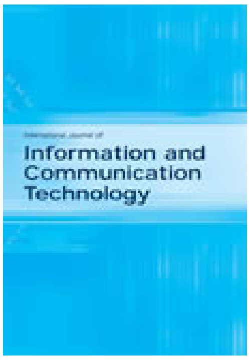


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Health management of young and middle-aged residents based on probabilistic optimisation BP neural network

Yuanyuan Wen

DOI: 10.1504/IJICT.2025.10071016

Article History:

Received:	27 March 2025
Last revised:	09 April 2025
Accepted:	10 April 2025
Published online:	27 May 2025

Health management of young and middle-aged residents based on probabilistic optimisation BP neural network

Yuanyuan Wen

College of Computer Science and Engineering,
Taizhou Institute of Science and Technology,
NJUST,
Taizhou, 225300, China
Email: yvonnewen1990@163.com

Abstract: The health status of the young and middle-aged population has a significant impact on the stable functioning of society. To address the issue of poor prediction accuracy in the current study, the BP neural network (BPNN) is first improved based on Bayesian optimisation (BO), and the parameter combination that maximises the conditional probability is selected to improve the fitting accuracy of the model. Then the factors affecting the health status are analysed, the influencing factors are decomposed and reconstructed using the improved variational modal decomposition (VMD) and fuzzy entropy algorithm, and the objective function is iteratively searched through BO probability theory to obtain the degree of the BPNN parameter that minimises the prediction error. Finally, corresponding health management suggestions are proposed for the prediction results. The simulation results indicate that the accuracy of the proposed method is 95.37%, which significantly improves the prediction accuracy.

Keywords: health prediction; Bayesian optimisation; probability theory; BP neural network; BPNN; variational modal decomposition; VMD.

Reference to this paper should be made as follows: Wen, Y. (2025) 'Health management of young and middle-aged residents based on probabilistic optimisation BP neural network', *Int. J. Information and Communication Technology*, Vol. 26, No. 14, pp.87–103.

Biographical notes: Yuanyuan Wen received her Master's degree from the Jiangsu University in 2014. She is currently a Lecturer at the Taizhou Institute of Science and Technology, Nanjing University of Science and Technology. Her research interests include big data analytics and artificial intelligence.

1 Introduction

As the backbone of society, the health status of young and middle-aged groups is not only related to the quality of life of individuals, but also profoundly affects the stability and development of society (Hu et al., 2024). However, with the accelerated pace of life and increased work pressure, the health problems of young and middle-aged residents are becoming more and more prominent, and the trend of the rejuvenation of chronic diseases and the prevalence of sub-health states have become social problems that cannot be ignored (Ma, 2018). Conventional health assessment systems are mostly based on empirical judgment and basic statistical analysis, making it difficult to gain precise insight into the subtle fluctuations and potential trends of individual health risks (Hong et al., 2014). With the booming growth of information technology, machine learning (ML) algorithms, as the core technology in the field of artificial intelligence, are penetrating into various industries with unprecedented momentum, and the field of health management has also ushered in a new opportunity for change. ML algorithms, with their powerful data mining, pattern recognition, and predictive analysis capabilities, are able to extract valuable information from massive and complicated health data, accurately identify individual health risk factors, construct highly personalised health prediction models, and show great potential in health management (Ramachandran, 2024).

Resident health management is to analyse personal health data and predict the risk of disease through predictive models, so as to achieve disease prevention and health management. Traditional research is based on the construction of predictive models of residents' health by statistical analysis methods (López-Martínez et al., 2020). Orang et al. (2024) analysed the influence indicators of population health through chaos theory and used the moving average method to predict influenza-like illnesses, but the prediction accuracy was not high. Chen et al. (2024) investigated the influencing variables of influenza disease by nonlinear regression analysis, and predicted the incidence trend of tuberculosis by auto-regressive moving average model (Asadi et al., 2012), but it could not respond well to the multi-factorial internal association. Zhang et al. (2017) used grey models to mine and analyse the incidence trends of residential diseases, but the prediction accuracy was not high. The traditional model is only applicable to exponential growth and short-term prediction, which leads to the model cannot maintain long-term stable prediction accuracy.

The ML algorithm learns the responsible nonlinear relationships between variables through an internal network, which leads to adaptive learning of internal features and greatly improves the prediction accuracy. Thenmozhi and Deepika (2014) collected physiological parameters such as ECG and blood pressure of the elderly through sensor technology, screened the main variables of the parameters using principal component analysis (PCA), and realised disease prediction through decision tree (DT) algorithm, but the computation consumes a lot of resources. Benkedjouh et al. (2015) decomposed the influence indicators of health by empirical modal decomposition (EMD) and used each intrinsic mode function (IMF) as an input to a support vector machine (SVM) to achieve health prediction, but the prediction error was large. Xing and Bei (2019) used different combinations of features and K-nearest neighbour (KNN) classification technique (Kataria and Singh, 2013) to build prediction models to improve the accuracy of predicting cardiovascular diseases. Nandy et al. (2023) learned the optimal representation of the training data by training a sparse self-coder (Qureshi et al., 2020), and then utilised an artificial neural network (ANN) to predict the learned health conditions, but the accuracy of the prediction was not high.

Compared to ML algorithms such as DT, SVM, KNN, etc., BP neural network (BPNN) over the training data learned feature representations that are able to predict unknown data with strong generalisation ability. Koshimizu et al. (2020) utilised PCA for the selection of the main influencing variables and BPNN for the prediction of hypertension development status due to meteorological factors. Tan et al. (2018) analysed numerous indicators affecting the health of the population and predicted the health status by BPNN optimised by genetic algorithm (GA) with a prediction accuracy of 87.15%.

Yan et al. (2022) improved the prediction accuracy by collecting the information of various physiological indicators of urban residents and using particle swarm optimisation of BPNN to achieve the prediction of the incidence of hypertension in residents. The BO algorithm models the objective function based on a probabilistic model, which is able to estimate the uncertainty of the prediction, converge quickly, and avoid the model falling into local convergence. Peng et al. (2019) processed the decomposition of impact indicators by VMD and used the decomposition result as the input of Bayesian optimisation (BO) BPNN, which significantly improved the prediction effect.

Based on the above in-depth analysis of the current state of research, it is clear that current health prediction management methods suffer from redundant variables and poor prediction accuracy. To cope with the above issues, this paper suggests a health management method for young and middle-aged residents based on probabilistic optimisation BPNN, and the main work of this method is reflected in the following aspects.

The main innovations and contributions of this work include:

1 Based on the BO algorithm and regularisation method, the BPNN is optimised, the hyperparameters of the model are smoothed by the regularisation term, and the network is simplified by constraining the size of the network weights and bias values, and the weights and bias values that can maximise the conditional probability are selected to improve the fitting accuracy of the model.

2 Construct the function model of residents' health index, analyse the factors affecting the health status, use the noise pre-processing combined with the decomposition error to adaptively determine the optimal number of modal decomposition layers of the VMD algorithm, and on this basis, decompose and reconstruct the influencing factors through the improved VMD algorithm and fuzzy entropy (FE), so as to reduce the complexity of the input variables.

3 Using the optimal weights and biases determined by the BO algorithm, the objective function is iteratively searched through probability theory to obtain the degree of network parameters that minimise the prediction error. Based on the prediction results, health management strategies for young and middle-aged residents are proposed to provide new ideas for improving the health of this population.

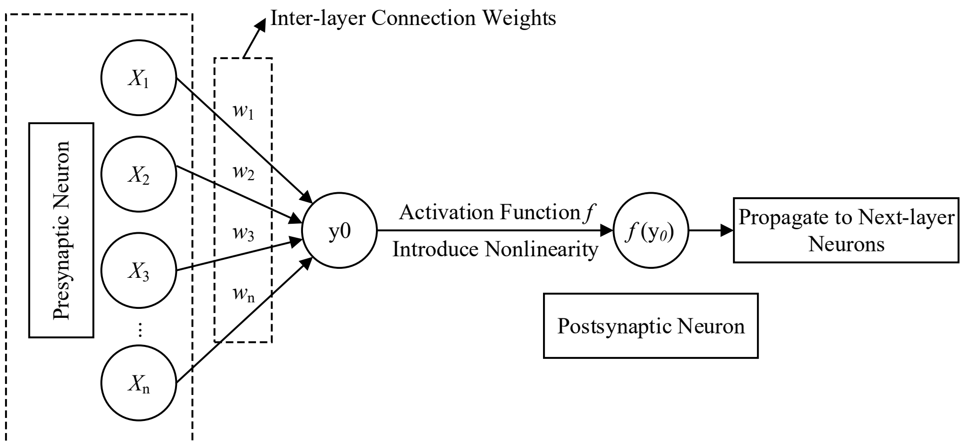
4 The experimental results show that the mean absolute percentage error (MAPE) of the proposed method is reduced by at least 12.07% compared with the other five methods, which has excellent prediction performance and provides strong support for health risk assessment, disease prediction, and personalised health management.

2 Relevant technologies

2.1 BP neural network

BPNN belongs to multi-layer feed-forward neural networks, which can learn and approximate complex nonlinear mapping relationships through the nonlinear activation function of the hidden layer, and solve the problems that are difficult to be handled by the traditional methods. The learning process of the BP NN includes the signal forward propagation and the error back propagation (Cui and Jing, 2019). In the forward propagation process, the training samples are transmitted from the input layer through the hidden layer to the output layer to obtain the actual output signal. If the output results do not meet the desired requirements, then enter the back propagation stage, through the constant correction of the neuron's weight threshold to reduce the prediction error, repeated cycles, so that the network output is constantly approaching the desired output. The BPNN feed-forward transmission process is shown in Figure 1.

Figure 1 The BPNN feed-forward transmission process



The input layer information x_j (j = 1, 2, …, m) enters to the implicit layer through the input port, the weights w_{jk} (k = 1, 2, …, q) are used to connect the input layer to the implicit layer, and the activation function of the implicit layer is set to $h_b(e)$ (b = 1, 2, …, q).

$$h_b(e) = g\left(\sum_{j=1}^{m} w_{jk} x_j\right) \tag{1}$$

On the basis of obtaining the implicit layer function $h_b(e)$, the output layer function y_p is obtained through the weights V_{kp} (p = 1, 2, …, n).

$$y_p = f\left(\sum_{k=1}^{q} V_{kp} h_b(e)\right) = f\left(\sum_{k=1}^{q} V_{kp}\, g\left(\sum_{j=1}^{m} w_{jk} x_j\right)\right) \tag{2}$$

2.2 BO algorithm

BO significantly reduces the number of calls to the objective function by actively selecting the 'most promising' points for evaluation by constructing a surrogate model (e.g., Gaussian process). PSO and GA require a large number of random samples or iterations (e.g., population evolution), which is inefficient, especially in high-dimensional or expensive evaluation problems. The core idea of BO is to describe the uncertainty of a function by constructing a probabilistic model and using this model to guide the search process (Victoria and Maragatham, 2021). The BO algorithm, as a global optimisation algorithm, always follows Bayes' theorem to fit the probability distribution of the objective function in finding the best hyperparameters of the model, i.e., the optimal solution of the objective function, as follows.

$$P\left(f \mid D_{1:n}\right)=\frac{P\left(D_{1:n} \mid f\right) P(f)}{P\left(D_{1:n}\right)} \tag{3}$$

where f is the black-box objective function, i.e., the unknown function between the hyperparameters to be optimised and the predicted effect of the model. $D_{1:n} = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ is the set of hyperparameter combinations x_n and corresponding function values y_n that have been collected, where $y_n = f(x_n) + \varepsilon_n$, ε_n are the collection errors. $P(D_{1:n}|f)$, $P(f)$, $P(D_{1:n})$, and $P(f|D_{1:n})$ are the likelihood probability distribution, prior probability distribution, marginal likelihood probability distribution, and posterior probability distribution of f, respectively.

2.3 Variational modal decomposition

VMD is a signal decomposition algorithm used to decompose a complex signal into multiple IMFs with specific centre frequencies and bandwidths. The core idea is to decompose the signal into a series of relatively smooth and physically meaningful sub-signals by means of variational optimisation methods (Lian et al., 2018). Compared with the EMD algorithm, the VMD considers a mechanism to suppress modal aliasing. It introduces a regularisation term to reduce the spectral overlap between different modal functions by constraining the bandwidth or frequency range of the modal functions, thus suppressing the modal aliasing phenomenon to a certain extent and making the decomposition results more accurate. The IMF is redefined in the VMD algorithm based on a rigorous mathematical theory, and is defined as an AM-FM signal, denoted by $u_k(t)$, and represented by the following equation.

$$u_k(t)=A_k(t)\cos\left(\phi_k(t)\right) \tag{4}$$

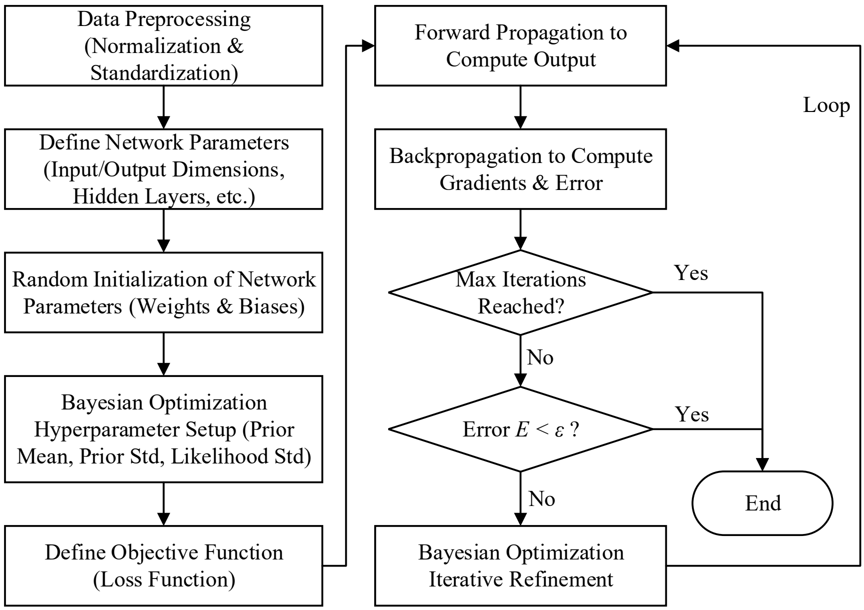
where $u_k(t)$ is a single harmonic signal, $A_k(t)$ is the instantaneous amplitude, $\phi_k(t)$ is the instantaneous phase, and the derivative of the instantaneous phase $\phi_k'(t)$ is the instantaneous frequency corresponding to $u_k(t)$.

3 Optimisation of BPNN based on BO and regularisation algorithms

3.1 Regularisation-based BPNN

The BPNN is prone to local optimisation rather than global optimisation during the training process, the prediction effect is not as good as the training effect, and they may be over-trained in order to achieve the preset conditions, i.e., the phenomenon of overfitting is generated. To address the above issues, this paper uses the regularisation method to generate the simplest model, but through the constraints on the size of the network weights and bias values to achieve the network simplest, based on which, the BO algorithm is used to optimise the weights and bias of the regularised BPNN (BOR-BPNN), in order to reduce the effect of the randomness of the initial state on the training effect and to improve the fitting accuracy of the model. The optimisation flow of BPNN is shown in Figure 2.

Figure 2 The optimisation flow of BPNN



Regularisation is a method to improve the generalisation ability of the BPNN by adding another function to the error sum of squares, which is used to penalise the complexity of the network. This paper adds a regularisation term containing the derivatives of the approximation function to smooth the resulting function. For BPNN, the error function on the training sample data is the sum of squares of the errors, as shown below.

$$F(x) = E_D = \sum_{q=1}^{Q} \left(t_q - a_q\right)^T \left(t_q - a_q\right) \tag{5}$$

where a_q is the output of the neural network when the input is x.

The above regularisation term can be written in the form of the sum of the squares of the network weights and biases as follows.

$$F(x) = \beta E_D + \alpha E_W = \beta \sum_{q=1}^{Q} \left(t_q - a_q\right)^T \left(t_q - a_q\right) + \alpha \sum_{i=1}^{n} x_i^2 \tag{6}$$

where the ratio α/β is used to control the effective complexity of the network solution. The larger the ratio, the smoother the BPNN response.

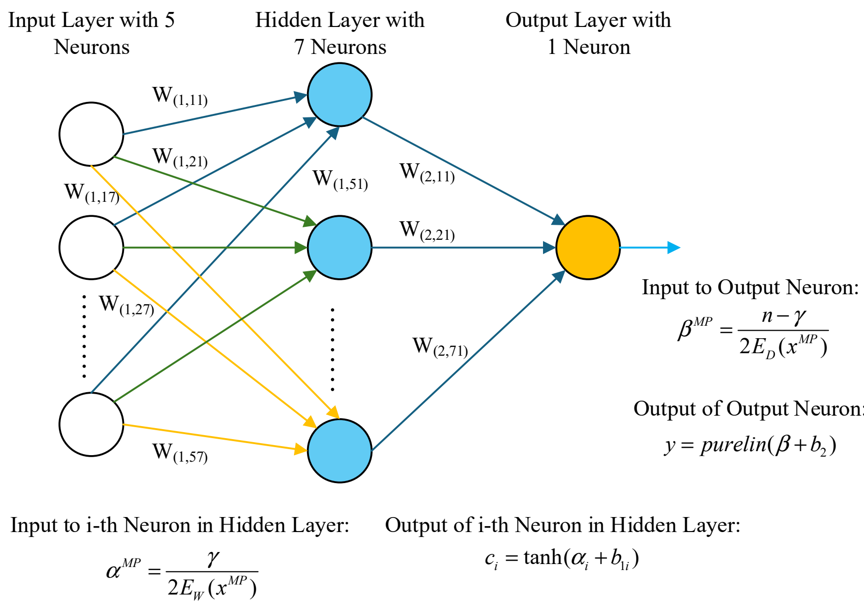
3.2 Optimal weights and bias search based on BO

After regularising the BPNN, the BO algorithm is used to optimise the regularised BPNN, and the optimised structure is shown in Figure 3. In this paper, assuming that the weights and biases of BPNN are random variables, for the given sample data, select the weights and biases that can maximise the conditional probability, and the BO algorithm calculates the following probabilities.

$$P(x \mid D, \alpha, \beta, M) = \frac{P(D \mid x, \beta, M) P(x \mid \alpha, M)}{P(D \mid \alpha, \beta, M)} \tag{7}$$

where x is a vector consisting of all the ownership values and bias values in the BPNN, D is the training sample data, α and β are the regularisation parameters; M is the number of neurons in each level of the BPNN.

Figure 3 The optimised structure of BPNN (see online version for colours)



Assuming that the error of the BPNN output obeys a Gaussian distribution, then the probability density $P(D|x, \beta, M)$ of the training sample data can be written as a likelihood function of the BPNN weights and bias, which gives the likelihood of the occurrence of the sample data, as follows.

$$P(D \mid x, \beta, M) = \frac{1}{Z_D(\beta)} \exp(-\beta E_D) \tag{8}$$

where $\beta = 1/(2\sigma_\varepsilon^2)$, σ_ε^2 is the variance of each element in the measurement noise ε, E_D is the sum of squares of the errors of the BPNN on the training sample data, and $Z_D(\beta) = (\pi|\beta)^{N/2}$. The weights and bias values of the BPNN are smaller values around 0, then a zero-mean Gaussian prior density can be chosen.

$$P(x \mid \alpha, M) = \frac{1}{Z_W(\alpha)} \exp(-\alpha E_W) \tag{9}$$

where $\alpha = 1/(2\sigma_w^2)$, σ_w^2 is the variance of each weight and bias, E_W is the sum of squares of the BPNN weights and biases, and $Z_W(\alpha) = (\pi|\alpha)^{n/2}$, n is the number of weights and biases in the BPNN.

Since the normalisation term $P(D|\alpha, \beta, M)$ is not a function of x, when seeking to maximise the weights and bias x of the a posteriori density $P(x|D, \alpha, \beta, M)$, the value of $P(D|\alpha, \beta, M)$ is not considered. Therefore, equation (7) can be rewritten as follows, where λ is the normalisation factor, $Z_F(\alpha, \beta)$ is a function of α and β, and $F(x)$ is the regularised error function as defined in Equation (6).

$$\begin{aligned} P(x \mid D, \alpha, \beta, M) &= \frac{\dfrac{1}{Z_W(\alpha)Z_D(\beta)} \exp\left(-(\beta E_D + \alpha E_W)\right)}{\lambda} \\ &= \frac{1}{Z_F(\alpha, \beta)} \exp\left(-F(x)\right) \end{aligned} \tag{10}$$

In summary, the regularisation error function can be derived from Bayesian statistics. Moreover, by maximising the a posteriori density, the possible values of the weights and biases x^{MP} can be found, which is equivalent to minimising the regularisation error function $F(x) = \beta E_D + \alpha E_W$. The regularisation error function can be derived from the Bayesian statistics. Since the error function has a quadratic form in the region near the minima, $F(x)$ can be expanded in a second-order Taylor series near its minima $F(x)$ (i.e., the point with gradient 0) as follows, where $H = \beta\nabla^2 E_D + \alpha\nabla^2 E_W$ is the Hessian matrix of $F(x)$, and H^{MP} is the estimate of the Hessian matrix at x^{MP}.

$$F(x) \approx F\left(x^{MP}\right) + \frac{1}{2}\left(x - x^{MP}\right)^T H^{MP}\left(x - x^{MP}\right) \tag{11}$$

By substituting the approximate terms in equation (11) into equation (10), the optimal values of and at the minima can be found, and by taking the derivative of equation (10) with respect to each of the logarithms of equation (10), and letting the derivative be 0, this paper have the following equation.

$$\begin{cases} \alpha^{MP} = \dfrac{\gamma}{2E_W\left(x^{MP}\right)} \\ \beta^{MP} = \dfrac{n-\gamma}{2E_D\left(x^{MP}\right)} \end{cases} \tag{12}$$

where $\gamma = n - 2\alpha^{MP} tr(H^{MP})^{-1}$ is the number of valid parameters and n is the number of all parameters in the BPNN. γ measures the number of weights and bias values in the BPNN that are effectively used to reduce the error function, ranging from 0 to n.

The minima of α and β are always changing. If the error function is always moved to the next smallest point as it moves across the surface, then the new estimate of the regularisation parameter will become more accurate. Eventually, a sufficient level of accuracy is reached such that the error function does not change significantly over the next iterations. At this point, the BPNN converges.

4 Health management of young and middle-aged residents based on BO BPNN

4.1 Analysis of the main factors affecting the health of young and middle-aged residents

To reduce the complexity of the input variables of the prediction model and improve the efficiency of health management, this paper first constructs a function model of the residents' health index, analyses the factors affecting the health status, and uses the improved VMD algorithm and FE to decompose and reconstruct the influencing factors to reduce the complexity of the input variables. Subsequently, the BO algorithm is used to determine the optimal weights and biases of the BPNN, and the objective function is iteratively searched through probability theory to obtain the network parameters that minimise the prediction error, thus improving the prediction efficiency. Finally, based on the prediction results, health management strategies for young and middle-aged residents are proposed to provide new ideas for improving the health of this population. The flow of the proposed health management approach is shown in Figure 4.

The health of young and middle-aged residents is influenced by many factors, and according to research (Zhou et al., 2014), changes in the natural patterns of their daily lives often signal physiological changes. Thus, dining, sleeping, exercise behaviour, and number of leisure activities were selected as the main reference factors affecting health status. The index function model for young and middle-aged residents was designed to be multi-parameter based, with fixed and variable components, depending on the influence of different factors on health status, as shown in equation (13).

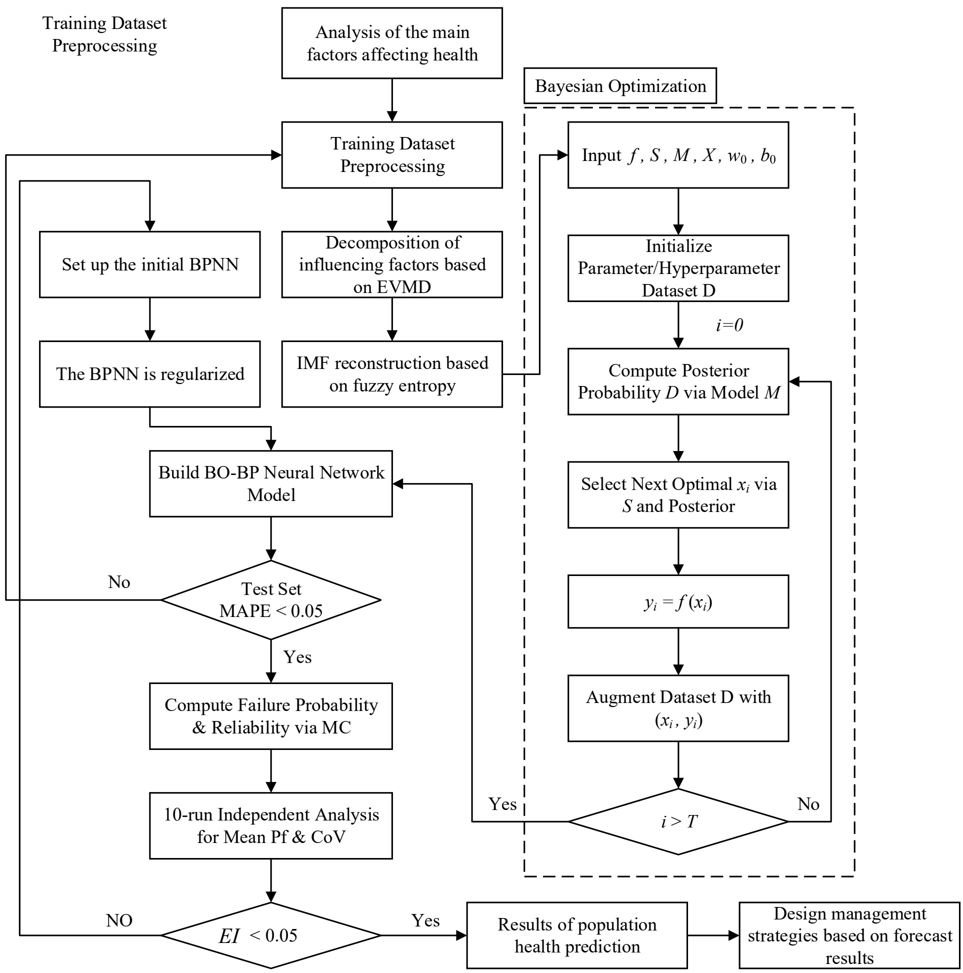
$$G(x)=\sum_{i=1}^{n} w_i g\left(x_i\right)+\sum_{i=1, j=1}^{n} \rho_{ij} f\left(x_i, x_j\right)+\xi \tag{13}$$

where $\sum_{i=1}^{n} w_i G(x_i)$ is the contribution component of each influencing factor to the health status of young and middle-aged residents, $\sum_{i=1, j=1}^{n} \rho_{ij} f(x_i, x_j)$ is the influence component, ξ is the adjustment error, w_i is the influence weight of the factor on the health status, $g(x_i)$ is the health index function of the different factors, which takes the value in the range of [0, 1], ρ_{ij} is the coefficient of interaction, and $f(x_i, x_j)$ is the function of the interaction between the different factors.

Due to the large number of factors in the activity, it is too complex to quantitatively define the interactions between each factor, and therefore activity interactions are treated as errors. Let $\xi'=\sum_{i=1, j=1}^{n} \rho_{ij} f(x_i, x_j)+\xi$, the health index function model of young and middle-aged residents can be obtained as follows. The value range of $G'(x)$ is [0, 1], different values represent different health states of residents, and ξ' is generally 0.

$$G'(x)=\sum_{i=1}^{n} w_i \frac{-\alpha_i\left(x_i-\delta_i\right)^2+\beta_i}{\beta_i}+\sum_{j=1}^{m} w_j \frac{2 \arctan x_j}{\pi}+\xi' \tag{14}$$

Figure 4 The flow of the proposed health management approach



4.2 *Influence factor decomposition reconstruction based on improved VMD and FE*

After constructing the health index function model for young and middle-aged residents, this paper utilises the improved VMD algorithm (EVMD) and FE (Wang et al., 2024) to decompose and reconstruct the influence indicators in order to reduce the input complexity of the prediction model. The performance of the VMD algorithm depends largely on the setting of the number of modes K. Before the variable VMD decomposition reaches the optimal decomposition K value, the invalid modes corresponding to the white noise component are removed, and a noise pre-processing process is realised. Iteration of variables the output frequency is directly determined by the input variables, as shown below.

$$\omega_k^{n+1} = \frac{\int_0^{\infty} \omega \times \left| \frac{\hat{y}(\omega) - \sum_{i \neq k} \hat{u}_i + \hat{\lambda}(\omega) \Big/ 2}{1 + 2\alpha(\omega - \omega_k)^2} \right|^2 d\omega}{\int_0^{\infty} \left| \frac{\hat{y}(\omega) - \sum_{i \neq k} \hat{u}_i + \hat{\lambda}(\omega) \Big/ 2}{1 + 2\alpha(\omega - \omega_k)^2} \right| d\omega} \tag{15}$$

where u is the modal component, w is the centre frequency, $\wedge$ is the Fourier transform, λ is the Lagrange multiplication operator, and α is the penalty operator.

The above equation is used to determine whether the white noise modal component contains redundant components. If present, the modal is retained, otherwise it is rejected as an invalid component of white noise. The decomposition energy error of the signal gradually increases when the signal is in the under-decomposition state and decreases when the optimal decomposition is reached. According to the method of setting the optimal number of decomposition layers by using the decomposition energy error, the optimal number of decomposition layers can be determined adaptively by taking the value of K when the decomposition energy error E_r is less than zero, so as to obtain the optimal IMF component $\{\chi_1, \chi_2, \ldots, \chi_n\}$ of the final index decomposition.

$$E_r = 2\sum_{i=1}^{K} E_{i+1} - \sum_{i}^{K} E_{i2}^2 \tag{16}$$

where E_{i+1} is the level $k + 1$ indicator energy and E_i is the level k indicator energy.

After obtaining $\{\chi_1, \chi_2, \ldots, \chi_n\}$, the subsequence is processed using equation (17) to obtain the embedding vector X_i, where n is the number of sequences, m is the number of embedding dimensions, and $\overline{\chi}$ is the mean of m consecutive χ_i's.

$$X_i = [\chi_i, \chi_{i+1}, \ldots, \chi_{i+m-1}] - \overline{\chi} \tag{17}$$

Subsequently the fuzzy similarity function $D_{i,j}^m$ is defined, the mean value of similarity $\varphi^m(r)$ is computed and finally the FE F_E of the sequence is computed.

$$D_{i,j}^m = \exp\left(-\frac{\left(d_{i,j}^m\right)^n}{r} \right) \tag{18}$$

$$\varphi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \frac{\sum_{j=1, i \neq j}^{N-m+1} D_{i,j}^m}{N-m} \tag{19}$$

$$F_E = \ln \varphi^m(r) - \ln \varphi^{m+1}(r) \tag{20}$$

where $d_{i,j}^m$ is the Chebyshev distance between $X(i)$ and $X(j)$; r is the similarity threshold.

The closer the F_E value is to 1, the more chaotic the subsequence is, and the higher the probability of generating new modes, and the closer the F_E value is, the higher the similarity is, and it has the same timing characteristics. Thus, the subsequences with close entropy values can be grouped and reconstructed, which can reduce the dimensionality of the data.

4.3 Health prediction and management of young and middle-aged residents based on BO BPNN

The input sequence after EVMD and FE processing is denoted as $x_1', x_2', \ldots, x_i'$. The health status of the young and middle-aged residents is denoted as excellent, good, moderate, and poor, i.e., y_1, y_2, y_3, y_4. The parameter optimisation problem of BPNN can be expressed by equation (21), where Q is the parameter to be optimised, L is the objective function, and Q^* is the ideal combination of hyperparameters.

$$Q^* = \arg\min L(Q) \tag{21}$$

$$L_{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{22}$$

where y_i is the actual value and $\hat{y}_i$ is the predicted value. Then the posterior probabilities of BOR-BPNN obtained by Bayes' theorem are as follows.

$$EI(x, D) = \int_{-\infty}^{\infty} \max\left(y^* - y, 0\right) p(y \mid x, D)dy \tag{23}$$

where y^* is the optimal value in the currently available sample and $P(y|x, D)$ is the posterior probability obtained from Bayes' theorem.

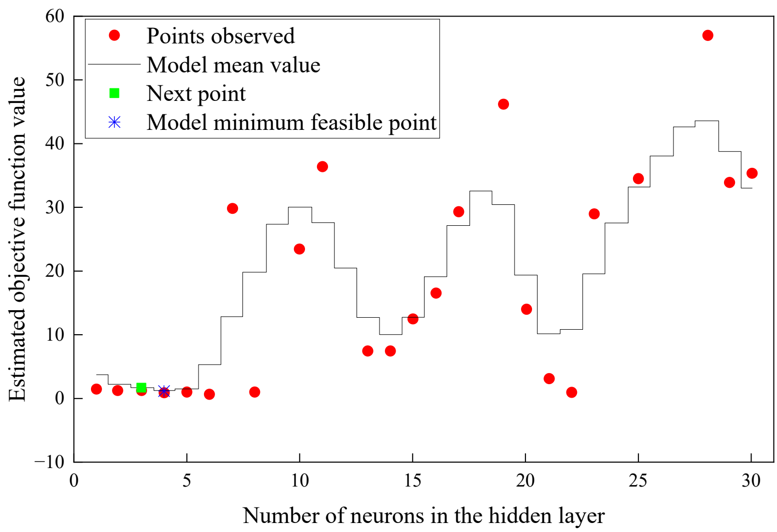
After obtaining the optimal weights and bias values of the BPNN through the BOR-BPNN proposed in Section 3, the objective function in the model that minimises the prediction error is determined, and the acquisition function S and the Bayesian probability model M are used to compute the target value, i.e., $y_i = f(x_i')$, by bringing $x_1', x_2', \ldots, x_i'$ into the objective function. Subsequently, the number of iterations T is set, the a posteriori probability $p(y|x, D)$ is calculated based on M and $x_1', x_2', \ldots, x_i'$, the next most promising combination of weights and bias w_i, b_i is selected using S, and w_i, b_i is brought into equation (2) to calculate the target value y_i, which serves as the historical information for updating M. Thus, the prediction error is continuously reduced until the objective function converges, and the predicted value of health index of the young and middle-aged residents with the smallest error is obtained.

Using the above prediction results, health management can be carried out for young and middle-aged residents. If the prediction shows that an individual has a higher probability of suffering from cervical spondylosis due to prolonged desk work, ergonomic desks and chairs can be purchased in advance to adjust the working posture and prevent the occurrence of the disease. At the same time, based on the prediction results, young and middle-aged residents can select suitable health insurance products to ensure that they have sufficient financial protection to cope with medical expenses when they may face health problems. In addition, based on the results of health forecasting, communities and enterprises can also carry out targeted health talks and intervention activities, such as providing mental health counselling courses for young and middle-aged people with high work pressure, helping them to relieve pressure and maintain a good psychological state, so as to realise all-round health management for young and middle-aged people and to enhance the overall health level.

5 Experimental results and analyses

The dataset of this paper is from WHO Global Health Observatory, 13,698 health data of young and middle-aged people aged 18–60 years old were crawled from this website using Scrapy, including basic personal information, physical health indicators, lifestyle, mental health indicators, medical history, etc., 80% of this dataset is used as a training set, 10% as a training set and 10% as a validation set. For BPNN, the batch size was 8, the number of iterations was 200 and the learning rate was 0.01 during the experiment. The prior mean of BO algorithm is 0, the prior standard deviation is 3, and the likelihood standard deviation is 0.7. The CPU in the experimental configuration is Intel (R) Core (TM) i5-7300, 2.5GHz, the graphics card is NVIDIA GeForce GTX1050, the RAM is 8.00GB, the operating system is Windows 10, and the software environment in which the algorithms run is MATLAB 2015b.

Figure 5 Experimental results of BO of BPNN (see online version for colours)



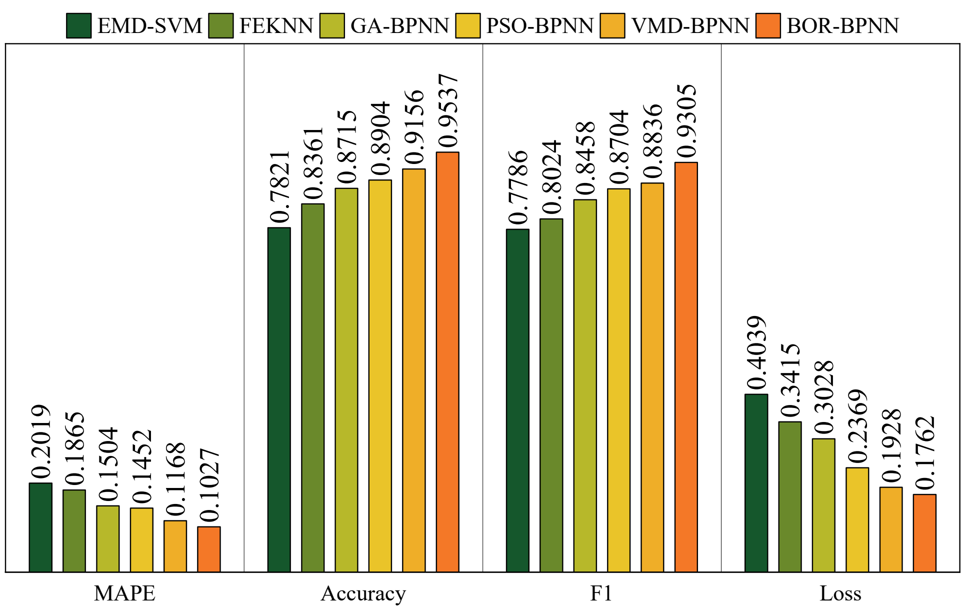
In this study, the BO algorithm is used to select the optimal number of neurons in the hidden layer, the range of optimisation variables is 1–30, and the optimisation objective is the mean square error between the predicted and actual values of the model, and the number of neurons with the smallest MAPE is selected through the iterative loops, and the results of the experiments are shown in Figure 5. The number of neurons in the range of 1–6 yields relatively small objective functions, and the number of neurons in the hidden layer is 4, which has the smallest mean square error, so the number of neurons in the hidden layer is 4 based on the results of the BO. After determining the optimal structure of the model, the model is re-trained by using the training set in order to learn the features and patterns of the data more adequately and to improve the model's generalisation ability.

The optimal structure of the model was determined and the model was retrained on the basis of that structure to realise the optimal model. Based on the comprehensive health quality score, the health level was categorised into four levels, where those with scores between 90–100 were recorded as excellent, those with scores between 80–89 were recorded as good, those with scores between 70–79 were recorded as moderate, and those with scores of 69 and below were recorded as poor. The partial assessment results obtained based on the grading of the health level and the errors are shown in Table 1. In the seven sets of sample data, the errors of the true and predicted values do not exceed 1, indicating that the true and predicted values of the health scores are well fitted and have a high prediction accuracy.

Table 1 Assessment results

True value of health quality score	*Predicted health quality score*	*Errors*	*Level*
84	84.3	0.3	Good
82	82.1	0.1	Good
91	90.5	0.5	Excellent
86	86.7	0.7	Good
94	94.2	0.2	Excellent
76	75.8	0.2	Moderate
87	86.5	0.5	Good

Figure 6 The prediction performance indexes of different health prediction methods (see online version for colours)



To further validate the effectiveness of the proposed prediction method BOR-BPNN, this paper selects EMD-SVM (Benkedjouh et al., 2015), FEKNN (Xing and Bei, 2019), GA-BPNN (Tan et al., 2018), PSO-BPNN (Yan et al., 2022), and VMD-BPNN (Peng et al., 2019) as the benchmark methods, and MAPE, accuracy, F1, and loss as the evaluation indexes, and a comparison of the prediction performance indexes of different health prediction methods is shown in Figure 6. The accuracy and F1 of BOR-BPNN are 95.37% and 93.05%, respectively, which are improved by 1.32%–15.19% compared with the benchmark method. The MAPE of BOR-BPNN is 0.1027, which is reduced compared with EMD-SVM, FEKNN, GA-BPNN, PSO-BPNN, and VMD-BPNN, respectively, by 49.13%, 44.93%, 31.72%, 29.27%, and 12.07%, indicating that the error between the true and predicted values of BOR-BPNN is small, which verifies the experimental results in Table 1. Comparing the Loss values of each method again, BOR-BPNN has the smallest loss of 0.1762, indicating that BOR-BPNN can speed up the convergence of the prediction model by optimising BPNN through BO, which reduces the function loss. In summary, BOR-BPNN has excellent prediction performance.

6 Conclusions

In today's society, the health problems of young and middle-aged residents are becoming more and more prominent with the accelerated pace of life and increased work pressure. The importance of health management as an important means of disease prevention cannot be overstated. To address the issue of high complexity of input variables in current health prediction management methods, this paper firstly optimises BPNN based on BO algorithm and regularisation method, and minimises the network by restricting the size of network weights and bias values, and selects weights and bias values that can maximise conditional probabilities to improve the fitting accuracy of the model. Then construct the function model of residents' health index, analyse the factors affecting the health status, and use EVMD algorithm and FE to decompose and reconstruct the influencing factors to reduce the complexity of input variables. Finally, using the optimal weights and biases determined by the BO algorithm, the objective function is iteratively searched through probability theory to obtain the degree of network parameters that minimise the prediction error. Based on the prediction results, health management strategies for young and middle-aged residents are proposed to provide new ideas for improving the health of this population. The experimental results show that the MAPE of the proposed method is 0.1027, which is reduced by 12.07%–49.13% compared with the baseline method, and opens up a new idea for personalised health management.

Acknowledgements

This work is supported by the Taizhou City Youth Science and Technology Talent Lifting Project in 2023.

Declarations

All authors declare that they have no conflicts of interest.

References

Asadi, S., Tavakoli, A. and Hejazi, S.R. (2012) 'A new hybrid for improvement of auto-regressive integrated moving average models applying particle swarm optimization', *Expert Systems with Applications*, Vol. 39, No. 5, pp.5332–5337.

Benkedjouh, T., Medjaher, K., Zerhouni, N. and Rechak, S. (2015) 'Health assessment and life prediction of cutting tools based on support vector regression', *Journal of Intelligent Manufacturing*, Vol. 26, pp.213–223.

Chen, Q., Zheng, X., Shi, H., Zhou, Q., Hu, H., Sun, M., Xu, Y. and Zhang, X. (2024) 'Prediction of influenza outbreaks in Fuzhou, China: comparative analysis of forecasting models', *BioMed Central Public Health*, Vol. 24, No. 1, p.1399.

Cui, K. and Jing, X. (2019) 'Research on prediction model of geotechnical parameters based on BP neural network', *Neural Computing and Applications*, Vol. 31, pp.8205–8215.

Hong, S., Zhou, Z., Zio, E. and Wang, W. (2014) 'An adaptive method for health trend prediction of rotating bearings', *Digital Signal Processing*, Vol. 35, pp.117–123.

Hu, L., Han, X., Chen, M. and Zhang, T. (2024) 'Association of waist circumference and BMI with premature death in young and middle-aged population', *Frontiers in Public Health*, Vol. 12, pp.21–27.

Kataria, A. and Singh, M. (2013) 'A review of data classification using k-nearest neighbour algorithm', *International Journal of Emerging Technology and Advanced Engineering*, Vol. 3, No. 6, pp.354–360.

Koshimizu, H., Kojima, R., Kario, K. and Okuno, Y. (2020) 'Prediction of blood pressure variability using deep neural networks', *International Journal of Medical Informatics*, Vol. 136, p.104067.

Lian, J., Liu, Z., Wang, H. and Dong, X. (2018) 'Adaptive variational mode decomposition method for signal processing based on mode characteristic', *Mechanical Systems and Signal Processing*, Vol. 107, pp.53–77.

López-Martínez, F., Núñez-Valdez, E.R., García-Díaz, V. and Bursac, Z. (2020) 'A case study for a big data and machine learning platform to improve medical decision support in population health management', *Algorithms*, Vol. 13, No. 4, p.102.

Ma, C. (2018) 'An investigation of factors influencing self-care behaviors in young and middle-aged adults with hypertension based on a health belief model', *Heart & Lung*, Vol. 47, No. 2, pp.136–141.

Nandy, S., Adhikari, M., Balasubramanian, V., Menon, V.G., Li, X. and Zakarya, M. (2023) 'An intelligent heart disease prediction system based on swarm-artificial neural network', *Neural Computing and Applications*, Vol. 35, No. 20, pp.14723–14737.

Orang, A., Berke, O., Poljak, Z., Greer, A.L., Rees, E.E. and Ng, V. (2024) 'Forecasting seasonal influenza activity in Canada – comparing seasonal auto-regressive integrated moving average and artificial neural network approaches for public health preparedness', *Zoonoses and Public Health*, Vol. 71, No. 3, pp.304–313.

Peng, W., Ye, Z-S. and Chen, N. (2019) 'Bayesian deep-learning-based health prognostics toward prognostics uncertainty', *IEEE Transactions on Industrial Electronics*, Vol. 67, No. 3, pp.2283–2293.

Qureshi, A.S., Khan, A., Shamim, N. and Durad, M.H. (2020) 'Intrusion detection using deep sparse auto-encoder and self-taught learning', *Neural Computing and Applications*, Vol. 32, No. 8, pp.3135–3147.

Ramachandran, K. (2024) 'Population health management through predictive analytics', *International Journal of Health Care Analytics*, Vol. 1, No. 1, pp.1–9.

Tan, X., Ji, Z. and Zhang, Y. (2018) 'Non-invasive continuous blood pressure measurement based on mean impact value method, BP neural network, and genetic algorithm', *Technology and Health Care*, Vol. 26, No. 1, pp.87–101.

Thenmozhi, K. and Deepika, P. (2014) 'Heart disease prediction using classification with different decision tree techniques', *International Journal of Engineering Research and General Science*, Vol. 2, No. 6, pp.6–11.

Victoria, A.H. and Maragatham, G. (2021) 'Automatic tuning of hyperparameters using Bayesian optimization', *Evolving Systems*, Vol. 12, No. 1, pp.217–223.

Wang, Y., Sun, J., Chen, S., Peng, W. and Zhang, D. (2024) 'Multi-scale reconstruction of rolling mill vibration signal based on fuzzy entropy clustering', *Vibroengineering Procedia*, Vol. 54, pp.22–27.

Xing, W. and Bei, Y. (2019) 'Medical health big data classification based on KNN classification algorithm', *IEEE Access*, Vol. 8, pp.28808–28819.

Yan, Y., Chen, R., Yang, Z., Ma, Y., Huang, J., Luo, L., Liu, H., Xu, J., Chen, W. and Ding, Y. (2022) 'Application of back propagation neural network model optimized by particle swarm algorithm in predicting the risk of hypertension', *The Journal of Clinical Hypertension*, Vol. 24, No. 12, pp.1606–1617.

Zhang, L., Wang, L., Zheng, Y., Wang, K., Zhang, X. and Zheng, Y. (2017) 'Time prediction models for echinococcosis based on gray system theory and epidemic dynamics', *International Journal of Environmental Research and Public Health*, Vol. 14, No. 3, p.262.

Zhou, J., Ru, X. and Hearst, N. (2014) 'Individual and household-level predictors of health related quality of life among middle-aged people in rural Mid-east China: a cross-sectional study', *BioMed Central Public Health*, Vol. 14, pp.1–10.