



Construction and optimization of traditional Chinese medicine constitution prediction models based on deep learning

ZHANG Xinge, XU Qiang, WEN Chuanbiao*, LUO Yue*

School of Intelligent Medicine, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan 611137, China

ARTICLE INFO

Article history

Received 10 July 2024

Accepted 02 September 2024

Available online 25 September 2024

Keywords

Traditional Chinese medicine (TCM) constitution

Deep learning

Constitution classification

Prediction model

Optimization research

ABSTRACT

Objective To cater to the demands for personalized health services from a deep learning perspective by investigating the characteristics of traditional Chinese medicine (TCM) constitution data and constructing models to explore new prediction methods.

Methods Data from students at Chengdu University of Traditional Chinese Medicine were collected and organized according to the 24 solar terms from January 21, 2020, to April 6, 2022. The data were used to identify nine TCM constitutions, including balanced constitution, Qi deficiency constitution, Yang deficiency constitution, Yin deficiency constitution, phlegm dampness constitution, damp heat constitution, stagnant blood constitution, Qi stagnation constitution, and specific-inherited predisposition constitution. Deep learning algorithms were employed to construct multi-layer perceptron (MLP), long short-term memory (LSTM), and deep belief network (DBN) models for the prediction of TCM constitutions based on the nine constitution types. To optimize these TCM constitution prediction models, this study introduced the attention mechanism (AM), grey wolf optimizer (GWO), and particle swarm optimization (PSO). The models' performance was evaluated before and after optimization using the F1-score, accuracy, precision, and recall.

Results The research analyzed a total of 31 655 pieces of data. (i) Before optimization, the MLP model achieved more than 90% prediction accuracy for all constitution types except the balanced and Qi deficiency constitutions. The LSTM model's prediction accuracies exceeded 60%, indicating that their potential in TCM constitutional prediction may not have been fully realized due to the absence of pronounced temporal features in the data. Regarding the DBN model, the binary classification analysis showed that, apart from slightly underperforming in predicting the Qi deficiency constitution and damp heat constitution, with accuracies of 65% and 60%, respectively. The DBN model demonstrated considerable discriminative power for other constitution types, achieving prediction accuracy rates and area under the receiver operating characteristic (ROC) curve (AUC) values exceeding 70% and 0.78, respectively. This indicates that while the model possesses a certain level of constitutional differentiation ability, it encounters limitations in processing specific constitutional features, leaving room for further improvement in its performance. For multi-class classification problem, the DBN model's prediction accuracy rate fell short of 50%. (ii) After optimization, the LSTM model, enhanced with the AM, typically achieved a prediction accuracy rate above 75%, with lower performance for the Qi deficiency constitution, stagnant blood constitution, and Qi stagnation constitution. The GWO-optimized DBN model for multi-class classification showed an increased prediction accuracy rate of 56%, while the PSO-optimized model had a decreased

*Corresponding author: LUO Yue, E-mail: louyue@cdutcm.edu.cn. WEN Chuanbiao, E-mail: 228237222@qq.com.

Peer review under the responsibility of Hunan University of Chinese Medicine.

DOI: [10.1016/j.dcmcd.2024.12.004](https://doi.org/10.1016/j.dcmcd.2024.12.004)

Citation: ZHANG XG, XU Q, WEN CB, et al. Construction and optimization of traditional Chinese medicine constitution prediction models based on deep learning. Digital Chinese Medicine, 2024, 7(3): 241-255.

accuracy rate to 37%. The GWO-PSO-DBN model, optimized with both algorithms, demonstrated an improved prediction accuracy rate of 54%.

Conclusion This study constructed MLP, LSTM, and DBN models for predicting TCM constitution and improved them based on different optimisation algorithms. The results showed that the MLP model performs well, the LSTM and DBN models were effective in prediction but with certain limitations. This study also provided a new technology reference for the establishment and optimisation strategies of TCM constitution prediction models, and a novel idea for the treatment of non-disease.

1 Introduction

The discipline of traditional Chinese medicine (TCM) constitutional medicine ^[1], a cornerstone of the theoretical framework of TCM, transcends mere categorization and description of individual constitutions. It profoundly elucidates the dynamic equilibrium between an individual's intrinsic physiological functions, psychological states, and the external environment, and how this balance shapes health status and disease susceptibility. The emphasis on "preventive treatment of diseases" (治未病) or treating before the onset of illness through constitutional conditioning, which resonates with modern medicine's concept of preventive medicine, emphasizing the importance of prevention over treatment.

The infusion of deep learning technology into TCM constitution prediction has invigorated the field. Its formidable data processing and feature extraction capabilities render predictions more precise and efficient. By constructing and optimizing deep learning models, researchers delve deeper into the intricate relationships between constitution and disease, enriching TCM constitutional research with abundant and insightful information.

Recent advancements in deep learning have significantly impacted the classification of TCM constitutions. (i) Based on tongue manifestations ^[2, 3]: deep neural network models are constructed to automatically learn tongue manifestation features like color, texture, and shape. Some use convolutional neural networks and other methods to improve classification accuracy ^[4]. Additionally, improved approaches like the metric classification method based on Triplet Loss address the small sample problem to enhance performance ^[5]. (ii) Based on facial images ^[6-8]: deep convolutional neural networks are employed to analyze facial images and extract related features associated with TCM constitution classification, providing a new, convenient identification means. (iii) Multimodal data fusion ^[9]: fusing multimodal data such as tongue manifestations and facial images, and using deep learning models to analyze them can compensate for single-modal limitations and boost classification accuracy. (iv) Combined with TCM theory ^[10]: deep learning models are combined with TCM theory to adjust their structure and parameters, and develop personalized diagnostic models in line with TCM thinking and rules. For

example, according to TCM theories such as Yin Yang and five elements, as well as Zang-Fu and meridians, the structure and parameters of deep learning models are adjusted to align with TCM's thinking mode and diagnostic rules. (v) Clinical application and development of auxiliary diagnosis systems ^[11]: research results are applied in clinical practice with auxiliary diagnosis systems developed. They collect relevant patient information and use deep learning models for analysis, improving the efficiency and accuracy of TCM constitution identification and helping with personalized plans.

The prediction of TCM constitution is of great significance. On one hand, it can accurately judge the constitution type to grasp the disease trend, formulate personalized treatment plans, and promote the modernization and internationalization of TCM theory. On the other hand, constitution prediction is crucial for health management. It is the cornerstone of preventing future health risks, providing customized suggestions, connecting physical and mental health, dealing with chronic diseases, and improving health literacy. At the scientific research level, it promotes the integration of multiple disciplines and provides a new perspective for precision medicine. To sum up, the necessity of constitution prediction is prominent in aspects such as early intervention, precision medicine, public health, mental health, and scientific and technological progress.

The research on building and optimizing TCM constitution prediction models based on deep learning, focusing on prediction analysis and model optimization, is expected to provide new ideas and methods for TCM constitution research and is of extraordinary significance for promoting the development of TCM and improving public health.

2 Data and methods

2.1 Data sources and preprocessing

2.1.1 Data sources Data were collected to identify the nine TCM constitutions, including balanced constitution, Qi deficiency constitution, Yang deficiency constitution, Yin deficiency constitution, phlegm dampness constitution, damp heat constitution, stagnant blood constitution, Qi stagnation constitution, and specific-inherited

predisposition constitution [12]. All data were ethically reviewed by the Medical Ethics Committee of the Affiliated Hospital of Chengdu University of Traditional Chinese Medicine (2022KL-024) [13].

Utilizing the 60-item Constitution in Chinese Medicine Questionnaires (CCMQ) [12] within the TCM constitution identification mini-program developed by our team, from January 21, 2020 to April 6, 2022, students from Chengdu University of Traditional Chinese Medicine were invited to complete their personal basic information and the entire questionnaire of 60 items during each of the 24 solar terms. Upon successful submission, they received their constitution identification results. These results were then subjected to a secondary manual verification process to ensure accuracy.

Inclusion criteria: (i) participants of any gender, aged between 18 and 60 years, who were either healthy or in a subhealth state; (ii) participants who were classified into one of the nine constitution types as defined in the "Classification and Judgment of TCM Constitution" scale, published by the China Association of Chinese Medicine in 2009 [12]; (iii) participants who were informed about and provided consent to participate in this study [13].

Exclusion criteria: (i) participants who were undergoing treatment for major illnesses such as cardiovascular and cerebrovascular diseases, diabetes, or cancer; (ii) participants who had non-major illnesses but required medication for maintenance; (iii) pregnant and lactating women, as well as individuals with psychiatric or neurological disorders [13].

2.1.2 Data preprocessing The numbers 1 to 9 were used to quantify the nine TCM constitutions, including the balanced constitution, Qi deficiency constitution, Yang deficiency constitution, Yin deficiency constitution, phlegm dampness constitution, damp heat constitution, stagnant blood constitution, Qi stagnation constitution, and specific-inherited predisposition constitution, respectively [12]. We quantified the 24 solar terms starting from the beginning of spring with numbers 1 to 24. The numbers 1 and 2 were used to quantify males and females, respectively. The number 0 was used to quantify the unfilled items in the CCMQ, and the numbers 1 to n were utilized to quantify the indicators of the items, respectively. According to the corresponding relationship between items on the questionnaire, the data characteristics of different constitutions were collated and sieved. In addition, data decentralization and normalization reduced the correlation between features so that other order-of-magnitude features were treated equally in model training.

The former process, decentering, involved subtracting the mean value of each feature from its dataset, resulting in a mean of zero for the data. This eliminated the influence of measurement scales on data analysis. Normalization further scaled the data to a specific, smaller range, such as [0, 1], using a proportional adjustment.

This ensured that each feature carried similar weight numerically, with a common approach being min-max normalization. These two processes collectively minimized feature correlations and enabled equal consideration of features with different magnitudes during model training, thereby significantly enhancing training efficiency and prediction accuracy. The specific computational formulas are shown in Equations (1) and (2):

$$x' = x - \text{mean} \quad (1)$$

$$x'' = \frac{(x - \text{min})}{(\text{max} - \text{min})} \quad (2)$$

In this context, x represents the original data, mean denotes the average value of the original data, x' is the data after the decentering process, min and max signify the minimum and maximum values of the original data, respectively, and x'' stands for the data after normalization.

2.2 Feature extraction

Utilizing variance thresholding and mutual information as feature extraction methods is a crucial step in data preprocessing. Variance thresholding involves setting a variance threshold to filter out features with low variance, i.e., those that exhibit minimal variation, thereby retaining features within the dataset that possess significant discriminative power. This approach contributed to removing redundant information and enhancing models' training efficiency. Conversely, mutual information quantifies the correlation between features and the target variable, enabling the selection of features with higher mutual information values relative to the target. This ensures that the model focused on the most relevant features of the target, ultimately improving prediction accuracy.

In this study, features were extracted for the nine constitutions, and the normalized results were compared to obtain the corresponding feature sets under the variance threshold method and the mutual information method. The feature sets obtained under the two methods were compared again to extract the repeated items, and the final feature set (the most representative and predictive feature set) was obtained. This process combines the variance threshold and mutual information methods to remove redundant information and ensures that the model focuses on the features most relevant to the target.

2.3 Model architecture design

We constructed three TCM constitution prediction models based on the deep learning algorithms: multi-layer perceptron (MLP), long short-term memory (LSTM) networks, and deep belief network (DBN) for different constitutions. Moreover, according to the experimental

results, the attention mechanism (AM), grey wolf optimizer (GWO), and particle swarm optimization (PSO) were applied to optimize the models. Then, we elaborated on the three TCM constitution prediction models and the optimization algorithms mentioned above.

2.3.1 MLP In the realm of statistics and machine learning, a significant milestone in biomimetics is the evolution of the perceptron. Originating from the MLP model proposed by MCCULLOCH et al. [14] in 1943, this model laid the foundation for the first mathematical framework based on physiological neurons.

MLP is adaptable to a wide range of tasks and data types. By adjusting its network structure and parameters, MLP can learn complex mapping relationships between inputs and outputs, making it an outstanding performer in tasks such as classification, regression, and clustering. This versatility underscores the rationale behind selecting MLP to construct the TCM constitution prediction model in this study.

MLP consists of three primary components. (i) The input layer: the nodes in this layer correspond one-to-one with the feature dimensions of the input data, serving solely as a conduit for transmitting external constitution data into the model. They do not participate in any computational processes; their function is purely to relay the input data forward. (ii) The hidden layer(s): the MLP comprises at least one hidden layer that receives the weighted sum (including the bias term) of the constitution data transmitted from the previous layer. It applies activation functions such as sigmoid, tanh, or ReLU to obtain its output. This process repeats through all hidden layers. (iii) The output layer: an appropriate activation function is applied upon receiving the computational results from the last hidden layer. Taking a binary classification problem as an example, the output layer is configured with two neurons, and the sigmoid function is chosen as the activation function. Based on the constitution feature data input from the input layer, this allows the model to determine whether to predict an output of “is XX constitution” or “not XX constitution.” During the forward propagation process, the values of weights and biases are fixed, which determines how the inputs influence the outputs.

2.3.2 LSTM The LSTM networks are a special architecture within the recurrent neural network (RNN) family. They were put forward by HOCHREITER in 1997 [15]. LSTM makes innovative optimizations based on the core features of RNN [16]. Inspired by the selective memory mechanism of the human brain, it has an input gate, a forget gate, an output gate and a cell unit. These gating units work together to selectively filter information [17], retain the parts relevant to the task and help with efficient processing of sequential data. They are located at the interfaces between other parts of the neural network and the memory unit. They optimize the performance of the neural network by controlling the flow and storage of

information instead of directly influencing the activities of other neuron nodes [18].

Its core computational process is divided into three phases [19]. (i) The forgetting phase. By combining h_{t-1} and x_t through the forget gate, a value between 0 and 1 is generated using the sigmoid function, determining which constitutional characteristic information in the memory cell state should be retained or forgotten. (ii) The selective memory phase. The sigmoid function controls the amount of information to be memorized, while the tanh function generates a candidate state \tilde{C}_t . This candidate constitutional characteristic information does not all enter the LSTM memory but are regulated by the input gate i_t , determining the extent to which constitutional characteristic information is retained for the next phase. Ultimately, the previously retained constitutional characteristic information ($f_t \cdot C_{t-1}$) and the newly required constitutional characteristic information after secondary adjustment and update ($i_t \cdot \tilde{C}_t$) are summed to update the learning and memory state of constitutional characteristics at the current time step, denoted as C_t . (iii) The output phase. The output gate utilizes the sigmoid function to determine the amount of constitutional characteristic information to be output. The tanh function standardizes the output constitutional characteristic vector, which is then multiplied by the output gate to obtain the hidden layer output h_t of the constitutional characteristic information at the current time step. Throughout this process, LSTM effectively handles long-term dependencies in constitutional characteristic sequence data by meticulously controlling information flow, enabling high-performance sequence modeling. A schematic diagram of the LSTM network architecture is shown in Figure 1.

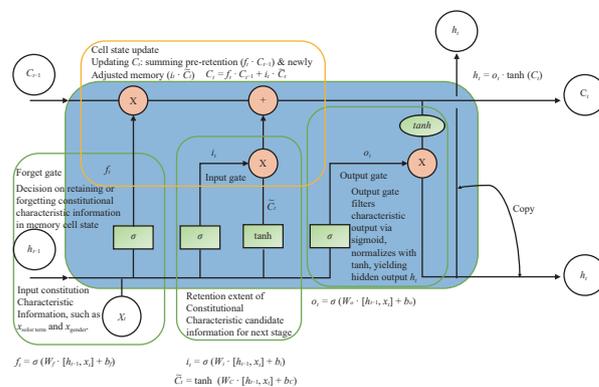


Figure 1 Schematic diagram of LSTM network architecture

The diagram, the σ function refers to the function. W_x (where $x = f, i, c, o$) represents the weight matrices, and b_x (where $x = f, i, c, o$) represents the bias parameters.

2.3.3 DBN The DBN is a multi-layered neural network architecture that is constructed from multiple layers of neurons, with each layer comprising both visible and hidden units. The visible units serve as the input to the

network, carrying the initial data feature information and forming the first gateway for data entry. In contrast, the hidden units play a pivotal role, tasked with feature extraction, deeply analyzing and processing the input data. Thus, they are often metaphorically referred to as feature detectors [20].

The core building block of DBN is the restricted Boltzmann Machine (RBM). RBM is a special form of the Boltzmann Machine (BM), an undirected graphical model that defines the energy state of a system based on the states of nodes and the connection weights between them in an undirected graph. The training of DBN proceeds layer by layer, with each layer's RBM being trained through unsupervised learning. This process gradually extracts deeper features from the data, ultimately constructing a robust feature representation model. DBN effectively learns the intrinsic laws and structures of the data through this layer-wise training, providing strong support for subsequent tasks. The schematic diagram specifically depicting the working principle of the DBN network structure for multi-class classification of landing points is shown in Figure 2 (similarly, for binary classification of landing points, the output function would be replaced from softmax to sigmoid function).

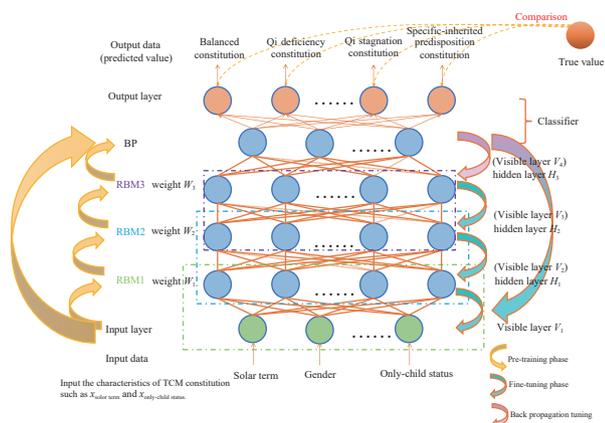


Figure 2 Schematic diagram of the DBN structure based on multi-class classification

In the input & layerwise training stage, the data enters the input layer, and the RBMs of each layer take the output of the RBM of the previous layer (the first layer is the original data) as input, and are trained with the contrastive scattering (CD or CD-k) algorithm. The extracted hierarchical features are propagated to the subsequent RBM layers to build a deep network, and each layer captures unique somatic feature information at the corresponding level. Feature extraction & propagation and deep network formation & characteristic capture are completed in this process. Next, replacement & prediction is performed to replace the last layer of the DBN with a multi-label output layer, and the algorithm is applied to predict multiple body labels. The algorithm is used to predict multiple body labels. Loss & optimization is performed using multi-class cross-entropy loss function, optimized by gradient descent/adam algorithm, and fine-tuned by using TCM constitution data. Finally, backpropagation & enhancement adjusts the weights and biases to improve the accuracy of the model's output of multiple body labels to achieve accurate prediction of TCM constitution.

2.4 Optimization algorithms

2.4.1 AM The AM originates from the imitation of the information processing mechanisms of the human visual system. Humans selectively focus on information due to the bottleneck in information processing and the differences among regions of the retina, so as to make efficient use of visual resources [21]. Inspired by this, when the computational and data processing capabilities of neural networks are limited, the attention mechanism allocates resources according to the importance of tasks, alleviates information overload, improves the efficiency and accuracy of task processing, and thus is widely used [22]. This study introduces the attention mechanism into the LSTM-based TCM constitution prediction model, aiming to optimize its ability to process constitution-related features while maintaining model performance. This integration is expected to better adapt to the complex and diverse scenarios encountered in TCM constitution prediction, enhancing the model's practical applicability and predictive accuracy.

2.4.2 GWO In GWO, there are different grey wolf categories in the initialized grey wolf population, namely alpha (α), beta (β), delta (δ), and omega (ω) [23]. The social status of different categories is different (Figure 3). During the hunting process, under the leadership of α -wolves, they tacitly track their prey and closely cooperate to make prey gradually incapacitated through pursuit, encirclement and harassment. Finally, attack decisively to complete the hunt [24].

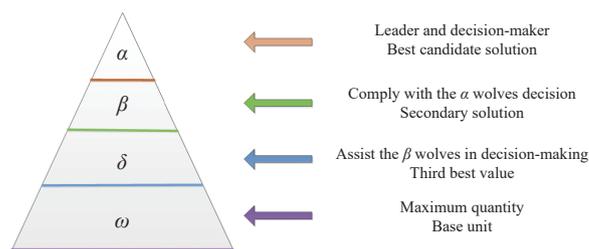


Figure 3 Grade distribution map of the grey wolf population

The GWO mimics the hunting behavior of grey wolf packs, utilizing the leadership structure (α , β , and δ) and the positional adjustments of the ω wolves [25]. It starts with a population of grey wolves randomly placed in the search space. Each wolf's performance is evaluated using a fitness function, selecting the top three to guide the search. The algorithm employs two coefficient vectors, A and C , to adjust the search step size and randomness. Vector A promotes exploration early on and refines search later, while vector C introduces diversity. The distances between ω wolves and the leaders are calculated to update the ω positions toward the optimal solution. As iterations progress, the positions of the α , β , and δ wolves are updated, influencing the ω wolves' positions to gradually approach the optimal solution. The algorithm

concludes when the maximum iterations are reached, outputting the position of the α wolf as the final optimized result.

2.4.3 PSO The PSO algorithm is a swarm intelligence optimization algorithm for evolutionary mechanisms based on the study of migration and foraging behavior of birds [26, 27]. In the standard application of the PSO algorithm, each individual in the swarm is transformed into a particle in a multidimensional search space, with their unique attributes characterized by three dimensions: position, velocity, and fitness [28]. The fitness quantitatively evaluates the particle performance and reflects the degree to which particles approach the optimal solution. The PSO algorithm can guide particles to approach the optimal solution more accurately through such a multidimensional description. The specific algorithm flow diagram is shown in Figure 4.

2.5 Evaluation indicators

It's essential to define the key definitions of false negative (FN), true negative (TN), false positive (FP), and true positive (TP). FN means wrongly predicting a sample as negative, TN is correctly predicting a sample as negative, FP refers to wrongly predicting a sample as positive, and TP indicates correctly predicting a sample as positive.

The performance evaluation indicators for the classification prediction model in this study are as follows:

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN})$$

$$\text{F1-score} = 2 \times (\text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall})$$

Receiver operating characteristic (ROC) curve: drawn with true positive rate (TPR) as the vertical axis and false positive rate (FPR) as the horizontal axis [29].

Precision-recall (PR) curve: drawn with precision as the vertical axis and Recall as the horizontal axis.

Area under the curve (AUC) value: represents the area under the ROC curve (AUC-ROC) or the area under the PR curve (AUC-PR), ranging from 0 to 1.

3 Results

3.1 Data distribution of the nine constitution

A total of 31 655 constitution identification datasets were obtained. The distribution of data of the nine constitution types is presented in Table 1.

3.2 Construction and implementation of each model

In the TCM constitution prediction, the relationship between the constitution and its influencing factors is non-linear. MLP can approximate and realize complex mappings. Therefore, this experiment constructed the MLP constitution classification prediction model based on the scikit-learn framework and showed it with the balanced constitution as an example (the same below). Preliminary preprocessing steps included data normalization, one-hot encoding for the target variable constitution type, the application of softmax function in the output

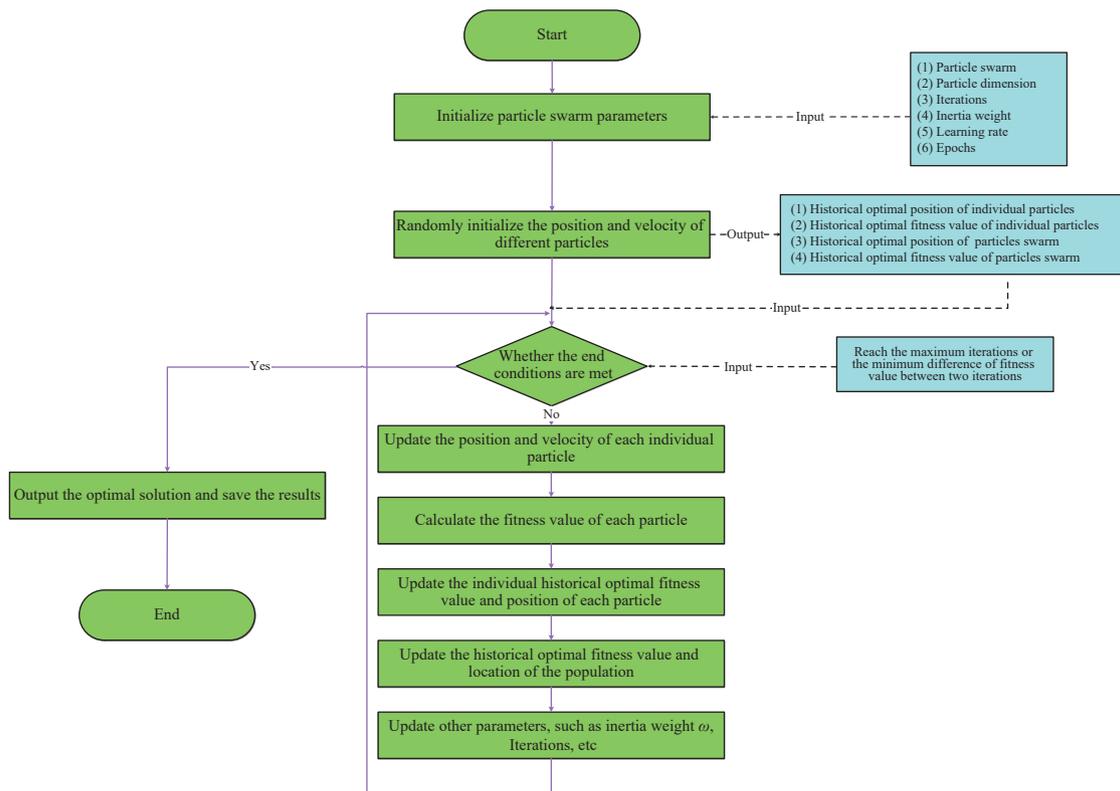


Figure 4 Flow diagram of PSO algorithm

Table 1 Data distribution of the nine constitution types

Constitution type	Data distribution
Balanced	10 959 (34.6%)
Qi deficiency	6 586 (20.8%)
Yang deficiency	2 837 (9.0%)
Yin deficiency	1 496 (4.7%)
Phlegm dampness	1 693 (5.4%)
Damp heat	2 070 (6.5%)
Stagnant blood	2 333 (7.4%)
Qi stagnation	2 730 (8.6%)
Specific-inherited predisposition	951 (3.0%)

layer, the utilization of a weight initializer for neural layers, the adoption of the Adam optimizer during model compilation, and the reduction of hidden layer parameters in the neural network and fine-tuning of model parameters through the integration of weight initializers and the Adam optimizer. Then, we compared the loss and accuracy values under different training conditions. Based on this analysis, we confirmed the selection of the MLP model utilizing both a weight initializer and the Adam optimizer for optimal performance. The multi-layer perceptron model structure was designed with five layers, including one input layer, three hidden layers, and one output layer (Figure 5A).

According to the feature extraction completed in section 2.2, in the input layer, this study selected 18 items, including solar terms, gender, whether it is only-child status, final educational attainment, mode of birth, feeding methods within four months after birth, allergy history, preference for light-taste food, preference for spicy food and love for eating spicy food, sleeping habits, exercise habits, emotional state, pressure situation, family atmosphere, overtime work situation, title-2, title-27, and title-53, as the input layer of MLP model to predict whether it was a balanced constitution or not. The expression is shown in Equation (3). The output layer was configured with two neurons, representing the classification results of a balanced constitution and a non-balanced constitution.

$$X_f = (x_{\text{solar term}}, x_{\text{gender}}, x_{\text{only-child status}}, \dots, x_{\text{title-2}}, x_{\text{title-27}}, x_{\text{title-53}}) \tag{3}$$

Multiple trials determined the number of neurons in the hidden layer and the ReLU function is chosen as the activation function. The softmax function is used in the output layer. The loss function is binary cross entropy. With Adam algorithm, we adjusted the learning rate to optimize the model performance. During the training process, parameters such as epochs and batch_size were adjusted several times, and different test_sizes to achieve the best model training performance.

Before building the LSTM network structure, the Z-score standardization method was applied to process

constitution data and transformed it into a standard normal distribution. Meanwhile, this study aimed to achieve a balanced 0 – 1 distribution in the TCM constitution types column by adjusting sample weights and employing undersampling techniques. Subsequently, an LSTM-based prediction model was constructed for TCM constitution prediction (Figure 5B). The input layer received the data and transformed the dimensions, the LSTM layer served as an intermediate layer to extract key features and applied the dropout technique to prevent overfitting (Figure 5C) [30]. Finally, the output layer generated the prediction results.

There were numerous parameters of the LSTM model, and parameters that significantly affected performance, such as learning rate, which needed to be optimized through debugging and 10-fold cross-validation. The training and testing sets were split in 9 : 1, 8 : 2, and 7 : 3 ratios, respectively. During the training process, the model underwent forward propagation, loss computation, sample weight calculation, cost-sensitive learning, back-propagation, and optimization, with the model being evaluated at the end of each training epoch to output the precision, recall, and F1-score. The sample weights were computed using the torch.tensor method, and cost-sensitive learning was implemented through the torch.mean method, considering the sample weight parameters specified in Table 2 for the nine constitution datasets.

Two models were constructed for multi-class and binary classification problems. The former used numbers from 0 to 8 to sequentially quantify different TCM consti-

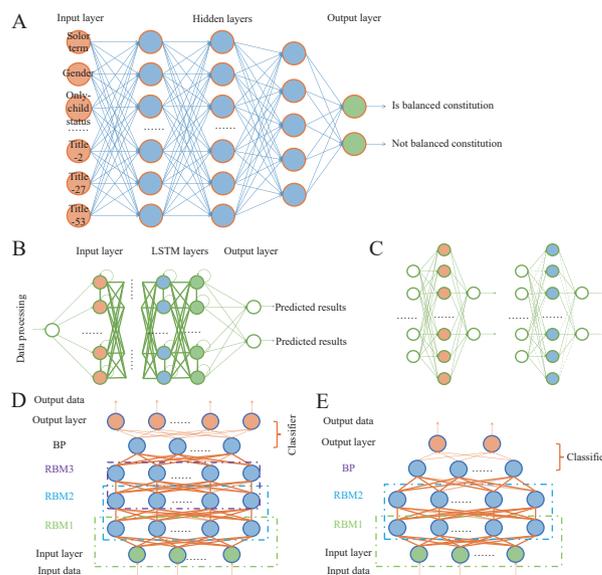


Figure 5 Construction and implementation of each model

A, network structure of MLP for TCM constitution classification prediction model (balanced constitution). B, network structure of LSTM for TCM constitution classification prediction model. C, a comparison of the differences before and after using the dropout technique in the hidden layer of the LSTM model. D and E, multi-label and two-label classification DBN model network structure diagram, respectively.

Table 2 Sample weight parameters for the nine constitutions datasets

Constitution type	Sample weight parameter (m, n)
Balanced	(1.00, 2.00)
Qi deficiency	(0.82, 3.82)
Yang deficiency	(0.80, 9.20)
Yin deficiency	(0.75, 20.25)
Phlegm dampness	(0.85, 17.85)
Damp heat	(0.63, 14.37)
Stagnant blood	(0.97, 25.03)
Qi stagnation	(0.62, 10.62)
Specific-inherited predisposition	(0.70, 32.30)

tutions to avoid index overbounds, while the latter adopted Z-score normalization to standardize the data. Both types of models incorporated the undersampling technique to balance the data distribution. Besides, the former (Figure 5D) trained RBM layer by layer to capture nonlinear relationships, introduced the dropout technique and regularization terms to improve performance, and output the probabilities using the softmax function; the latter (Figure 5E) fine-tuned weights through back-propagation (BP) layers to optimize performance^[31] and used the sigmoid function to output probabilities.

3.3 Prediction performance analysis of each model

In this case, the precision, recall, F1-score, and accuracy were used to evaluate model performance. The output results of three TCM constitution prediction models are shown in Table 3 - 5.

Table 3 Performance evaluation of MLP classification prediction models for the nine TCM constitutions

Constitution type	Precision	Recall	F1-score	Accuracy
Balanced	0.773 8	0.777 9	0.773 9	0.777 9
Qi deficiency	0.707 1	0.794 0	0.721 1	0.794 0
Yang deficiency	0.864 2	0.909 3	0.875 8	0.909 3
Yin deficiency	0.911 8	0.945 8	0.927 4	0.945 8
Phlegm dampness	0.897 5	0.935 1	0.913 7	0.935 1
Damp heat	0.878 3	0.929 7	0.900 5	0.929 7
Stagnant blood	0.880 3	0.920 7	0.885 4	0.920 7
Qi stagnation	0.839 6	0.916 3	0.876 2	0.916 3
Specific-inherited predisposition	0.971 0	0.972 2	0.971 6	0.972 2

According to the MLP model, the true and predicted value sequence was calculated for each constitution, and 100 cases of data were randomly selected to draw the fitting curve (Figure 6A). The threshold of predicted value was set at 0.5. When y_{pred} is higher than 0.5, it is predicted to be XX constitution; on the contrary, it is non-XX constitution.

Table 4 Performance evaluation of LSTM classification prediction models for the nine TCM constitutions

Constitution type	Precision	Recall	F1-score	Accuracy
Balanced	0.672 8	0.679 1	0.676 0	0.780 2
Qi deficiency	0.669 0	0.562 7	0.611 2	0.627 5
Yang deficiency	0.667 8	0.705 3	0.686 0	0.676 1
Yin deficiency	0.655 4	0.717 8	0.658 2	0.670 5
Phlegm dampness	0.808 4	0.750 0	0.778 1	0.772 9
Damp heat	0.642 9	0.655 3	0.649 0	0.647 3
Stagnant blood	0.721 5	0.686 7	0.703 7	0.691 6
Qi stagnation	0.758 2	0.721 3	0.739 3	0.732 6
Specific-inherited predisposition	0.878 4	0.896 6	0.887 4	0.884 4

Table 5 Accuracy and AUC of the DBN two-label output model for the nine TCM constitutions

Constitution type	Accuracy	AUC
Balanced	0.780 0	0.851 8
Qi deficiency	0.650 0	0.701 8
Yang deficiency	0.740 0	0.789 9
Yin deficiency	0.840 0	0.906 7
Phlegm dampness	0.810 0	0.877 6
Damp heat	0.600 0	0.669 5
Stagnant blood	0.720 0	0.781 9
Qi stagnation	0.820 0	0.869 5
Specific-inherited predisposition	0.900 0	0.948 7

According to Table 3, applying the MLP model for TCM constitution prediction demonstrated outstanding classification performance on the dataset of the nine constitution types. Specifically, the model achieved high scores of critical evaluation metrics such as accuracy, recall, and F1-score, validating its effectiveness in constitution prediction. Notably, when applied to the specific-inherited predisposition constitution dataset, the model exhibited a remarkable precision rate of 0.9722, accompanied by an F1-score of 0.9716.

As shown in Figure 6B and 6C, the AUC value of the LSTM prediction model for the damp heat constitution was about 0.68, indicating room for improvements in the prediction performance of the damp heat constitution dataset.

DBN multi-label output model experiments compared the prediction effect under different test_sizes. Taking test_size = 0.3 as an example, this study randomly selected 100 cases of data to draw the fitting curve of the true value and the predicted value, as well as the error curve of the two, showing in detail the matching degree between the true value and the predicted value and the distribution of the error (Figure 6D and 6E).

The experimental results reveal that the accuracy and AUC of 9 : 1, 8 : 2, and 7 : 3 were 0.42 (0.834 4), 0.43

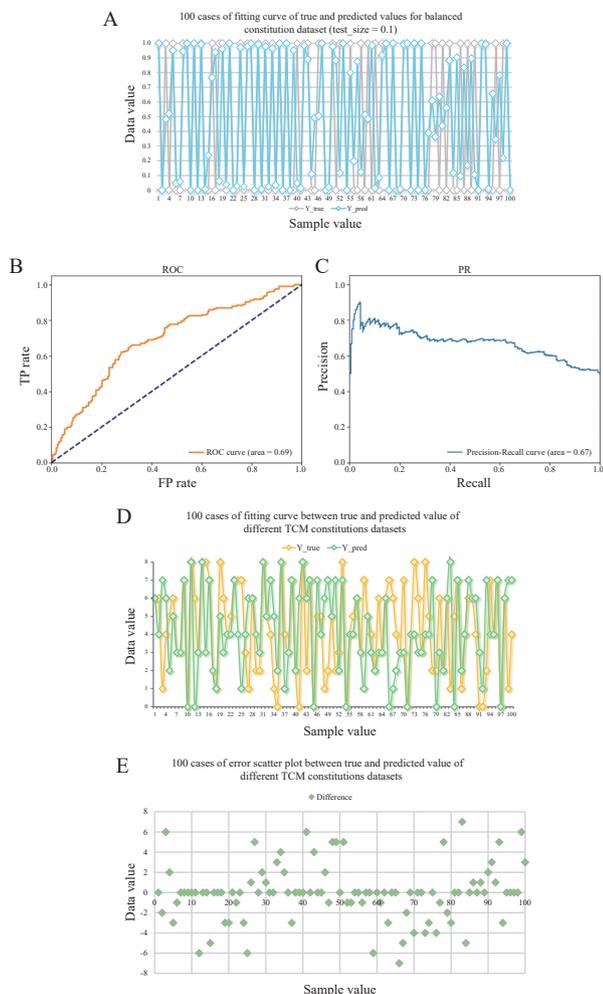


Figure 6 Analysis of the prediction results of each model A, fitting curve of true and predicted values for the balanced constitution dataset. The grey line is the true value, and the blue line is the predicted value. B and C, ROC curve and PR curve of damp heat constitution dataset (test_size = 0.1), respectively. D and E, fitting curve and error scatter plot of true and predicted value for the nine TCM constitutions datasets, respectively. The yellow line represents the true value, and the green line represents the predicted value. The difference calculation formula of the error value in Figure 6E: difference = $Y_{pred} - Y_{true}$. The horizontal axis represents the data index value, and the vertical axis represents the error value between the predicted and true values.

(0.8270), and 0.44 (0.8336), respectively. The classification performance of the model is adequate, but the recognition accuracy is not ideal.

The experiment conducted a comparative analysis of the performance of the DBN model under a multi-label output configuration, particularly focusing on various data splitting ratios, emphasizing test_size = 0.3. This investigation unveiled both the remarkable performance and underlying limitations of the model. By visually presenting the fitting curves between true and predicted values, along with error distribution plots, we clearly understood the model's robustness in the overall classification task. However, it was also observed that the model's precision in individual predictions had not yet to reach an optimal

level. This discovery prompted a deeper delve into the reasons behind such performance.

The DBN two-label output model used a specific-inherited predisposition constitution as an example to create a bar chart depicting the true value and the predicted value. This visual representation illustrates the data volume and proportions of each indicator (Figure 7).

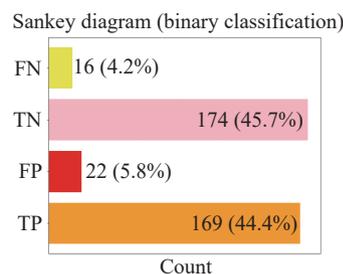


Figure 7 The true and predicted value of specific-inherited predisposition constitution dataset (test_size = 0.2)

The accuracy and AUC were used as indicators for model evaluation. The accuracy for each constitution ranged from 0.6 to 0.9, and the AUC was generally high, indicating that the model performed well in TCM constitution classification prediction (Table 5).

Table 5 presents a comprehensive performance evaluation results of a DBN with a two-label output model in identifying the nine TCM constitution types, encompassing accuracy and AUC. These metrics provide valuable insights into the model's varying performance across constitution types. Specifically, the balanced constitution, Yin deficiency constitution, phlegm dampness constitution, and Qi stagnation constitution exhibit relatively high accuracy and AUC, indicating the model's success in detecting these types. Conversely, the Qi deficiency constitution, Yang deficiency constitution, damp heat constitution, and stagnant blood constitution yielded lower accuracy and AUC, highlighting the challenges faced by the model in recognizing these complex or symptomatically overlapping constitution types.

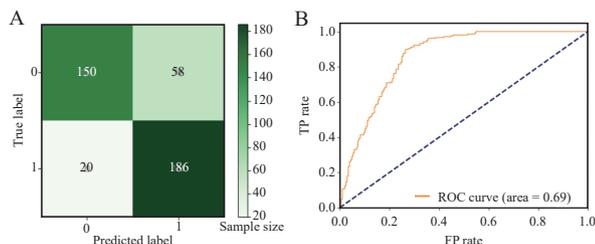
3.4 Comparison of models before and after optimization

3.4.1 Model optimization based on AM This study employed principal component analysis (PCA) for feature extraction to enhance the model's performance. Using damp heat constitution as an example, a longitudinal comparison of the optimized and pre-optimized LSTM model (Table 6) revealed that the pre-optimized model performed better in precision, recall, F1-score, and accuracy. Additionally, the AUC reached 0.8596, indicating excellent classification performance. The results presented in Table 6 underscore the substantial enhancement in prediction accuracy across various constitution types achieved by the LSTM-attention model compared with the baseline LSTM model.

Table 6 Longitudinal comparison results of LSTM models before and after optimization

Constitution type	LSTM	LSTM-attention
Balanced	0.780 2	0.818 4
Qi deficiency	0.627 5	0.616 8
Yang deficiency	0.676 1	0.767 6
Yin deficiency	0.798 0	0.753 3
Phlegm dampness	0.772 9	0.817 1
Damp heat	0.647 3	0.811 6
Stagnant blood	0.691 6	0.749 5
Qi stagnation	0.732 6	0.745 4
Specific-inherited predisposition	0.884 4	0.874 3

To further analyze the errors, this study plots the optimized confusion matrix and ROC curve (Figure 8). The results indicate that the false positive class rate of the optimized model has decreased to 18.8%, resulting in more accurate identification of negative class samples. Additionally, the AUC increased to 0.86, a 19.8% improvement, holding a promise for more precise identification of both positive and negative class samples. In particular, it performed even better when dealing with unbalanced data. The ROC curve was closer to the top left corner, indicating a low FP class rate and a high TP class rate, significantly improving model performance.

**Figure 8** Confusion matrix and ROC curve of optimized damp heat constitution dataset

A, confusion matrix. B, ROC curve. The test samples collected before and after optimization are consistent.

3.4.2 Model optimization based on swarm intelligence optimization algorithm

This experiment employed the swarm intelligence optimization algorithm to optimize the DBN multi-label output model. Following PCA feature extraction, the model was trained by debugging hyperparameters such as the number of hidden layer units, the precision, recall, F1-score, and accuracy were used to assess the performance of the GWO-DBN and PSO-DBN models. This paper took `test_size = 0.1` as an example, and the evaluation results after optimization are shown in Table 7.

3.5 Error analysis results

According to Table 3, the balanced constitution confusion matrix was plotted for error analysis (Figure 9). It can

Table 7 Evaluation results of DBN multi-label output model optimized based on GWO and PSO algorithms

Algorithm	PCA extracted feature number	Precision	Recall	F1-score	Accuracy
		GWO	0.573 0	0.548 3	0.540 8
PSO	30	0.165 0	0.162 4	0.152 6	0.360 4
	55	0.181 9	0.185 6	0.179 9	0.369 5

be concluded from Figure 9A that the model had the problem of more FP in prediction, while both the accuracy and recall were low, indicating that the model has difficulties in identifying positive class samples. This may be related to the imbalance of data categories, where the overabundance of negative samples caused the model to favor negative class predictions.

According to the experimental evaluation results in Table 4, taking the Yang deficiency constitution, Yin deficiency constitution, and damp heat constitution as examples, the confusion matrix of the three datasets was drawn for error analysis. From the confusion matrix in Figure 9B - 9D, the model accuracy of the damp heat constitution dataset was relatively low. This means that among the samples predicted as positive classes by the model, the proportion of TP classes was relatively low. Recall reflects the model's ability to recognize positive class samples. The recall rate of the current model of the damp heat constitution dataset was low, indicating that the model missed many positive class samples. The model may need to be improved to increase the recognition rate of positive classes. Among them, the false alarm rates of the model for the Yang deficiency constitution, Yin deficiency constitution, and damp heat constitution datasets reached 33.1%, 20.9%, and 42.5%, respectively. That was the situation of wrongly predicting negative class samples as positive classes were relatively frequent, which may lead to unnecessary costs or interference in practical applications. Generally, the model accuracies of the Yang deficiency constitution, Yin deficiency constitution, and damp heat constitution datasets were relatively low. For the Yang deficiency constitution, the F1-score was 0.6860; for the Yin deficiency constitution, it was 0.6582; and for the damp heat constitution, it was 0.6490. This suggests an imbalance between the precision and recall, indicating the need for further model optimization.

The confusion matrix was also generated for error analysis (Figure 9E - 9G). The multi-label output confusion matrix depicted the classification performance under different `test_sizes`, and the color depth indicating the number of instances. For example, the two-label output model classified stagnant blood constitution into positive and negative classes (Figure 9H). As Figure 9E - 9H shows, the multi-label output model exhibited more

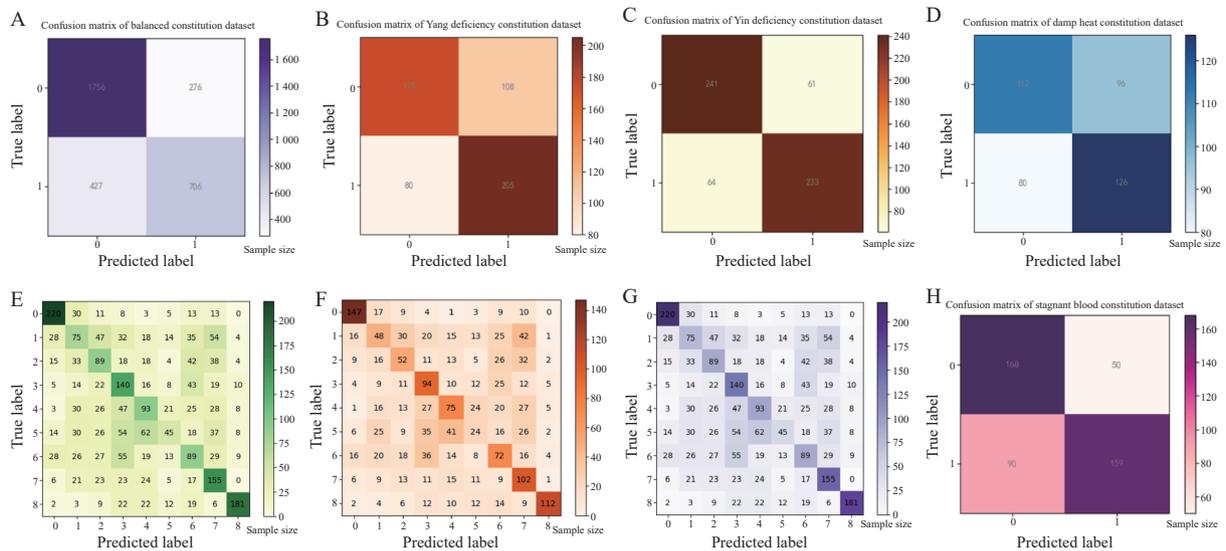


Figure 9 Error analysis of each model on different TCM constitutions datasets

A, confusion matrix for balanced constitution of the MLP model. B – D, confusion matrix for Yang deficiency constitution, Yin deficiency constitution, and damp heat dataset of the LSTM model, respectively. Label 1 is the positive class and 0 is the negative class. E – G, confusion matrix for the multi-label output DBN model case with test_sizes of 0.1, 0.2, and 0.3, respectively. H, confusion matrix for stagnant blood constitution dataset of the two-label output DBN model.

significant errors in dealing with rare and similar TCM constitutions, with the prediction error increasing as the test_size rose. This might be related to the uneven distribution of the dataset, the alignment between the model complexity and data volume, and parameter optimization.

3.6 Performance comparison of TCM constitution prediction of models

Classification prediction performance is the key index to measure the quality of the model. In model training, its classification and prediction performance might be affected by many factors.

As shown in Figure 10A, the three models showed different prediction performances in the face of different constitution types. The MLP model presented a high accuracy in multiple constitution types, especially in the prediction of Yin deficiency constitution, phlegm dampness constitution, damp heat constitution, stagnant blood constitution, and Qi stagnation constitution, and its prediction accuracy is more than 90%, showing its advantages in TCM constitution prediction tasks. In contrast, although the LSTM model performed reasonably well in the prediction of balanced constitution and specific-inherited predisposition constitution, with accuracy of 0.7802 and 0.8804, respectively, the accuracy of other constitution types were relatively low, especially in the prediction of Qi deficiency constitution and damp heat constitution, with accuracies of 0.6275 and 0.647 3, respectively. The DBN model for binary classification problem was outstanding in predicting Yin deficiency constitution and specific-inherited predisposition constitution, with prediction accuracies of more than 80%. Still, the

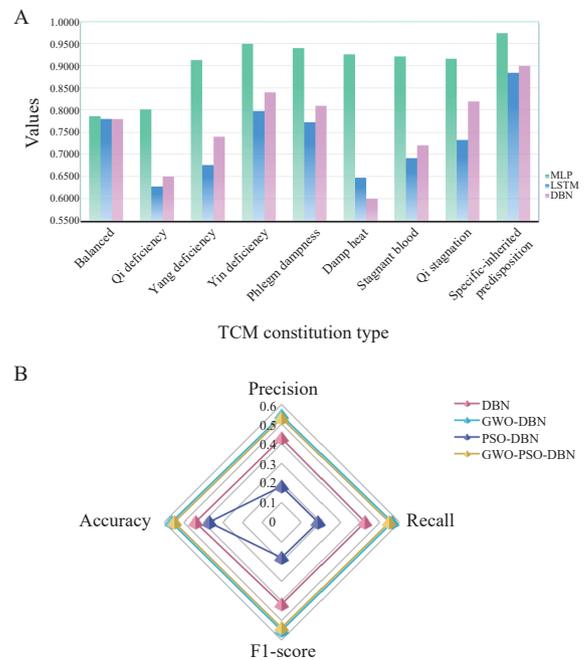


Figure 10 Comparative analysis of the performance before and after model optimization

A, comparison of accuracy rate distribution of the nine constitutions on different models. B, comparison and analysis of evaluation results of TCM constitution datasets on different models.

accuracies of the damp heat constitution and stagnant blood constitution were low, especially the damp heat constitution, with an accuracy rate of only 0.60. Based on the model prediction performance analysis in section 3.3, the DBN model for the multi-class classification problem showed a stable AUC of around 0.82 across different partitioning ratios. Yet, the model's prediction accuracy remained relatively low, falling below 0.50.

In the DBN model optimization, the GWO and PSO algorithms were introduced in this paper. By simulating the behavior of a grey wolf, the GWO algorithm effectively adjusted the key parameters, such as the number of hidden layer nodes and learning rate of the DBN model, increased the model prediction accuracy to 56%, and reduced the risk of overfitting. However, the PSO algorithm performed poorly in DBN model optimization, and the prediction accuracy rate dropped to 37%, which may be limited by its searchability and convergence speed to find the global optimal solution. Combining the two algorithms, the prediction accuracy of the GWO-PSO-DBN model was 54%, which did not exceed that of the GWO-DBN model, indicating that the GWO algorithm is more critical in the DBN model optimization (Table 7). Figure 10B compares the effects of different optimization algorithms on the DBN model.

4 Discussion

4.1 In-depth analysis of the underlying reasons for performance differences

In the process of predicting TCM constitutions, there are aspects that need improvement in the models. For example, the classification performance for some constitution types like the balanced constitution and Qi deficiency constitution still requires enhancement. The reasons behind this involve aspects such as model complexity and parameter settings. The complexity of multi-label classification itself increases the difficulty of the task. Besides, issues like data imbalance, limited sample size, and the sensitivity of deep learning models to hyperparameters all have an impact on the model's ability to predict minority categories, its generalization ability, and prediction accuracy. Moreover, the overlap of symptoms among different constitution types and the variability within the same constitution type also pose challenges for the model to make accurate judgments.

When optimizing the multi-label output tasks of the DBN, significant differences can be observed in the effects of the GWO algorithm and the PSO algorithm. On one hand, the GWO algorithm simulates the social structure of grey wolves and adopts a hierarchical search method, which enables it to balance global exploration and local search. During the iteration process, it can adaptively adjust the control parameters, making it more suitable for dealing with complex multi-modal problems. On the other hand, the PSO algorithm relies on the competition and cooperation among particles, as well as the individual and global historical optimal solutions to guide the search direction, and it is prone to falling into local optimal solutions. When dealing with the complex optimization challenges brought by multi-label output models, its global search ability is relatively weak, and the performance of this algorithm is greatly affected by parameter adjustment.

4.2 Theoretical explanation of the advantages and limitations of the models

4.2.1 LSTM-attention model (i) Advantages. By introducing the AM, this model can dynamically allocate weights according to the correlation between the input data and the prediction target, focus on key features, and filter out redundant information. Compared with the LSTM model before optimization, its average prediction accuracy has increased by 4.24%, and the generalization ability of the model in different data distributions and in the presence of noisy data has also been enhanced. (ii) Limitations. This model has problems such as high computational complexity and low operational efficiency, which will lead to an increase in training time, limited inference speed, and also increase the difficulty and cost of parameter adjustment and optimization. In addition, this model is sensitive to noisy data and has certain limitations in capturing long-term dependencies, and highly depends on high-quality data annotation. Otherwise, it will affect the learning effect and prediction accuracy.

4.2.2 The situation of GWO and PSO algorithms in optimizing the DBN model (i) GWO. Its advantages are reflected in its powerful global search ability and fast convergence speed, which are beneficial for determining the optimal parameter configuration in the initial stage of constitution data analysis, thereby improving the prediction accuracy of the DBN model. However, this algorithm faces the problem of parameter sensitivity. Given the diversity and complexity of constitution data, precise exploration of the parameter space is required. Otherwise, the algorithm is likely to fall into local optimal solutions, or the search process may be lengthy and inefficient. Moreover, when facing high-dimensional fitness datasets, the search space of this algorithm will expand dramatically, putting pressure on limited computing resources and increasing the risks of falling into local optimal solutions and soaring computing costs. Critically, the dynamic nature of fitness data imposes higher requirements on the GWO algorithm, requiring it to find the current optimal solution and predict and adapt to future data trends. However, there are still deficiencies in this regard at present. (ii) PSO. Although this algorithm has the potential for global optimization, its limitations are quite obvious in the context of constitution data analysis. Premature convergence is the primary problem. Especially when there are noise and outliers in the data, these factors can easily mislead the particle swarm to move in the wrong direction, thus prematurely terminating the search process. In addition, fine-tuning the parameters of the PSO algorithm is a complex and delicate task, and different parameter combinations have unpredictable impacts on the algorithm's performance, bringing many challenges to practical applications. More importantly, this algorithm tends to converge rapidly towards the current

optimal solution, which will lead to a rapid loss of population diversity. In the context of constitution data, the reduction of population diversity means that the algorithm may overlook potential better solutions, especially those hidden in complex correlation patterns. Therefore, its overall performance is often not as stable and reliable as that of the GWO algorithm.

4.3 Application value and limitations

These models and optimization algorithms have certain value in the prediction of TCM constitutions. For example, the LSTM-attention model helps to promote the development of personalized medicine and the modernization process of TCM practice. Selecting appropriate optimization algorithms can improve the prediction accuracy and generalization ability of the model, thus assisting in the classification and prediction of TCM constitutions.

However, each model and algorithm has certain limitations. For example, due to its high computational requirements and strict data requirements, the LSTM-attention model is limited in application scenarios that require quick responses. Whether used alone or in combination, the GWO and the PSO algorithms will face different challenges when dealing with the complexity of constitution data, thus affecting the accuracy and efficiency of data analysis.

4.4 Future improvement directions

4.4.1 Model level It is planned to optimize the overall performance of TCM constitution prediction by rebalancing training data, adjusting model parameters, integrating new features or algorithms, etc. For example, new feature extraction methods can be adopted to reduce the error rate of the model and improve its accuracy and reliability in practical applications. In addition, considerations can also be given to increasing the data volume, optimizing the model structure, adjusting hyperparameters, and applying advanced feature extraction methods, or using ensemble learning methods to optimize the classification accuracy of the DBN model.

4.4.2 Algorithm level Future research can further explore and optimize existing algorithms, combine them with advanced technologies such as feature selection and ensemble learning, and explore more advanced optimization algorithms or hybrid strategies. The focus should be on solving key problems such as effectively balancing the contributions of each part in the hybrid algorithm, dynamically adjusting the parameters of each algorithm at different stages to adapt to the dynamic characteristics of fitness data, and overcoming the inherent defects of a single algorithm, such as premature convergence, parameter sensitivity, and loss of diversity. Comprehensive consideration should be given to the

characteristics of data, the advantages and disadvantages of algorithms, and the requirements of practical applications to provide strong support for the in-depth mining and precise analysis of constitution data.

5 Conclusion

This study constructed three deep learning models (MLP, LSTM, and DBN) for TCM constitution prediction. Before optimization, MLP excelled in most constitution types, LSTM showed potential despite data lacking temporal features, and DBN had discriminative power but faced limitations. After optimization, the AM-enhanced LSTM improved accuracy, while DBN with GWO rose to 56% and GWO-PSO-DBN reached 54%. The unique features and limitations of each model guide future optimizations. This study provides a new idea for exploring the intelligent prediction and optimization of TCM body mass based on deep learning to enable accurate constitution prediction. Future efforts in model fusion, data handling and feature engineering will improve performance, enabling real-time health insights and personalised TCM care, which are crucial for “preventive treatment of disease” and improving public health.

Fundings

National Natural Science Foundation of China (81904324), and Sichuan Science and Technology Department Project (2022YFS0194).

Competing interests

The authors declare no conflict of interest.

References

- [1] WANG Q. TCM constitutional theory. Beijing: China Traditional Chinese Medicine Publishing House, 2021.
- [2] ZHOU H, HU GQ, ZHANG XF. Preliminary study on traditional Chinese medicine constitution classification method based on tongue image feature fusion. *Beijing Biomedical Engineering*, 2020, 39(3): 221-226.
- [3] HUANG EM, TAN MM, YUAN XL. An automatic discrimination system for traditional Chinese medicine constitution based on tongue image features. *Modern Computer*, 2023, 29(7): 116-120.
- [4] MA N, HAO XX, XING JF, et al. Application of convolutional neural network model in TCM treatment based on syndrome differentiation. *China Digital Medicine*, 2023, 18(4): 37-42.
- [5] SUN M, ZHANG XF. Tongue image classification based on TripletLoss metric. *Beijing Biomedical Engineering*, 2020, 39(2): 131-137.
- [6] HUAN EY. Research of facial constitution recognition based on deep neural networks. Gangdong: South China University of Technology, 2020.

- [7] YANG S. Construction of TCM constitution identification model based on tongue, face and inquiry information fusion. Shanghai: Shanghai University of Traditional Chinese Medicine, 2021.
- [8] ZHANG JY. A study on the physical identification and objectification of facial morphological features of five shaped people in traditional Chinese medicine. Changsha: Hunan University of Chinese Medicine, 2023.
- [9] HU JW. The deep learning of multimodal data for traditional Chinese medicine constitution identification. Shanghai: Shanghai University, 2021.
- [10] XIN JL. Research of health status identification algorithm based on TCM theory of state. Fuzhou: Fujian University of Traditional Chinese Medicine, 2020.
- [11] LI YT. Research on key technology of intelligent assistant tongue diagnosis. Xiangtan: Xiangtan University, 2020.
- [12] China Association of Chinese Medicine. Classification and Judgment of TCM Constitution (ZYYXH/T157-2009). World Journal of Integrated Traditional and Western Medicine, 2009, 4(4): 303-304.
- [13] LUO Y, JIANG LX, YANG SW, et al. Grey correlation analysis on influencing factors of Yang deficiency constitution. *Digital Chinese Medicine*, 2023, 6(2): 151-159.
- [14] MCCULLOCH WS, PITTS W. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 1943, 5(4): 115-133.
- [15] HOCHREITER S, SCHMIDHUBER J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735-1780.
- [16] PENG YQ. Research on SSE 50ETF price forecast and option strategy based on GA-LSTM model. Xi'an: Northwest University, 2021.
- [17] LI CY. Research on stock price prediction and quantitative stock selection based on CNN-LSTM. Xi'an: Northwest University, 2021.
- [18] GRAVES A, SCHMIDHUBER J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, 2005, 18(5/6): 602-610.
- [19] YAN B. Research on prediction method of foundation pit deformation by fusing auto-encoder depth feature extraction and LSTM network. Taian: Shandong Agricultural University, 2022.
- [20] YUE H. The feature extraction and classification of underwater acoustic signal based on deep learning methods. Changsha: National University of Defense Technology, 2017.
- [21] SINGH R, CHANDRA S, DHUSIA K, et al. Capacitating surveillance and situational awareness with measure of visual engagement using eyetracker. 2016 International Conference on Computing, Communication and Automation (ICCCA). IEEE, 2016: 1161-1165.
- [22] YU L. Research on CNN-LSTM air quality prediction based on attention mechanism. Dalian: Dalian Jiaotong University, 2023.
- [23] CHENG S, WANG R, WU GH, et al. Swarm intelligence optimization algorithms. *Journal of Zhengzhou University*, 2018, 39(6): 1-2.
- [24] ZHANG XF, WANG XY. Comprehensive review of grey wolf optimization algorithm. *Computer Science*, 2019, 46(3): 30-38.
- [25] MIRJALILI S, MIRJALILI SM, LEWIS A. Grey wolf optimizer. *Advances in Engineering Software*, 2014, 69(3): 46-61.
- [26] CHEN CC. Research on swarm intelligence optimization algorithm and its application in several problems of smart farming. Changchun: Jilin University, 2021.
- [27] KENNEDY J, EBERHART R. Particle swarm optimization. Proceedings of ICNN'95-International Conference on Neural Networks. IEEE, 1995: 1942-1948.
- [28] FU R. Application of LSTM model based on improved particle swarm optimization in stock price prediction. Jinan: Shandong University, 2022.
- [29] PAN XW. The construction and assessment study of article importance evaluation model based on the node attributes of article similarity network. Shenyang: China Medical University, 2022.
- [30] QIN QL, YAO CX, JIANG Y. Research on cerebrovascular disease prediction model based on the long short term memory neural network. International Conference on Smart Health. Cham: Springer, 2019: 247-256.
- [31] WANG R, YAN F, LU J, et al. COVID-19 trend forecasting by using dropout-LSTM model. *Journal of University of Electronic Science and Technology of China*, 2021, 50(3): 414-421.

基于深度学习的中医体质预测模型构建及优化研究

张新格, 许强, 温川飙*, 罗悦*

成都中医药大学智能医学学院, 四川 成都 611137, 中国

【摘要】目的 为满足个性化健康服务需求, 从深度学习角度挖掘中医体质数据特征并构建模型以探索预测新方法。**方法** 收集并整理 2020 年 1 月 21 日至 2022 年 4 月 6 日期间成都中医药大学学生按二十四节气划分的数据。这些数据用于识别 9 种中医体质, 包括平和质、气虚质、阳虚质、阴虚质、痰湿质、湿热质、血瘀质、气郁质和特禀质。利用深度学习算法, 构建多层感知机 (MLP)、长短期记忆网络 (LSTM) 和深度置信网络 (DBN) 中医体质预测模型。同时本文引入注意力机制 (AM)、灰狼优化算法 (GWO) 和粒子群优化算法 (PSO) 对以上 3 种模型进行优化。利用精确率、准确率、召回率和 F1 分数评估优化前后的模型性能。**结果** 该研究共分析了 31 655 份数据。(1) 优化前: MLP 模型除平和质和气虚质外的预测准确率均达 90% 以上; LSTM 模型预测准确率均达到 60% 以上, 表明其在中医体质预测任务中的潜力可能因数据缺乏显著时序特征而未被充分挖掘; DBN 模型在处理二分类问题时, 除在气虚质和湿热质的体质预测上稍显逊色, 预测准确率分别为 65% 和 60%, 其余体质的预测准确率和模型受试者工作特征 (ROC) 曲线下面积 (AUC) 分别达到 70% 以上和 0.78 以上, 表明模型具有一定的体质区分能力, 但其在特定体质的特征处理上存在局限, 模型性能仍有提升空间; 处理多分类问题时, DBN 模型的预测准确率不足 50%。(2) 优化后: 经 AM 优化后的 LSTM 模型预测准确率达 75% 以上, 但气虚质、血瘀质和气郁质除外; DBN 模型处理多分类问题时, 分别引入 GWO 和 PSO 算法优化后的模型, 前者预测准确率较优化前增至 56%, 后者预测准确率较优化前降至 37%。结合以上两种算法优化后的 GWO-PSO-DBN 模型预测准确率较优化前增至 54%。**结论** 本研究构建了 MLP、LSTM 和 DBN 模型来预测中医体质, 并基于不同的优化算法对其进行了改进。结果表明, MLP 模型具有较好的预测效果, LSTM 和 DBN 模型预测效果较好, 但存在一定的局限性。本研究为中医体质预测模型的建立和优化策略提供了新技术参考, 为中医治未病提供了新思路。

【关键词】 中医体质; 深度学习; 体质分类; 预测模型; 优化研究