International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 8061-8074

https,//doi.org/10.70517/ijhsa463689

Intelligent management of human health conditions using integrated learning algorithms

Ying Zhu[1], Zekun Chen[2] and Maoquan Su[1,3,*]
[1] School of Management, Shandong Second Medical University, Weifang, Shandong, 261000, China
[2] Administrative Office of the Dean, Weifang People's Hospital, Weifang, Shandong, 261000, China
[3] The First Affiliated Hospital of Shandong Second Medical University (Weifang People's Hospital), Weifang, Shandong, 261000, China
Corresponding authors: (e-mail: zhuyingflx@163.com).

Abstract Human health status data is characterized by high dimensionality, complex indicators, and interactive relationships, and the traditional single prediction model faces the problems of insufficient accuracy and robustness. In this paper, we propose an intelligent management method of human health status based on Stacking integrated learning, which constructs a feature system from five dimensions: physical health, mental health, lifestyle and behavioral health, social adaptation and environmental health, and disease prevention and health management. The study firstly screened 20 key feature variables by Lasso regression and stepwise regression, and then designed a dynamic weight estimation algorithm based on the Breiman method, combining long-term historical data and short-term neighboring data to optimize the weight configuration. The experimental results show that compared with a single model, the proposed Stacking integration model performs well in a number of metrics such as AUC, accuracy, and F1-Score, with an AUC value of 0.97287 and an accuracy of 0.93149. Through 10-fold cross-validation for 10 consecutive tests, the model demonstrates less volatility than the individual base classifiers, which verifies that the method is highly stable in the high stability in human health status prediction. The Stacking model in integrated learning significantly improves the accuracy and generalization ability of the prediction results by effectively integrating the advantages of various learners, provides reliable technical support for the intelligent assessment and management of health conditions, and is of great practical value in promoting personalized health management and precision medicine.

Index Terms human health condition, feature screening, integrated learning, Stacking model, weight estimation, intelligent management

I. Introduction

With the acceleration of people's life rhythm, people's attention to health problems is increasing day by day, while with the development of artificial intelligence, the application of intelligent algorithms in health management makes intelligent management of health conditions gradually become an emerging management mode [1]-[3]. Intelligent management of human health condition refers to the use of intelligent technology, combined with medical, health management and other professional knowledge, to carry out all-round, personalized health management for individuals, with personalized customization, real-time monitoring, remote services and cross-border cooperation [4]-[7]. The rise of intelligent health management has brought new opportunities and challenges for individual health management [8].

The application of intelligent health management provides more intimate and convenient services for people's health [9], [10]. Through the support of intelligent algorithms, it can collect, analyze and interpret individual health data, as well as targeted health interventions, to achieve the purpose of improving health, preventing diseases and managing chronic diseases, which can help us to better understand our own health status and effectively prevent and manage diseases [11]-[14]. At the same time, more personalized medical services can be obtained to improve the efficiency and quality of medical care [15]. However, while enjoying convenient and fast smart health management services, personal privacy protection and data security have become a new issue [16], [17]. Therefore, we need to develop appropriate laws and norms to protect individual privacy and information security [18].

Health management is an important part of modern healthcare system, which can provide scientific basis for disease prevention and health promotion through comprehensive assessment and dynamic monitoring of individual health status. Currently, health management faces many challenges: health data are diverse and heterogeneous, health assessment lacks standardization and precision, and traditional single algorithms are difficult to capture complex health factor interactions. These problems lead to the lack of accuracy of health assessment results, which cannot meet the needs of personalized health management. In recent years, machine learning methods have been increasingly used in the health field, providing new ideas for health condition assessment. However, single machine learning algorithms often have shortcomings such as limited generalization ability and sensitivity to noise, making it difficult to fully capture the complexity of health conditions. Integrated learning is expected to overcome the limitations of a single algorithm by combining multiple underlying algorithms to improve prediction accuracy and stability. In health condition assessment, it is particularly important to accurately identify key health features and build reliable prediction models. Traditional health assessment methods mainly rely on single or limited dimensional indicators, making it difficult to comprehensively reflect the multifaceted nature of individual health. At the same time, health data often have problems such as redundancy and multicollinearity, which affect model performance. Therefore, how to effectively screen health characteristics and construct a multidimensional health assessment system has become a key issue in health management research. In addition, health status has dynamic change characteristics, how to effectively integrate long-term health trend and short-term health fluctuation information in the model to realize dynamic prediction is also a difficult point in current research. Integrated learning, with its powerful synthesizing ability, offers the possibility to solve the above problems. In particular, the Stacking integration method, through the hierarchical learning strategy, is able to fully utilize the advantages of different underlying models to improve the overall prediction performance.

Based on the above analysis, this study proposes an intelligent management method of human health status based on integrated learning. The study first constructs a comprehensive human health characteristic system, covering five dimensions: physiological, psychological, behavioral, social and preventive; Second, a combination of Lasso regression and stepwise regression is used to screen key health features; then, a dynamic weight estimation algorithm based on the Breiman method is designed to fuse the long-term and short-term health data for prediction; finally, a variety of base learners are integrated through the Stacking method to optimize the model performance. The study empirically analyzes the large-sample population health data to validate the effectiveness of the proposed method, and assesses the model stability through cross-validation.

II. Relevant theories

II. A.Feature engineering

Feature engineering (FE) is the process of transforming raw data into features that can better express the nature of the problem, and applying these features to predictive models can improve the accuracy of the model's prediction of future data. How to be able to decompose and aggregate the raw data to better express the nature of the problem is the key purpose of feature engineering, which simply means to discover the features that have an important effect on the dependent variable [19]. In addition, feature engineering is also an important step in converting raw data into datasets.

II. A. 1) Feature extraction

In machine learning, a feature is a parameter that is used to describe some aspect of a data object, such as a person's age or pupil color, the mean or plurality of a student's grades. Features of information are the basis of data analysis, they help to describe objects and distinguish between different representations of an object. They are the source of the “knowledge” learned by the machine and are essential for generating accurate models that produce good results.

In general, features are divided into the following categories:

(1) Categorical features: Categorical features are discrete values that distinguish an object intuitively from other objects, such as color, species, etc. Some objects may have a particular categorical feature, which has only two values: yes or no.

(2) Ordered features: Ordered features can also classify objects, and also represent the sorting of objects in a certain characteristic, such as high, moderate, low satisfaction, or the education level of elementary school, secondary school, bachelor's degree, master's degree, doctoral degree and so on.

(3) Numerical features: Numerical features, on the other hand, are a numerical value or a set of numerical values, which represent a certain quantifiable characteristic of an object. Numerical features are usually human-defined concepts, which are formulated and standardized to provide a detailed description of some abstract contents, for example, it is far from enough to describe the temperature only with hot and cold, and the quantified temperature value can give a more intuitive feeling.

In real-world applications of machine learning, it is also common for categorical features to be represented as numerical features, because some categorical features are regarded as influencing factors, but computers cannot understand descriptions such as the color red in human language, so they are usually coded to convert them, for example, using 1~7 to represent seven different colors as inputs to the model, and it is worth noting that the order of numerical values implied by these cases is often not the same as that of the numerical values in human language.

It is worth noting that in such cases, the implied order of the values is often meaningless and only serves as a distinction between different features.

II. A. 2) Feature standardization

Usually a normalization operation, also called normalization, is performed on the data before training the model. In neural network problems, due to the difference in scale of the original features, the contour plot of the loss function may be elliptical, and the gradient direction is perpendicular to the contour. If not normalized, then due to the large difference in the values of different features in the feature vector, it will lead to a “flat” objective function, so that in the gradient descent, the direction of the gradient will deviate from the direction of the minimum value, and go through a lot of detours, i.e., the training time is too long. If the standardization, the objective function will appear more “round”, so that the training speed is greatly accelerated, eliminating unnecessary time. In short, standardization can accelerate the speed of gradient descent to find the optimal solution, and standardization has the potential to improve the accuracy of the model. There are three main methods of common standardization:

(1) Linear standardization: also known as deviation standardization, is a linear transformation, able to convert the data to the interval [0, 1], the transformation formula is:

$$X^* = \frac{X - X_{\min}}{X_{\min} - X_{\max}} \tag{1}$$

In the above equation, X is the original data, X^* is the normalized data, and $X_{\min}$ and $X_{\max}$ are the minimum and maximum values in the original data, respectively. This normalization method is more suitable in the case of concentrated values, which can simplify the calculation for model training, its disadvantage is that if the maximum and minimum values are unstable, it will also make the normalization result unstable, that is, if the data has anomalies or the extreme value is the maximum value, then the normalization result will also be greatly affected, so in practice, it can be used with empirical constants instead of max and min, suitable for the data not involving distance measure, covariance calculation, data does not meet the normal distribution, from the engineering aspect, generally also used in the project with small amount of data.

(2) Z-score standardization: this method requires the use of the mean and standard deviation of the original data to do standardization of the data, after standardization of the data in line with the mean value of 0, the standard deviation of 1 of the standard normal distribution, but the interval after the change is uncertain, the transformation formula is:

$$X^* = \frac{X - \mu}{\sigma} \tag{2}$$

In the above equation X is the original data, X^* is the normalized data, μ is the mean of the original data, and σ is the standard deviation of the original data [20]. Z-score normalization is related to the overall distribution of the original data, and each data point can have an impact on the results, in addition, the Z-score normalized data can retain the valuable information in the original outliers. Valuable information in the original outliers, making the algorithm less sensitive to the outliers, which is different from linear normalization. The advantage of this method is that it can improve the convergence speed of the model and simplify the calculation. It is usually used in classification and clustering algorithms and when distance is needed to measure similarity.

(3) Non-linear normalization: Non-linear transformation is often used in scenarios where data differentiation is more pronounced, when some values in the data are large but some are very small, the original values can be mapped by some mathematical functions. The actual use of which formula needs to be selected depending on the specifics of the data distribution. The formulas for the logarithmic transformation and the inverse tangent transformation are as follows, respectively:

$$X^* = \frac{\lg(X)}{\lg(X_{\max})} \tag{3}$$

$$X^* = \frac{2\arctan(X)}{\pi} \tag{4}$$

The formula requires that the original data is greater than or equal to 1. The normalized data in the inverse tangent transformation ranges from [-1, 1], and if there is a special need for the normalized interval in [0, 1], then the original data should be guaranteed to be greater than or equal to 0.

II. B.Machine learning models

II. B. 1) Logistic regression

Logistic regression, which belongs to the category of supervised learning, plays an important role in machine learning. It covers the subfields of binary regression and multivariate disordered regression. Binary regression is particularly commonly used in academic discussions of corporate financial crisis prediction given that it responds to problems with only two possible outcomes, i.e., dichotomous tasks. Logistic regression modeling, which focuses on exploring how specific attributes affect the likelihood of a particular outcome, demonstrates greater flexibility in dealing with multivariate relationships. Although it requires a random distribution of the dataset, it focuses on managing the interaction effects between the variables rather than strict pre-conditioning. Logistic regression is capable of picking performance through the following formula:

$$LR\left(\frac{p}{1-p}\right)=\beta_0+\beta_1X_1+\beta_2X_2+...+\beta_nX_n \tag{5}$$

where X_n represents the n th independent variable and $\beta_n(n=1,2,\ldots,n)$ represents the corresponding parameter estimates of the variable.

II. B. 2) Neural Networks

Neural network modeling is an approach that emulates the principles of biological neural networks, which centers on the use of a combination of linear and nonlinear methods to process multivariate eigenvalues in order to achieve the desired computational results. As the input data is passed through the layers of the network, the corresponding output is eventually produced. A basic multilayer perceptron (MLP) model usually consists of an input layer, a hidden layer, and an output layer, and its mathematical formulation is roughly as follows:

$$\lambda=M[0]\cdot x[0]+M[j]\cdot x[j]+\cdots M[b]\cdot x[b]+p \tag{6}$$

Within the MLP architecture, each neuron in the hidden and output layers is equipped with a unique activation function common choices include tanh and relu, among others. The activation function comes into play after the generation of the hidden layer and its main task is to perform a nonlinear transformation of the input signal to accommodate the learning needs of processing complex nonlinear data sets.

II. B. 3) Random Forests

A random forest (RF) is a collection of unique and mutually independent decision trees working together to make decisions through collective intelligence. A subset of features is randomly selected at each training session and a new tree is constructed based on these subsets as a way to weaken the decisive influence of specific features on the final result. Integration learning, an important pillar in modern machine learning strategies, covers a variety of strategies such as bagging, boosting and stacking.

Due to the randomness of sample and feature selection, a random forest composed of multiple decision trees can effectively prevent overfitting and thus achieve lower prediction errors. However, in the face of limited data volume, insufficient features, or missing data, random forests may not be able to demonstrate excellent classification performance. Training a random forest involves constructing numerous decision trees and performing integrated voting, a process that can take a long time. In addition, Random Forests have a built-in feature importance assessment mechanism, which can help select features that have a significant impact on the predictions by quantifying the average impact of each feature on the predictions of different categories to reveal the differences in the importance of the features.

II. B. 4) XGBoost

XGBoost is a novel gradient enhancement algorithm due to its efficient parallel training and significant improvements in ML-based applications.XGBoost is a variant of the integrated method GBDT, which combines a gradient enhancement optimization strategy with a DT classifier, i.e., combining multiple DTs into a gradient boosting framework that iteratively optimizes the training objective.GBDT is a additive combination of m based learner model, if the tree model trained in the m th iteration is $f_m(x_i)$, then the expression of GBDT is:

$$\hat{y}_i^{(m)}=\sum_{m=1}^{M}f_m(x_i)=\hat{y}_i^{(m-1)}+f_m(x_i),f_m\in F,i\in N \tag{7}$$

where $\hat{y}_i^{(m)}$ denotes the prediction result of the sample i after the m th iteration, $\hat{y}_i^{(m-1)}$ denotes the prediction result of the previous $m-1$ tree, and $f_m(x_i)$ denotes the model for the first m th tree.The objective of the XGBoost function consists of two parts, the loss function and the regularization term, and the formula is as follows:

$$L^{(m)} = \sum_{i=1}^{N} l(y_i, \hat{y}_i^{(m)}) + \sum_k \Omega(f_m) \tag{8}$$

The $L^{(m)}$ represents the difference between the true value y_i and its predicted value $\hat{y}_i^{(m)}$. In order to alleviate the overfitting problem, XGBoost adds a regularization term $\sum_k \Omega(f_m)$ to the loss function, and by combining the two formulas mentioned above, we can obtain the following XGBoost loss function:

$$L^{(m)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(m-1)} + f_m(x_i)) + \Omega(f_m) \tag{9}$$

The above formula is then subjected to a second-order Taylor expansion, removing the constant term, and the formula is derived as follows:

$$\begin{aligned} L^{(m)} &= \sum_{i=1}^{n} l\left[y_i, \hat{y}_i^{(m-1)} + f_m(x_i)\right] + \Omega(f_m) + C \\ &= \sum_{i=1}^{n}\left[l(y_i, \hat{y}_i^{(m-1)}) + g_i f_m(x_i) + \frac{1}{2} h_i f_m^2(x_i)\right] + \Omega(f_m) \\ &= \sum_{i=1}^{n}\left[g_i f_m(x_i) + \frac{1}{2} h_i f_m^2(x_i)\right] + \gamma T + \frac{1}{2}\lambda w_j^2 \end{aligned} \tag{10}$$

where g_i, h_i denotes the first and second order derivatives of the objective function, respectively:

$$\begin{aligned} g_i &= \partial_{\hat{y}^{(t-1)}} l(yi, \hat{y}^{(t-1)}) \\ h_i &= \partial^2_{\hat{y}^{(t-1)}} l(yi, \hat{y}^{(t-1)}) \end{aligned} \tag{11}$$

And the regularization term $\Omega(f_m)$ is used to measure the complexity of the tree, which consists of two parts: the number of leaf nodes and the leaf node weights, respectively. In the expansion, T represents the number of leaf nodes and w denotes the weight of leaf nodes, which is controlled by the coefficients γ and λ in order to prevent overfitting. The regularization term expression is as follows:

$$\Omega(f_x) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \tag{12}$$

Define a tree $f_m(x) = w_{a(x)}, w \in R^T, q : R^d \to \{1, 2, \cdots, T\}$, which consists of two parts: weight vectors of the leaf nodes w and mapping relations of the leaf nodes q. Substitute $f_m(x)$ and the regularization term expansion into the above equation:

$$\begin{aligned} &\sum_{i=1}^{n}\left[g_i f_m(x_i) + \frac{1}{2} h_i f_m^2(x_i)\right] + \gamma T + \frac{1}{2}\lambda w_j^2 + C \\ &= \sum_{i=1}^{n}\left[g_i \omega_{q(x_i)} + \frac{1}{2} h_i \omega_{q(x_i)}^2\right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \\ &= \sum_{j=1}^{n}\left[\left(\sum_{i \in I_j} g\right)\omega_j + \frac{1}{2}\left(\sum_{i \in I_j} h_i + \lambda\right)\omega_j^2\right] + \gamma T \end{aligned} \tag{13}$$

Define $G_j = \sum g, H_j = \sum h$, and the final objective function of XGBoost can be obtained by substituting G_j and H_j into the loss function:

$$L^{(m)} = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j+\lambda}+\gamma T+C \tag{14}$$

Solving the objective function, the optimal solution is reached when $w_j^* = -\frac{G_j}{H_j+\lambda}$, $Obj = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j+\lambda}+\gamma T$ minimizes the loss function.

II. B. 5) LightGBM

LightGBM, as a tree model, adopts a GOSS strategy to cope with sample weight differences, which cleverly utilizes a random sampling strategy aimed at maintaining the balance among training samples to reshape the global distribution properties of the dataset.The efficacy of LightGBM is significantly affected by the hyper-parameter settings, and requires fine tuning of the parameters to achieve the optimal performance, and suffers from the problem of high memory consumption. The prediction result of a LightGBM model containing T decision trees can be expressed as:

$$y = \sum_{t=1}^{T} f_t(x) \tag{15}$$

x represents the input data and y represents the actual output result.

II. B. 6) Stacking Integration Models

The Stacking strategy is an integration approach that aims to combine the strengths of multiple models to improve the overall prediction efficacy.Stacking is a unique integration technique that incorporates the features of Boosting and Bagging. In Stacking, multiple base learners first learn the original data in parallel, and subsequently the learning results generated by these base learners are fed into a second level of the model for higher level fitting and integrated decision making [21].

Stacking is modeling the stacking of the original data fit by the model, first the original data is learned by the base learner, and then several base learners are output to the original data, and the outputs of these models are stacked according to the columns, which constitutes the (m, p)-dimensional new data, with m representing the number of samples, and p representing the number of base learners, and then the new sample data is given to the second level model for fitting. The schematic diagram of Stacking principle is shown in Fig. 1.

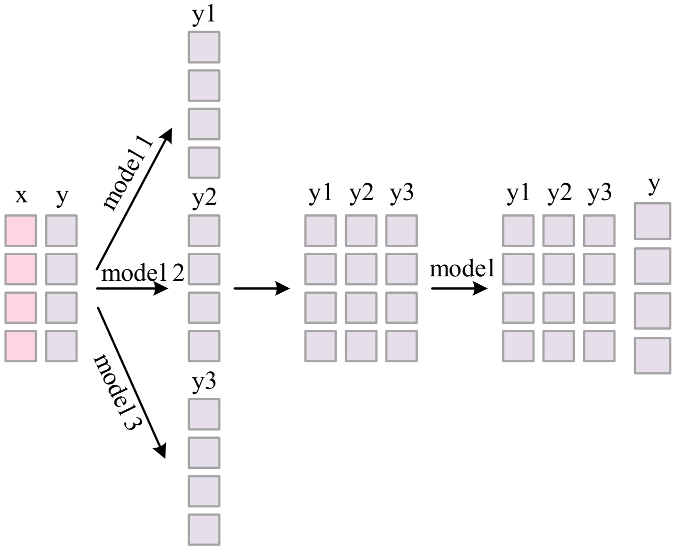


Figure 1: Schematic diagram of the Stacking principle

The 4-fold cross-validation Stacking model training is shown in Figure 2. Firstly, using K-fold cross validation approach, the data is sliced in 4 folds, then 4 groups of data sets are derived, then the models are arranged to carry out training activities for each group of training sets respectively, and finally the prediction processing of the validation set is completed, so as to obtain the output results [22]. Due to the implementation of the 4-fold cross-validation operation, a total of 4 groups of data are included, so the validation sets formed are also one-to-one correspondence, and then the results predicted by each model for the validation sets of their respective groups are stacked according to the rows, the predicted value of the complete sample data will be obtained, this is only for a model, different learners are the same, each model according to the predicted value of the predicted value in this way, and then merged according to the columns.

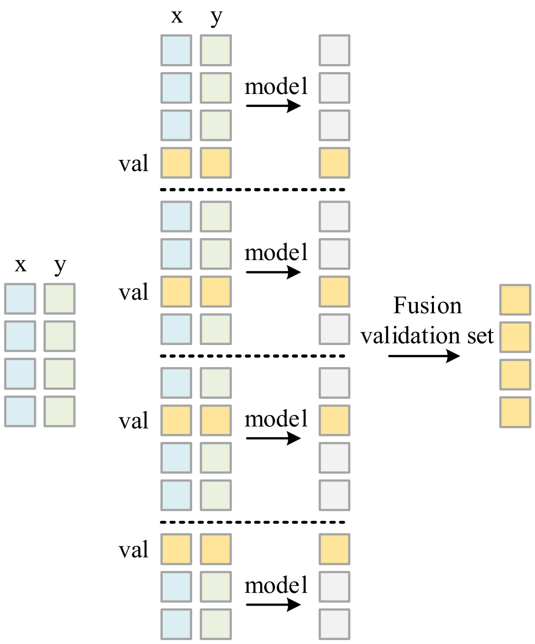


Figure 2: Training diagram of the 4-fold cross-validation Stacking model

III. Machine learning-based human health condition feature screening

III. A. Data sources

The data used in this section of the experiment came from households in a neighborhood in City A. The respondents were aged from 18 to 60 years old, and there were about 15,000 respondents in about 10,000 households. The contents of the questionnaire include: basic personal information, family information, health status, and physical measurements.

III. B. Feature engineering

III. B. 1) Feature construction

In this paper, human health condition is regarded as a label for model training, and human health condition is analyzed in five dimensions: physical health, mental health, lifestyle and behavioral health, social adaptation and environmental health, and disease prevention and health management. The feature construction of human health condition is shown in Table 1.

III. B. 2) Feature screening

Therefore, Lasso regression was used to screen variables to solve the problem of multicollinearity. The specific process is as follows: firstly, the Lasso regression is carried out with the variable "health" as the label, when parameter λ gradually changes from large to small, more and more variables will be screened out into the model, the results of Lasso regression are shown in Table 2, four decimal places are retained here, it can be seen that when λ>1621.32, due to the excessive penalty, no variable enters the model at this time, and when λ=1621.32, only one constant term enters the model. And when λ drops further to 1459.65, the variable enters the model.

After that for selecting the specific values of the fine-tuning parameters for k-fold cross-validation, the default value is used here, i.e., 10-fold cross-validation, and the cross-validation plot is shown in Figure 3.

Table 1: The characteristics of human health condition construction

Dimension	Feature	Representation
Physiological health	Cardiovascular function	CF
	Respiratory function	RF
	Metabolic state	MS
	Immune system	IS
	Digestive health	DH
	Bone and muscle state	BMS
	Nervous system function	NSF
Mental health	Emotional stability	ES
	Cognitive ability	CA
	Pressure management ability	PMA
Lifestyle and behavioral health	Dietary nutrition	DN
	Sports habit	SH
	Quality of sleep	QS
	Addictive behavior	AB
Social adaptation and environmental health	Social support network	SSN
	Economic situation	ES
	Occupational health	OH
	Sanitation	S
Disease prevention and health management	Risk of chronic disease	RCD
	Vaccination coverage	VC
	Health screening compliance	HSC
	Genetic factor	GF

Table 2: Lasso returns partial results

ID	λ	Variable number	R^2	New variable
1	1621.32	0	0.0000	BMS
2	1459.65	1	0.0196	NSF
3	1332.26	2	0.0365	OH
5	1213.68	3	0.0512	S
7	1005.85	4	0.0893	RCD
12	695.75	5	0.1532	CF
13	533.96	6	0.1521	RF
7	387.32	9	0.1763	MS
8	352.61	10	0.1802	IS
16	322.65	13	0.1846	SH

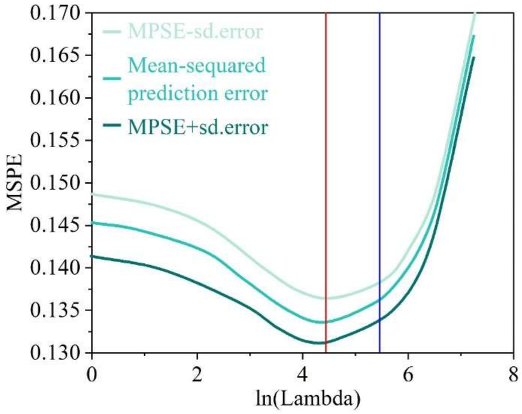


Figure 3: Cross verification diagram

The results show that when λ=89.655 can minimize the MSPE, so this parameter size again Lasso regression, λ=89.655, Lasso regression part of the results as shown in Table 3, where retained the four decimal places, a total of 90 screening variables.

Table 3: λ=89.655, lasso returns partial results

Selected variable	Coefficient
CF	0.0755
RF	-0.0002
MS	0.0269
IS	0.0000
DH	-0.0081
BMS	0.0000
NSF	-0.0026
AB	-0.0043
SSN	-0.0065
ES	0.0032

It can be seen that the number of variables is still large, so combined with the stepwise regression method for further feature screening, i.e., the dataset after the screening of Lasso regression using stepwise regression, based on the p-value of the variables to further screen the dataset's features. The results of stepwise regression are shown in Table 4. There are 20 feature variables left in the final dataset.

Table 4: Stepwise regression

Variable name	Coefficient	T value
CF	0.1089	0.0000
RF	0.0277	0.0000
MS	0.029	0.0000
IS	0.0002	0.0155
DH	0.0375	0.0000
BMS	-0.0082	0.0033
NSF	-0.0417	0.0121
ES	-0.0017	0.0000
CA	-0.0061	0.0322
PMA	0.03	0.0000
DN	-0.0059	0.1521
SH	0.0096	0.003
QS	0.0014	0.0000
AB	-0.0137	0.0026
SSN	-0.1344	0.001
ES	-0.0139	0.0033
OH	-0.0329	0.0000
S	-0.0064	0.0068
RCD	-0.0414	0.0000
VC	0.0178	0.0045

IV. Stacking-based weight estimation algorithm

The base model learning algorithm is denoted as algorithm $\xi_1,...,\xi_K$, and the metamodel learning is denoted as algorithm ξ. The output of the structural response prediction result for the i th base learning algorithm at moment T is denoted as $\xi_i(X_T)=\hat{y}_{T,i}$, where $i=\{1,...,K\}$. Simplify $\hat{y}_{T,i}$ by noting $\hat{y}_i=[\hat{y}_i^1,...,\hat{y}_i^F]^T$, where F is the number of variables of the multivariate prediction task, i.e., the number of sensors to be predicted. The output of the K base model is spliced as $Y^{oc}=(\hat{y}_1,\ldots,\hat{y}_K)$.

IV. A. Breiman method

An integrated strategy that combines generalized linear models and Stacking methods to assign weights to each base model based on its performance on historical prediction tasks.The Breiman method is formulated as follows:

$$\xi(X)=\sum_{i=1}^{K}\theta_i\hat{y}_i \tag{16}$$

where the scalar $\hat{y}_i$ is the prediction of the i th base model and θ_i is the corresponding weight. For $i\in\{1,...,k\}$, the weights assigned to each model need to satisfy the following conditions:

$$\sum_{i=1}^{k}\theta_i=1,\theta_i\geq 0 \tag{17}$$

That is, the weights θ_i contributed by the base model should be non-negative and the sum of the weights θ_i is 1.

In the multivariate task prediction integration task, the approach for model Breiman, whose output is a scalar $\hat{y}_i$ model, needs to be extended to an integration approach for vectors $\hat{y}_i$:

$$\xi(X)=\sum_{i=1}^{K}\theta_i\hat{y}_i \tag{18}$$

For the task of multivariate prediction, it is necessary to separately consider the contribution of each target from each model and then assign the weights, the formula for multivariate $\hat{y}$ prediction task integration in this paper is:

$$\hat{y}=\xi(X)=\sum_{i=1}^{K}\theta_i\hat{y}_i \tag{19}$$

where, for $\forall i\in\{1,...,K\}$, θ_i is the weight matrix of the multivariate vector $\hat{y}_i$ for the output of the i th base model:

$$\theta_i=\begin{pmatrix}\theta_i^1 & & \\ & \ddots & \\ & & \theta_i^F\end{pmatrix} \tag{20}$$

For $j=\{1,...,F\}$, θ_i^j is the weight of the contribution of the model at the i th model on the prediction of each target $\hat{y}_i^j$.

Thus the output of the final integrated model can be written as:

$$\xi(X)=\begin{bmatrix}\sum_{i=1}^{k}\theta_i^1\hat{y}_i^1\\ \vdots\\ \sum_{i=1}^{k}\theta_i^F\hat{y}_i^F\end{bmatrix} \tag{21}$$

where the $\xi(X)$ output of the integrated model has a multivariate vector with the j th row corresponding to the integrated output value of the structural response of the j th sensor, obtained by weighting and summing the outputs of each model for that sensor. And the weights θ_i^j need to satisfy the following two conditions:

$$\begin{gathered}\theta_i^j\geq 0,\forall i\in\{1,...,K\},\forall j\in\{1,...,F\}\\ \sum_{i=1}^{K}\theta_i^j=1,\forall j\in\{1,...,F\}\end{gathered} \tag{22}$$

That is, the weights contributed by each model need to be non-negative and the sum of the contribution weights of each model for each prediction target is 1. The estimation of the optimal weights θ^* under the above constraints is then carried out by least squares on the dataset D:

$$\theta^* = \arg\min_\theta \| \xi_D^\theta (Y^\infty) - y \|_2 \tag{23}$$

IV. B. Dynamic integration weight estimation

The Breiman approach usually refers to the estimation of individual base model weights using historical data, which refers to data collected in the past that may contain a wider range of situations, a larger range of variations, and a larger number of samples. Training with historical data can help the model learn more comprehensive and pervasive laws and patterns, thus enhancing the model's generalization ability. In contrast, data from adjacent time periods contain information closer to the current moment, reflecting the most recent trends and changes. Training with recent data allows the model to better capture current data patterns and features, which helps the model predict current and future situations more accurately. Inspired by long and short-term memory neural networks, this paper treats the weight estimation based on historical data D_L as “long-term weight” estimation, and the integration based on neighboring data D_S as “short-term weight” estimation. Since the health status of warm humans varies on a daily cycle, one day's data is used as the unit data volume for short-term estimation. In order to discuss the impact of long-term weight estimation and short-term weight estimation on the new prediction task, a schematic diagram of the weight estimation method is shown in Figure 4.

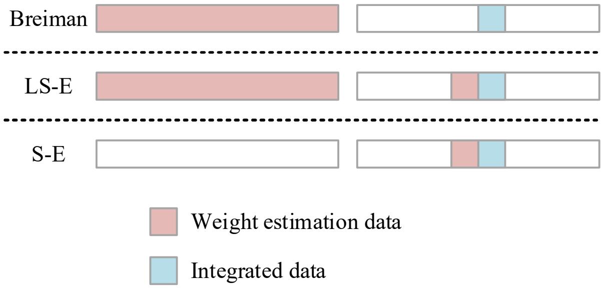


Figure 4: Schematic diagram of the weight estimation method

V. Performance testing experiments

V. A. Stacking model fusion based on integrated learning classifiers

The superiority of Stacking model is demonstrated in non-integrated learning, and the following further integration of integrated learning explores the role of Stacking in integrated learning. In this section, P1+P2+P3 are selected as the predicted feature set, Xgb,ADA,ET,GBDT,LightGBM,RF are selected as the base classifiers, and LR is selected as the classifier in the second stage. The integrated learning and Stacking model comparison is shown in Table 5. From the table, it can be concluded that in the base classifier as integrated learning, the fusion has the best result with an AUC value of 97.1%. Overall, the Stacking integrated learning model adopts K-fold cross-validation to reduce the risk of overfitting of the model, and uses the prediction results of multiple base learners for secondary training, which overcomes the limitations of a single learner to a certain extent, and is able to effectively integrate the scope of application and advantages of a variety of learners, and has obvious advantages in improving the accuracy and generalization of the prediction results.

Table 5: The integration learning is compared to the stacking model

	AUC	Accuracy	Precision	F1-Score
GBDT	0.96122	0.92655	0.92544	0.92577
ET	0.96747	0.93082	0.93198	0.93109
ADA	0.95282	0.91313	0.91197	0.91076
Xgb	0.966	0.93533	0.93478	0.93519
LightBGM	0.96761	0.93575	0.93401	0.93464
RF	0.96781	0.93553	0.93478	0.9355
Ours	0.97085	0.93553	0.93478	0.93519

V. B. Modeling Stacking Algorithm Based on Hybrid Classifier

V. B. 1) Highest Prediction Accuracy Algorithm Set Extraction

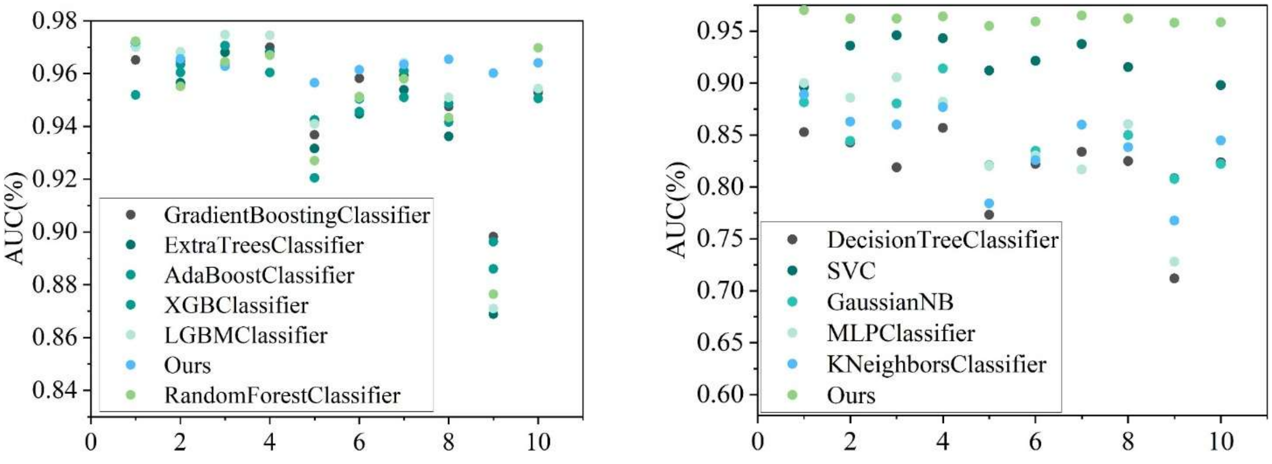
Through the Stacking model two phase study, in the first phase, the set of algorithms with the highest prediction accuracy is selected for different algorithms and here the prediction results are evaluated using AUC. The following are the top 10 ranked algorithm sets. The set of prediction accuracy algorithms is extracted as shown in Table 6. It is found that the highest prediction accuracy is the method of this paper. The model of this paper (0.97287) is able to improve the prediction accuracy as compared to Stacking integration model based on non-integrated learning method and Stacking integration model based on integrated learning method.

Table 6: Prediction algorithm collection

AUC	Accuracy	Optimal algorithm set
0.97287	0.93149	Ours
0.97197	0.93159	ET,Xgb
0.97252	0.93473	ET,ADA,Xgb,LGB
0.97252	0.93423	ET,Xgb,LGB
0.97195	0.93383	ET,Xgb,LGB,DT
0.97197	0.93253	ET,Xgb,LGB,SVC
0.97177	0.93433	ET,ADA,Xgb,LGB,SVC
0.97155	0.93391	ET,ADA,Xgb,LGB,Bayes
0.97175	0.93401	ET,Xgb,LGB,Bayes
0.97154	0.93506	ET,ADA,Xgb,RF

V. B. 2) Stacking Forecast Result Analysis

Integration learning has a high accuracy of classification prediction for health information adoption compared to non-integration learning, and the Stacking algorithm can integrate the base classifiers more efficiently than other algorithms, which can improve the accuracy of the model to a large extent, and the accuracy and precision of the model have been improved in both non-integration learning and integration learning, but the effect brought about by this may be against the improvement of non-integration learning more obvious. The effective algorithms in the first of the two phases of Stacking are also extracted.



(a) Integration

(b) Nonintegration

Figure 5: The classification accuracy of the experiment was compared

V. B. 3) Stability comparison of models

In order to further validate the stability of this paper's model in human health status information adoption prediction, this part uses 10-fold cross-validation, which is run 10 times consecutively on the data set. The AUC values are calculated, and the comparison of classification accuracy under multiple experiments is shown in Fig. 5 (Fig. a is integrated and Fig. b is non-integrated). Combined with the figure, it can be seen that the corresponding AUC curve of this paper's method has the smallest fluctuation compared to a single base classifier, which indicates that the predictive model of this paper's method has better stability. Meanwhile, when the base classifiers are integrated learning, most of the AUC curves remain above the model, which is of some reference significance. When the base classifier is non-integrated learning, the model is always located at the top, and the effect performs well. It shows that the method of this paper can well improve the accuracy of information adoption, and the comparison of the left and right graphs can be found that the fluctuation amplitude of integrated learning is lower than that of non-integrated learning, which indicates that the stability of integrated learning in the adoption of information on the human health status is generally stronger, and the prediction accuracy is generally higher than that of non-integrated learning.

VI. Conclusion

The intelligent management method of human health status based on integrated learning shows better prediction performance than the traditional single algorithm. Aiming at the multidimensional and complex characteristics of health data, a comprehensive assessment system is constructed from five dimensions of physical health and mental health through feature engineering and 20 key features are screened out, which effectively deals with the feature redundancy problem. The hybrid classifier model based on Stacking realizes the complementary advantages of multiple algorithms, and the optimal model has an AUC of 0.97085, an accuracy of 0.93553, and an F1-Score of 0.93519, which is significantly better than a single integrated learning algorithm. The dynamic weight estimation algorithm enhances the adaptability of the model to changes in health status by fusing long-term historical data with short-term adjacent data. 10-fold cross-validation tests show that the proposed method has the least volatility, reflecting strong stability and generalization ability. The secondary integration of integrated learning on the basis of non-integrated learning effectively improves the model accuracy, especially the enhancement of non-integrated learning is more significant. Extracted by algorithm preference, it is found that the optimal set of algorithms can achieve an AUC value of 0.97287 and an accuracy rate of 0.93149.This method provides technical support for the comprehensive assessment and accurate prediction of the human health condition, and is of great practical value for the construction of personalized health management system.

References

[1] Nguyen, K. T., Medjaher, K., & Tran, D. T. (2023). A review of artificial intelligence methods for engineering prognostics and health management with implementation guidelines. Artificial Intelligence Review, 56(4), 3659-3709.

[2] Chen, M., & Decary, M. (2020, January). Artificial intelligence in healthcare: An essential guide for health leaders. In Healthcare management forum (Vol. 33, No. 1, pp. 10-18). Sage CA: Los Angeles, CA: Sage Publications.

[3] Kobrinskii, B. A., Grigoriev, O. G., Molodchenkov, A. I., Smirnov, I. V., & Blagosklonov, N. A. (2019). Artificial intelligence technologies application for personal health management. IFAC-PapersOnLine, 52(25), 70-74.

[4] Moafa, K. M. Y., Almohammadi, N. F. H., Alrashedi, F. S. S., Alrashidi, S. T. S., Al-Hamdan, S. A., Faggad, M. M., ... & Al-Anzi, A. K. (2024). Artificial Intelligence for Improved Health Management: Application, Uses, Opportunities, and Challenges-A Systematic Review. Egyptian Journal of Chemistry, 67(13), 865-880.

[5] Racine, E., Boehlen, W., & Sample, M. (2019, September). Healthcare uses of artificial intelligence: Challenges and opportunities for growth. In Healthcare management forum (Vol. 32, No. 5, pp. 272-275). Sage CA: Los Angeles, CA: SAGE Publications.

[6] Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V., & Biancone, P. (2021). The role of artificial intelligence in healthcare: a structured literature review. BMC medical informatics and decision making, 21, 1-23.

[7] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology, 2(4)

[8] Baldominos, A., De Rada, F., & Saez, Y. (2018). DataCare: Big data analytics solution for intelligent healthcare management. International Journal of Interactive Multimedia & Artificial Intelligence, 4(7).

[9] Amoon, M., Altameem, T., & Altameem, A. (2020). Internet of things sensor assisted security and quality analysis for health care data sets using artificial intelligent based heuristic health management system. Measurement, 161, 107861.

[10] Yang, Y., Siau, K., Xie, W., & Sun, Y. (2022). Smart health: Intelligent healthcare systems in the metaverse, artificial intelligence, and data science era. Journal of Organizational and End User Computing (JOEUC), 34(1), 1-14.

[11] Swamy, T. J., & Murthy, T. N. (2019, January). eSmart: an IoT based intelligent health monitoring and management system for mankind. In 2019 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-5). IEEE.

[12] Mujawar, M. A., Gohel, H., Bhardwaj, S. K., Srinivasan, S., Hickman, N., & Kaushik, A. (2020). Nano-enabled biosensing systems for intelligent healthcare: towards COVID-19 management. Materials Today Chemistry, 17, 100306.

[13] Torens, C., Adolf, F., Faymonville, P., & Schirmer, S. (2017). Towards intelligent system health management using runtime monitoring. In AIAA Information Systems-AIAA Infotech@ Aerospace (p. 0419).

[14] Tian, S., Yang, W., Le Grange, J. M., Wang, P., Huang, W., & Ye, Z. (2019). Smart healthcare: making medical care more intelligent. Global Health Journal, 3(3), 62-65.

[15] Masood, I., Wang, Y., Daud, A., Aljohani, N. R., & Dawood, H. (2018). Towards Smart Healthcare: Patient Data Privacy and Security in Sensor-Cloud Infrastructure. Wireless Communications and Mobile Computing, 2018(1), 2143897.

[16] Karunarathne, S. M., Saxena, N., & Khan, M. K. (2021). Security and privacy in IoT smart healthcare. IEEE Internet Computing, 25(4), 37-48.

[17] Nidhya, R., Kumar, M., Maheswar, R., & Pavithra, D. (2022). Security and privacy issues in smart healthcare system using internet of things. IoT-Enabled Smart Healthcare Systems, Services and Applications, 63-85.

[18] Keshta, I. (2022). AI-driven IoT for smart health care: Security and privacy issues. Informatics in medicine Unlocked, 30, 100903.

[19] Daniel Voskergian,Burcu Bakir Gungor & Malik Yousef. (2025). Engineering novel features for diabetes complication prediction using synthetic electronic health records. Frontiers in Genetics,16,1451290-1451290.

[20] Mohammed Amine Bouqentar,Oumaima Terrada,Soufiane Hamida,Shawki Saleh,Driss Lamrani,Bouchaib Cherradi & Abdelhadi Raihani. (2024). Early heart disease prediction using feature engineering and machine learning algorithms.Heliyon,10(19),e38731-e38731.

[21] Xuhui Zhang,Wenyu Yang,Wenjuan Yang,Benxin Huang,Zeyao Wang & Sihao Tian. (2025). Occupational Risk Prediction for Miners Based on Stacking Health Data Fusion. Applied Sciences,15(6),3129-3129.

[22] Abd Algani Yousef Methkal,Ritonga Mahyudin,Kiran Bala B.,Al Ansari Mohammed Saleh,Badr Malek & Taloba Ahmed I.. (2022). Machine learning in health condition check-up: An approach using Breiman’s random forest algorithm. Measurement: Sensors,23.