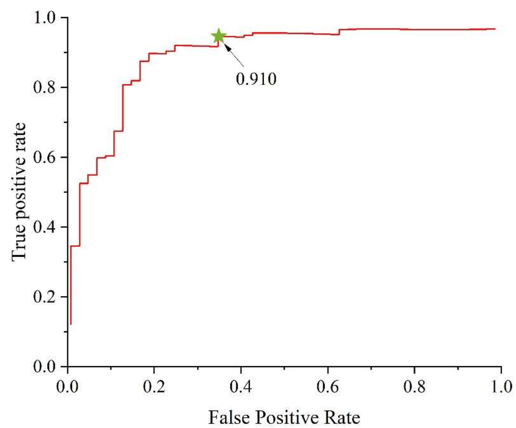

Figure 4: Lightgbm Roc curve

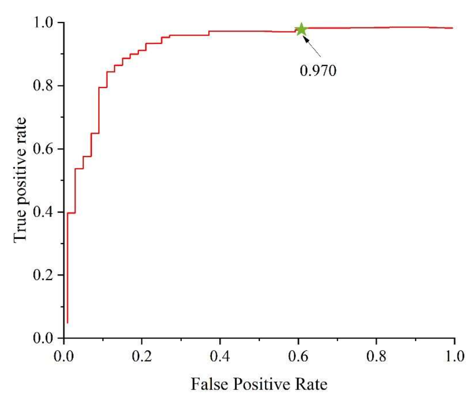

Figure 5: The stacking integrated roc curve

The evaluation indicators such as model accuracy rate and true case rate are specifically shown in Table 4:

Table 4: Model prediction effect contrast

Classifier	Accuracy rate	Accuracy ratio	True rate	AUC value
Random forest	0.924	0.843	0.984	0.941
Beels	0.666	0.679	0.532	0.669
Support vector machine	0.693	0.664	0.705	0.754
Lightgbm	0.931	0.876	0.962	0.925
Stacking	0.953	0.877	0.972	0.949

III. A. 3) Analysis of results of model application

Various evaluation metrics of the models revealed that the Stacking integrated model had the highest accuracy in the check accuracy rate and Lightgbm had the highest accuracy in the base learner. The true instance rate, which is the proportion of actual sustainability levels predicted by the model to practically all sustainability levels, was highest for Random Forest and second highest for Stacking Integration. The accuracy rate, on the other hand, reflects the model's ability to predict all samples, i.e., the ability to recognize correct samples and incorrect samples, and in the accuracy rate, Stacking integration accuracy is higher than several other and learners; In the ROC curve reflecting the overall effect of the model, the area under the Stacking integration curve is basically equal to that of the Random Forest, indicating that Stacking performs best with the Random Forest, and the model has the best overall effect.

III. B. SHAP-based analysis of factors influencing sustainable development

III. B. 1) Sustainable development impact factor identification

In this paper, based on the integrated Stacking algorithm, regression modeling of the rural pension system sustainability index and its influencing factors was conducted, and combined with the SHAP interpretation framework at two levels. From the global perspective, the SHAP value of each index is arranged in descending order from large to small to get the ranking of the contribution rate of each index, and the index with larger SHAP value is relatively more important. From the local analysis to plot the scatter plot of SHAP value of each indicator to further study the change of the contribution rate of each indicator is shown in Table 5. After comprehensive analysis, it is found that the speed of development of the system, the affordability of the economy at all levels, the rationality of the institutional setup, and the efficiency of management services are the main influencing factors for the sustainable development of the pension system.

Table 5: Impact factor contribution details

Index	SHAP value	Sort
System development speed	4.87	1
Economic affordability at all levels	1.67	2
Rationality of institutional setting	1.37	3
The efficiency of management services	1.11	4
The integrity of the supervisory mechanism	0.89	5
Insurance beneficiary	0.76	6
The level of satisfaction of life	0.59	7
Intergenerational fairness	0.48	8

The evolution of the contribution of the stability of the system and the normality of organizational management is shown in Figures 6 and 7. From the figures, it can be seen that the stability of the system has an inverted U-shaped relationship with the sustainable development of the pension system, and the normality of organizational management has an M-shaped relationship with the development of the pension system.

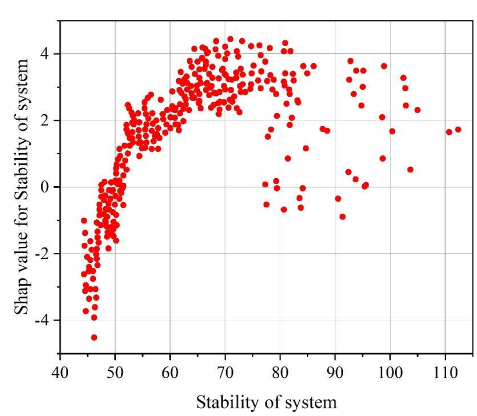

Figure 6: The stability rate of the system changes

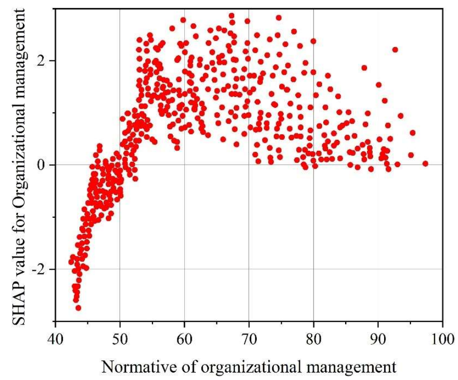

Figure 7: Changes in the normative contribution rate of organizational management

III. B. 2) Exploring the Interaction of Influential Factors

In order to further explore the interactions among the indicators, this paper utilizes Python to draw scatter plots of variable interactions based on SHAP values as shown in Figures 8 and 9. The horizontal coordinates of the two plots indicate the change in the value of the indicator, and the left axis of the vertical coordinate indicates the size of the corresponding SHAP value of the indicator. Pink dots indicate that the contribution of the indicator is negative and red dotted indicates that the contribution of the indicator is positive. From Figures 8 and 9, it is found that the stability of the system has an inverted U-shaped relationship with the standardization of organizational management, and is positively related to the level of sustainable development of the pension system. In Figure 8, the higher the level of stability of the system, the more it is affected by sustainable development. In figure 9, as the level of sustainable development of the pension system rises to a certain level, the marginal effect of the stability of the system on the norms of organizational management decreases, i.e., when the stability of the system continues to increase, it does not necessarily continue to significantly contribute to the growth of the norms of organizational management.

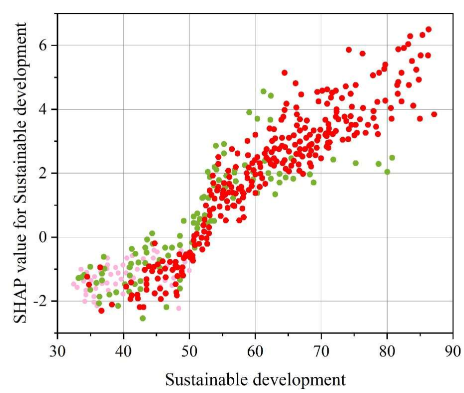

Figure 8: Interaction influences evolution

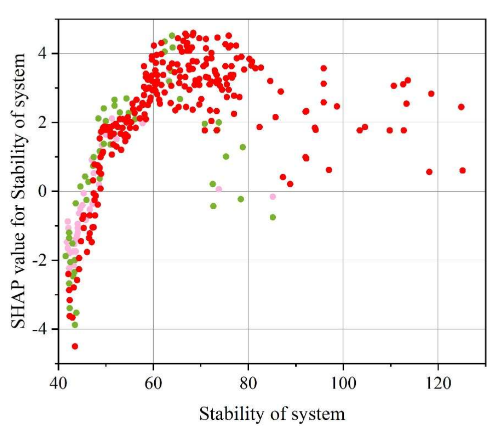

Figure 9: Interaction influences evolution

III. B. 3) Cluster analysis of SHAP values

Since the stability of the system plays a decisive role in the process of sustainable development of the pension system, the relationship between this factor and other indicators is further explored. This paper clusters the SHAP values of the stability indicators of the system, divides the study as a whole into two samples with higher and lower innovation dynamics, and separately counts the contribution rate of the factors influencing the sustainable development of the pension system as shown in Table 6.

From the table, it is found that the contribution rate of economic affordability at all levels is relatively high in the development region where the development speed of the system is lower. Speed is the first driving force of development and is the cornerstone for enhancing the development of rural cultural pension systems, and the lower the economic burden capacity at all levels, the greater the need for development and the more important the speed of system development.

Table 6: The contribution rate of the sample influence factor is counted

Index	SHAP value	Sort	SHAP value	Sort
	The first is the same		Second sample	
System development speed	0.72	6	0.83	6
Economic affordability at all levels	4.57	1	5.93	1
Rationality of institutional setting	1.44	2	1.88	2
The efficiency of management services	0.85	5	0.87	5
The integrity of the supervisory mechanism	0.59	8	0.64	7
Insurance beneficiary	1.24	3	1.79	3
The level of satisfaction of life	0.61	7	0.89	4
Intergenerational fairness	0.91	4	0.59	8

III. C. Strategies to enhance the sustainability of rural cultural pension systems

1. Optimizing the population structure and easing the aging of the rural population

Population aging leads to the disappearance of the demographic dividend, and in order to cope with the unfavorable situation of a lack of working population in the future, optimizing the population structure at multiple levels and from multiple perspectives is the key. Under the guidance and regulation of the family planning policy, the total fertility rate of China's population has been significantly lower than the level of population replacement. Due to demographic inertia, the total population still continues to grow, but at a significantly slower rate, and the process of population ageing is accelerating. However, there are obvious regional economic, social and cultural differences in China, with fertility concepts differing from place to place and fertility rates varying. In this regard, we should strengthen the hierarchy and adaptability of our policies, gradually relaxing the population policy in areas where the fertility rate is too low and moderately encouraging childbearing, while strictly enforcing the family planning policy in areas where the fertility rate is too high, increasing the penalties for exceeding the birth limit, and promoting the transformation of the rural population structure towards a younger age structure in a multi-pronged manner. Optimization should be carried out from multiple perspectives, including the age structure of the population, gender structure, industrial structure, and educational structure, in order to avoid the aging of the rural population and the risk of unsustainable pensions that it will bring about: moderately adjusting the fertility policy to improve the age structure of the population; continuing to promote the policy of caring for girls, and strengthening the strength of the policy in support of the control of the gender structure of the population; advocating the reasonable "going out" and "staying in" policies; and encouraging a rational "going out" and "staying in" policy. Advocating reasonable "going out" and "staying in", coordinating the development of the three major industries in rural areas, and upgrading the industrial structure of the rural population; paying attention to education in rural areas, especially in remote and impoverished areas, and making every effort to solve the problem of irrational allocation of educational resources in urban and rural areas, so as to promote the upgrading of the educational structure of the rural population.

2. Adjusting the key parameters of the new rural social security system at the right time to improve the sustainability of rural pensions.

The new rural social pension insurance contribution rate, subsidy rate, number of months of pension payment and other key system parameters directly affect the dynamic balance of rural pension income and expenditure. With reference to the actual development of the rural economy and society, the contribution level of the new rural social pension insurance should be set flexibly on the basis of five grades of contributions, and the contribution level of the new rural social pension insurance should be set flexibly on the basis of multiple grades and adjustable, so as to guide the insured farmers to make a reasonable choice (instead of concentrating on the lowest grade), improve the enthusiasm of farmers in making contributions, and contribute to the increase of the income of the rural pensions;

and the differentiated government subsidy rate should be set according to the different degree of economic development of different regions, so as to achieve the mutual benefit of the new rural insurance policy in Inter-regional mutual aid, is conducive to promoting the balance of income and expenditure of the new rural insurance; individual account pension counting months can be based on the changes in the average life expectancy of China's rural residents using indexing and other quantitative analysis methods of gradual adjustment to adapt to the accelerating development of the aging trend in the countryside to cope with the risk of longevity, in order to improve the sustainability of rural pensions.

3. Actively responding to population aging and rationally channeling rural pension pressure

In the face of the accelerated aging of the population, the surge in rural pension demand and the continued expansion of the rural pension gap and other realities, a positive attitude should be taken to cope with the aging of the rural population, the development of the rural aging business, and the promotion of the employment of the elderly re-employment, for the channeling of the rural pension pressure to reduce the risk of unsustainability of rural pensions has an obvious positive role in promoting. Re-employment of older persons is an important way of realizing their own value, creating social wealth, promoting their physical and mental health and reducing the pressure on old-age pensions. As the physical condition of the elderly has deteriorated, especially the rural elderly, years of physical labor have accelerated physical wear and tear, but they have accumulated a wealth of relevant experience in agricultural production, labor, management and so on. Therefore, the development of industries or positions suitable for the re-employment of rural older persons (such as organizers of public welfare institutions, agricultural technology consultants, community administrators, etc.) can raise the income of rural older persons, improve their standard of living, create corresponding wealth for society, and further rationalize the pressure on rural old-age care.

IV. Conclusion

In this paper, through the established sustainability indicators of rural cultural pension system, the integrated Stacking machine learning algorithm is utilized to evaluate the sustainability, and the SHAP explanatory model is combined to identify the influencing factors of the sustainability strategy, and the following conclusions are drawn based on the study:

(1) In terms of accuracy, the Stacking integrated prediction accuracy is 0.953, which ranks the highest among all learners. Based on the comparison results of accuracy rate, check accuracy rate, true example rate and AUC value, it can be seen that Stacking integrated model is superior to all other models except random forest.

(2) Combined with the SHAP explanatory model, it is found that the development speed of the system, the economic affordability at all levels, the reasonableness of the institutional setup, and the efficiency of the management services are the main influencing factors for the sustainable development of the pension system. The stability of the system and the standardization of the organization and management have an inverted “U”-shaped relationship, which is positively related to the level of sustainable development of the pension system.

Funding

1. One of the research outcomes from the Major Special Project (WH24-2-6) in the second batch of the 2024 Ningbo Cultural Research Program.

2. One of the research achievements of the General Scientific Research Project of Zhejiang Provincial Department of Education in 2024 (Y202456187).

References

[1] Abruquah, L. A., Yin, X., & Ding, Y. (2019). Old age support in urban China: the role of pension schemes, self-support ability and intergenerational assistance. International journal of environmental research and public health, 16(11), 1918.

[2] Wang, J., Zhang, G., & Tian, L. (2024). Reconstruction of China’s population age and education structure from 2000 to 2020—based on the data from the 5th to 7th national population censuses in China. China Population and Development Studies, 1-13.

[3] Tu, W. J., Zeng, X., & Liu, Q. (2022). Aging tsunami coming: the main finding from China’s seventh national population census. Aging clinical and experimental research, 34(5), 1159-1163.

[4] Williamson, J. B., Fang, L., & Calvo, E. (2017). Rural pension reform in China: A critical analysis. Journal of aging studies, 41, 67-74.

[5] Wang, Y., Yang, Q., Xin, L., & Zhang, J. (2019). Does the New Rural Pension System promote farmland transfer in the context of aging in rural China: Evidence from the CHARLS. International Journal of Environmental Research and Public Health, 16(19), 3592.

[6] Dong, K., & Wang, G. (2016). China’s pension system: achievements, challenges and future developments. Economic and Political Studies, 4(4), 414-433.

[7] You, J., & Niño-Zarazúa, M. (2019). The intergenerational impact of China's new rural pension scheme. Population and Development Review, 47-95.

[8] Li, C., Chi, I., Zhang, X., Cheng, Z., Zhang, L., & Chen, G. (2015). Urban and rural factors associated with life satisfaction among older Chinese adults. Aging & Mental Health, 19(10), 947-954.

[9] Cai, Y., & Cheng, Y. (2015). Pension reform in China: Challenges and opportunities. China's Economy: A Collection of Surveys, 45-62.

[10] Huang, W., & Zhang, C. (2021). The power of social pensions: evidence from China's new rural pension scheme. American Economic Journal: Applied Economics, 13(2), 179-205.
[11] Insurance, E. (2018). New Rural Social Endowment Insurance in China. Studies On Contemporary China, 3, 65.
[12] Gui, S. (2018). New Rural Social Endowment Insurance in China A Critical Analysis. Studies on Contemporary China, 65-72.
[13] Peng, H. (2023, September). Exploring the Relationship Between China's New Rural Pension Scheme and Rural Development. In 2023 9th International Conference on Humanities and Social Science Research (ICHSSR 2023) (pp. 1018-1027). Atlantis Press.
[14] Yingying, H. (2022). Analysis on the Effect of China's Rural Endowment Insurance System Reform. Academic Journal of Humanities & Social Sciences, 5(3), 7-9.
[15] Ma, S., Ma, S., & Cao, Q. (2018). The development of rural endowment patterns in China. Asia-Pacific Journal of Business & Commerce, 10(2), 22-43.
[16] Erfina Adhitia,Siti Utami Dian & Sembiring Falentino. (2024). Sentiment analysis of reviews related to Unesco Global Geopark (UGG) Ciletuh Palabuhanratu on google maps using the naive bayes algorithm. BIO Web of Conferences,148,
[17] S. Dinakaran & P. Ranjit Jeba Thangaiah. (2023). Effective hybrid feature subset selection for multilevel datasets using decision tree classifiers. International Journal of Advanced Intelligence Paradigms,24(1-2),206-228.