

REVIEW

Open Access



Machine learning in TCM with natural products and molecules: current status and future perspectives

Suya Ma^{1†}, Jinlei Liu^{1†}, Wenhua Li^{2†}, Yongmei Liu¹, Xiaoshan Hui¹, Peirong Qu¹, Zhilin Jiang¹, Jun Li^{1*} and Jie Wang^{1*}

Abstract

Traditional Chinese medicine (TCM) has been practiced for thousands of years with clinical efficacy. Natural products and their effective agents such as artemisinin and paclitaxel have saved millions of lives worldwide. Artificial intelligence is being increasingly deployed in TCM. By summarizing the principles and processes of deep learning and traditional machine learning algorithms, analyzing the application of machine learning in TCM, reviewing the results of previous studies, this study proposed a promising future perspective based on the combination of machine learning, TCM theory, chemical compositions of natural products, and computational simulations based on molecules and chemical compositions. In the first place, machine learning will be utilized in the effective chemical components of natural products to target the pathological molecules of the disease which could achieve the purpose of screening the natural products on the basis of the pathological mechanisms they target. In this approach, computational simulations will be used for processing the data for effective chemical components, generating datasets for analyzing features. In the next step, machine learning will be used to analyze the datasets on the basis of TCM theories such as the superposition of syndrome elements. Finally, interdisciplinary natural product-syndrome research will be established by unifying the results of the two steps outlined above, potentially realizing an intelligent artificial intelligence diagnosis and treatment model based on the effective chemical components of natural products under the guidance of TCM theory. This perspective outlines an innovative application of machine learning in the clinical practice of TCM based on the investigation of chemical molecules under the guidance of TCM theory.

Keywords Machine learning, Deep learning, Traditional Chinese medicine, Natural products, Chemical components, Multidisciplinary intersection

Introduction

Machine learning, which involves learning relationships from data using computer science, has been successfully applied to solve complicated tasks such as computer vision, speech recognition, and natural language processing [1, 2]. The large amount of TCM data which utilized natural product (herbal medicine) to remedy disease produced from long-term clinical diagnoses, treatment and experiments, can be utilized for future research with machine learning. Machine learning verifies similarities among datasets by identifying characteristic regularities

[†]Su-ya Ma, Jin-lei Liu and Wen-hua Li contributed equally to this work

*Correspondence:

Jun Li

gamyylj@163.com

Jie Wang

wangjie0103@126.com

¹ Guang'anmen Hospital, China Academy of Chinese Medicine Sciences, Beijing 100053, China

² Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China



between input data and output results, and it has been applied for research on natural products, disease diagnosis and treatment, etc. [3]. Deep learning is an extension of machine learning that involves processing and validation of large training datasets between input and output units [4]. Deep learning has gained increasing importance for effective processing of large amounts of data and identifying patterns or functions hidden deep inside biological data. It has rapidly developed and successfully applied in many fields, including image recognition, robotics, speech recognition, and life sciences [5]. Deep learning uses different structural network models for different types of data and different application situations, and the primary models in deep learning include the convolutional neural network (CNN), Elman recurrent neural network (RNN), long short-term memory (LSTM), and generative adversarial network (GAN). Unlike CNN, RNN, and LSTM, which are deep neural network models, GAN is an unsupervised learning algorithm that learns by playing two deep neural networks against each other.

TCM is one of the oldest healthcare systems in the world and is being increasingly used as a complementary medicine system worldwide [6]. As a fully institutionalized part of Chinese healthcare, TCM is widely used with western medicine in China. Natural products which contain various chemical ingredients are employed to cure disease under the guidance of TCM theory. TCM theories such as the eight diagnostic principles to differentiate, the five elements theory, and the visceral manifestation theory can be collected by four traditional examination methods: looking, listening and smelling, asking, and touching, which could obtain pulse, face, tongue, urine, and stool information to provide essential information for diagnosis and natural products treatment. The process of diagnosis that guides treatment is called syndrome differentiation, which reflects the temporary state of a syndrome defined on the basis of the symptoms and signs identified by the four traditional examination methods. In this approach, wherein a clinical condition defined as a specific disease in western medicine can manifest in different syndrome elements in the same patient and may require varying treatment over time [7]. When this guiding TCM theory could be reproduced by machine learning, natural products which treat disease basing on chemical ingredients and molecules will be more powerful in remedying disease, that will bring major changes to human health and the quality of human life and save more lives. Although machine learning in TCM have been studied before [8–10], in the present study, we provide a systemic summary both of machine learning and its application in TCM, and proposed a promising future research direction. We summarized the principles and processes underlying deep

learning and traditional machine learning algorithms, analyzed the development and application of machine learning in TCM research, and proposed a promising research direction that integrates machine learning with TCM theory, natural products research, and computational simulation to provide an intelligent artificial intelligence diagnosis and treatment model based on the effective chemical components of natural products and molecules under the guidance of TCM theory.

Machine learning algorithms-deep learning

Convolutional neural network

CNN is a deep feedforward neural network with the characteristics of local connection and weight-sharing that uses a stack of convolution layers to extract features. It is widely used in image classification [11], facial recognition [12], semantic segmentation [13], object detection [14], and natural language processing [15]. The core idea of CNN involves a local receptive field, weight-sharing, and a pooling layer. The architecture of CNN consists of a sequence of layers that function as follows: when data is input into the CNN, the convolution layer extracts the features and the pooling layer aggregates the local features extracted by the convolution layer to obtain global features. Finally, the fully connected layer is combined to classify and output the results (Fig. 1). The convolution kernel (also named as filter) is utilized in the convolution layer with sizes of 1×1 , 3×3 , or 5×5 . Activation is applied after each convolution layer. The ReLU function, a mathematical formula that chooses the maximum of either z or 0 and is designated as $f(z) = \max(0, z)$, is often utilized. Then, the pooling layer is used to reduce location sensitivity, minimize the number of parameters and computation in the network, and to control overfitting [16]. The most common pooling function is the MAX pooling function, which uses the maximum value from each cluster of neurons at the prior layer to form a new neuron in the next layer. Other functions such as average pooling are also applicable [16]. The pooling layer reduces the input dimension of the subsequent network layer, reduces the size of the model, improves the calculation speed, improves the robustness of the feature map, and prevents overfitting. Three hyperparameters—depth, stride, and padding are used to control the size of the output data volume. The depth is consistent with the number of filters used, while the stride parameter reflects the number of pixels by which the filter moves each time it slides. The filling layer reflects the filling at the edge of the data volume, which can be filled with 0 or the mean value. CNN minimizes losses by adjusting the network parameters iteratively, and improves the accuracy of the network through frequent iterative training [17].

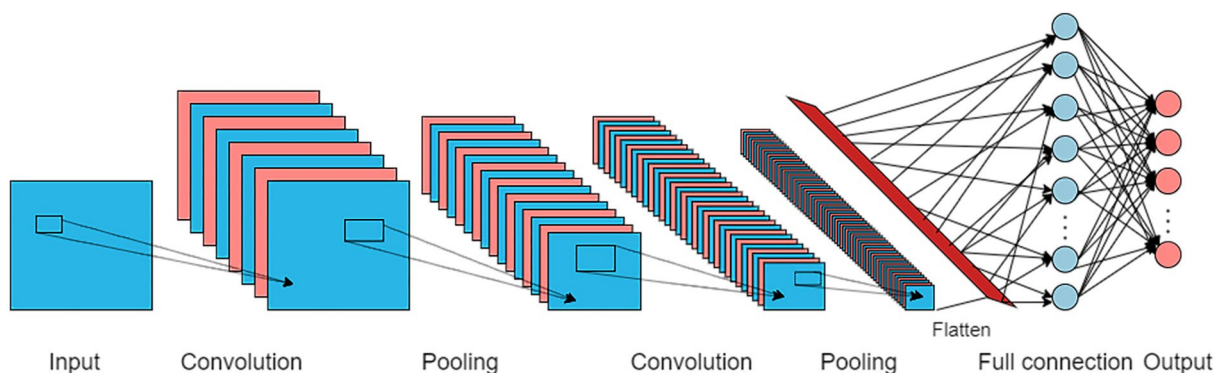


Fig. 1 The basic structure of CNN, including input, convolution, pooling, full connection and output layers

Elman recurrent neural network

The Elman RNN, also named as a simple recurrent network, was the first RNN among feedback neural networks and was specifically designed for processing time-dependent sequential data. The Elman RNN is often used in natural language processing [18]. It shows both current and past features of time series, adapts to the long-term historical changes in data, stores past information to solve context-dependent tasks, and provides predictions simultaneously with existing observations [2]. The Elman RNN consists of input, recurrent, hidden, and output layers. The standard connections of each layer which are similar to a feedforward network, are applied synchronously to propagate information from one layer to another by calculating a nonlinear function [19]. The input layer plays a signal transmission role, and the output layer plays a weighting role. The hidden and output layers usually employ the sigmoid nonlinear function as the activation function [20]. The recurrent layer is utilized to memorize the output value of the hidden layer at the previous moment, which can be regarded as a one-step delay operator. By the recurrent layer, the output of the hidden layer can self-connected to the input through delay and storage facilitated, and this self-connection makes the network can capture historical information. The addition of the internal feedback network increases the capacity of the network to handle dynamic information, thereby allowing dynamic modeling [21] (Fig. 2).

Long short-term memory network

The LSTM is an advanced variant of RNN with the capability of preserving long-term dependencies by using internal feedback [22]. Essentially, the LSTM layers prevent older information from gradually vanishing [23]. The LSTM is a popular RNN and has been successfully applied in many fields such as speech recognition, image description, and natural language

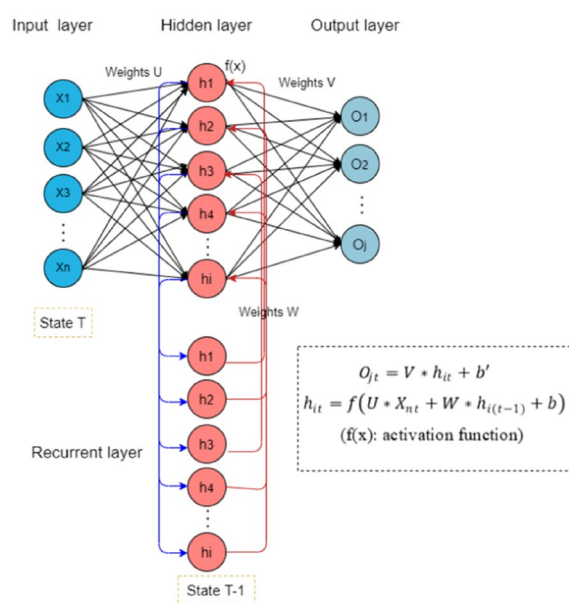


Fig. 2 The basic structure of Elman RNN, consists of input, recurrent, hidden, and output layers. U, V and W are the weights of input layer, output layer, recurrent layer separately. Parameter b represents bias term of hidden layer, b' represents bias term of output layer. Hidden layer: $h_{it} = f(U * X_{nt} + W * h_{i(t-1)} + b)$. Output layer: $O_{it} = V * h_{it} + b'$

processing. The LSTM can make use of gating mechanisms to mitigate gradient exploding and gradient vanishing when learning long-term dependencies [24]. This model introduces an intermediate type of storage using memory cells. A memory cell is a composite unit built from simpler nodes in a specific connectivity pattern, with the novel inclusion of multiplicative nodes. Each memory cell is equipped with an internal state and a number of multiplicative gates, namely, the input, forget, and output gates. The input gate determines whether a given input should impact the internal state; the forget gate determines the extent to which

the internal state should be flushed; and the output gate determines the extent to which the internal state of a given neuron should be allowed to influence the cell's output. The LSTM uses two activation functions: the tanh function and the sigmoid function. The repetitive module of the Elman RNN contains only one tanh function, while the repetitive module in the LSTM contains four interacting activation functions (three sigmoid and one tanh) (Fig. 3).

Generative adversarial network

GAN is a promising framework composed of two components: a generator and a discriminator. The generator generates false data samples and tries to deceive the discriminator. The discriminator tries to distinguish between true and false samples, which compete with each other in the training phase [25]. By repeating these steps, the generator and discriminator continue to improve in their respective tasks (Fig. 4). The generator and discriminator are usually implemented by a neural network with

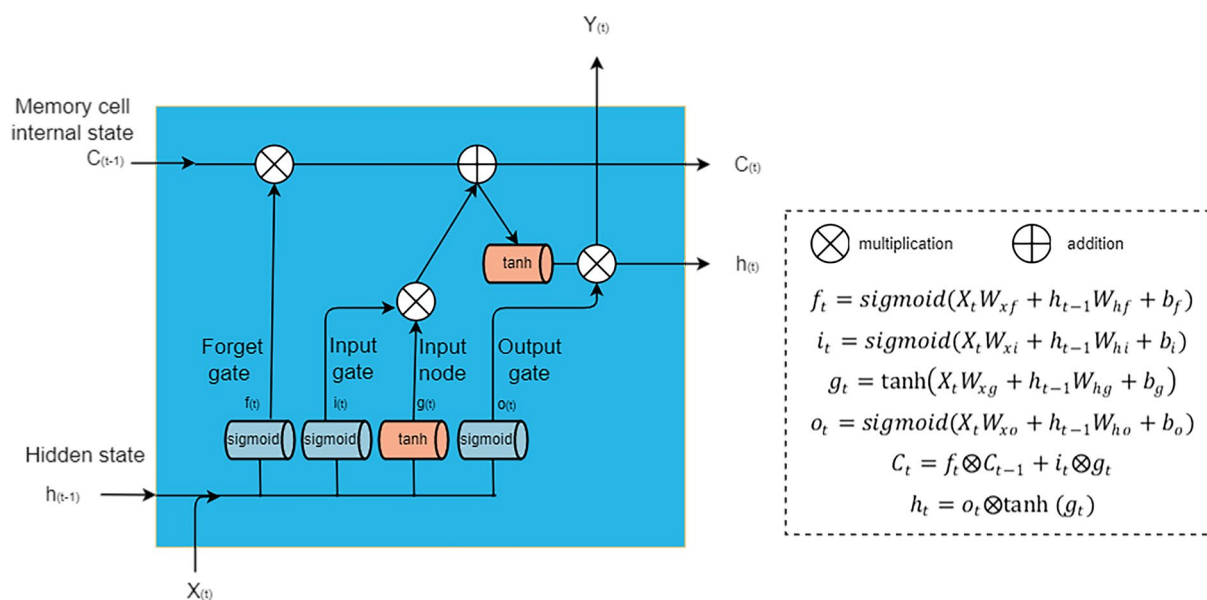


Fig. 3 The basic structure of LSTM. Input $X(t)$, output $Y(t)$; W represents weight, b represents bias term. Hidden state: $h_t = o_t \otimes \tanh(g_t)$; input node: $g_t = \tanh(X_t W_{xg} + h_{t-1} W_{hg} + b_g)$; memory cell internal state: $C_t = f_t \otimes C_{t-1} + i_t \otimes g_t$. Input gate: $i_t = \text{sigmoid}(X_t W_{xi} + h_{t-1} W_{hi} + b_i)$; forget gate: $f_t = \text{sigmoid}(X_t W_{xf} + h_{t-1} W_{hf} + b_f)$; output gate: $o_t = \text{sigmoid}(X_t W_{xo} + h_{t-1} W_{ho} + b_o)$

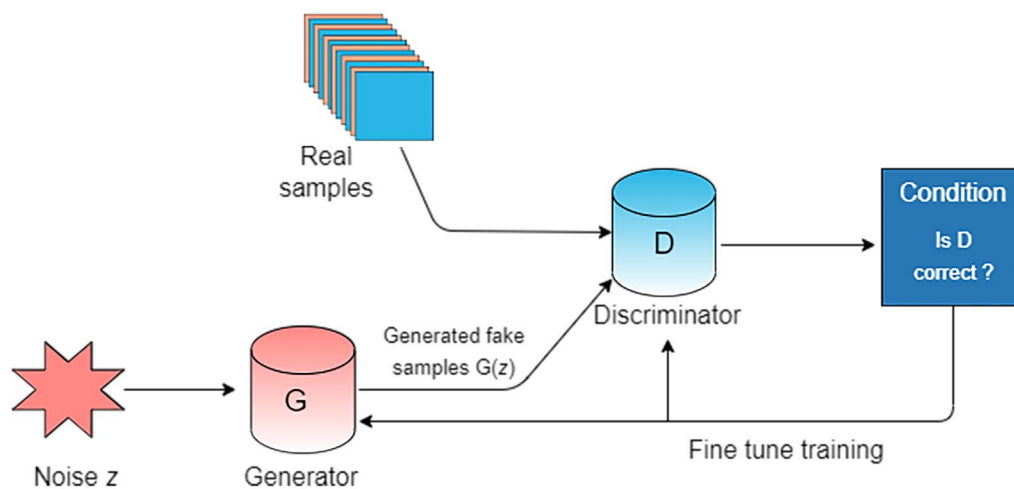


Fig. 4 The basic structure of GAN structure, consisting of a generator and a discriminator

simultaneous training, and both of them are trained by playing minimax games from game theory. The generator maximizes the cross-entropy loss (i.e., $\max \log(D(x'))$), while the discriminator minimizes the cross-entropy loss (i.e., $\min -y \log D(x) - (1 - y) \log(1 - D(x))$). The adversarial loss created by the discriminator provides a clever approach to incorporate unlabeled samples into training and impose higher-order consistency. This model has achieved state-of-the-art performance in many image-generation tasks, including text-to-image synthesis, super-resolution, and image-to-image translation [26].

Machine learning algorithms-traditional machine learning algorithms

Multilayer perceptron

Multilayer perceptron (MLP) uses a neural network with fully connected layers in a nonlinear model, including an input layer, several hidden layers, and an output layer with the capability to calculate the weighted sum of its inputs and then apply an activation function to transform a signal to the next neuron [27] (Fig. 5). The ReLU, sigmoid, and tanh functions are the common activation functions in the MLP. The sigmoid function, which takes a real-value input and “squashes” it in a range between 0 and 1, was often used previously. Like the sigmoid function, the tanh function also squashes its inputs, transforming them into elements on the interval between -1 and 1. However, the ReLU function has now emerged as a more popular nonlinear function with a mathematical formula that chooses the maximum of either x or 0 [16]. The ReLU function is significantly more amenable to optimization than the sigmoid or the tanh function.

MLP utilizes the back-propagation method with stochastic gradient descent to training [28].

Support vector machine, decision tree, random forest

The support vector machine (SVM) algorithm is a generalized linear classifier model for binary classification [29]. By using a non-parametric max-margin classification technique, it can classify data into two groups [30]. However, since practical research usually involves nonlinear problems, high-dimensional linear separable problems should be used instead of low-dimensional linear inseparable problems. Kernel functions such as the Gaussian, linear, and polynomial kernel functions are used to address this issue. SVM is based on the principle of structural risk minimization and shows excellent characteristics in nonlinear and small-sample problems. Combinations of methods or modification of separating hyperplanes, classification margins, boundaries, etc., can improve the generalization of SVM [30], avoiding the under-fitting and overfitting problems in previous attempts at neural network learning and yielding high generalization ability [31]. Decision tree (DT) is a basic classification and regression method in machine learning that mainly includes feature selection, decision tree generation, and pruning [32, 33]. Common DT algorithm models include the ID3, C4.5, and CART algorithms [34]. The DT algorithm is simple and intuitive, easy to understand, shows enough flexibility and expression ability. The random forest (RF) algorithm is an extension of DT with a high-performance speed [35, 36]. RF can perform prioritization of features by assigning different weight coefficients to different categories [37]. RF works by sequentially injecting training data and feature vectors

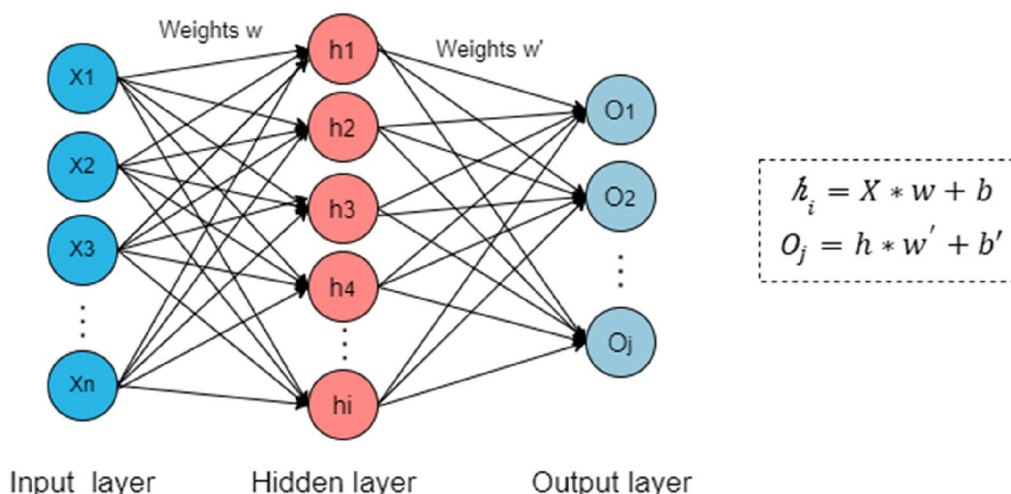


Fig. 5 The basic structure of MLP, including input layer, hidden layer, and output layer. The parameter w and w' represent weight, b and b' represent bias term. Hidden layer: $h_i = X * w + b$; output layer: $O_j = h * w' + b'$

into each of the base learners, identification of the best subset of features, and achieving the highest performance among all the aggregated base learners by increasing the impact factor of the best-feature subset in the classifier [38].

Comparisons between algorithms

From the original AlexNet in 2012 through the VGG in 2014 and ResNet in 2015, CNNs have been predominantly used in the field of computer vision [39] and natural language processing [40]. CNNs show good fitting effect and high accuracy. The weight-sharing feature of CNNs reduces the number of parameters, while their shift-invariant feature enhances the robustness of the network and shows an anti-disturbance effect. However, the shift-invariant feature also means that slight changes in the object will not activate the recognition of the object neurons. Moreover, the CNN model pooling layer loses a lot of valuable information, and the lack of a memory function as well as the limited data size and high computational requirements are other shortcomings of CNN [41] (Table 1). The Elman RNN has shown good ability in capturing the dynamics of sequences via recurrent connection, e.g., as in natural language processing [22]. The Elman RNN is effective and shows good generalization ability and has been widely used for solving practical problems [21]. However, it is prone to show gradient explosion and gradient vanishing, and cannot address the problems of long-term dependencies and parallel training [42]. In comparison with the Elman RNN, LSTM can achieve better analysis results in longer sequences, solving the vanishing gradient problem and stability problems in the time dimension of Elman RNN. However, the use of LSTM for processing longer sequence data is still difficult. Moreover, its calculation is time-consuming [22] (Table 1). As a generative model, GAN only uses back-propagation, which improves efficiency. GAN is an unsupervised learning method and is good at generalization. It can be used when the probability density cannot be calculated. The main disadvantage of GAN is the unstable training process and the difficulty in achieving Nash equilibrium. GAN can be used for various learning tasks, especially in the field of computer vision, but it is not suitable for processing discrete forms of data, such as text [25] (Table 1).

MLP is a simple and easy-to-implement algorithm with good generalization ability that is often used for identification, classification, and prediction [32]. However, there are two main problems associated with the development of MLP networks: architecture optimization and training. The definition of the architecture is a critical point because the lack of connections can reduce the ability of the network to solve the problem of insufficient

adjustable parameters, while too many connections may lead to overfitting of the training data. Therefore, training for large datasets is very time-consuming with MLP [27] (Table 1). SVM is more suitable for binary classification with small sample sizes and shows better robustness and generalization ability [43]. However, SVM is sensitive to parameters and kernel functions, and it is not suitable for multi-classification research in the case of non-optimization. SVM is often used in data classification and regression [32] (Table 1). As common traditional machine learning algorithms, DT and RF are based on simple principles and are easy to implement and can be used for data classification and regression [32, 44]. However, DTs are unstable since small variations in the data may result in the generation of a completely different tree. On the other hand, although RF shows good capability to reduce data noise, it is prone to overfitting when training large amounts of data [45]. RF has a simple structure, and is ease of understanding, performs higher efficiency than similar methods [37] (Table 1).

Deep learning offers absolute advantages in computer vision and natural language processing. Its powerful processing capabilities for features such as image and time series features are beyond the reach of traditional machine learning algorithms. Deep learning methods such as generative adversarial algorithms and reinforcement learning ensure continuous improvements in the calculation accuracy, which is also beyond the reach of traditional machine learning algorithms. However, for small datasets, deep learning is prone to overfitting and shows no advantage over traditional machine learning algorithms. Thus, the development of artificial intelligence techniques for TCM will require a combination of deep learning and traditional machine learning. Deep learning is preferred in feature extraction, such as semantic segmentation data fitting and image feature extraction. In contrast, traditional machine learning algorithms such as MLP and RF may be more suitable for small data classification and regression problems (Table 1).

Applications of machine learning in TCM research

Applications of machine learning in natural products development

Deep learning is widely used in the research and development of natural products (Table 2). Natural products which contain many effective chemical components with great potential value are the main methods to treating diseases in TCM. Approximately 70–95% of people in the developing world continue to rely on natural products as their primary pharmacopeia [47]. Thus, the development of natural products is of a great importance in clinical therapy, especially in combination with machine learning, is an innovative, forward-looking, and applicable

Table 1 Comparisons between machine learning algorithms

Machine learning algorithm	Advantages	Limitations	Applications
Deep learning			
CNN	High accuracy; Weight-sharing; Relieves the model overfitting problem; Shift-invariant feature enhances the robustness of the network	The pooling layer loses a lot of valuable information; Substantial hardware and dataset size requirements [41]; No memory function; Shift-invariant feature also prevents the neuron that recognizes the object from being activated when the object changes slightly	Computer vision [40, 41]; Natural language processing [40]
Elman RNN	Strong ability to extract time series features; Better generalization ability	Prone to show gradient exploding and gradient vanishing; Unable to solve the problem of long-term dependencies and parallel training [42]	Time series data, e.g., natural language processing [22]
LMST	Achieve better analysis results in longer sequences; Solving vanishing gradient problem and stability problems in the time dimension of Elman RNN	When processing longer sequences data, LSTM is still difficult; Time-consuming [22]	For processing longer time series data than Elman RNN, such as in natural language processing [22]
GAN	Generative model; Can still be used when the probability density is not calculated; Good at generalization	Unstable training process; Difficult to achieve Nash equilibrium; Not suitable for processing discrete forms of data	Data augmentation [46]; Text-to-image synthesis [26]; Image-to-image translation [26]; Computer vision [25]
Traditional machine learning			
MLP	Simple model; Easy implementation, and good generalization ability	architecture optimization; For training large datasets is very time-consuming [27]	Identification; Classification and prediction [32]
SVM	Suitable for small-sample binary classification research; Good robustness and generalization ability	Sensitive to parameters and kernel function; Inappropriate for multi-classification research in non-optimized cases	Classification and Regression problems [32]
DT	Simple to understand and to interpret; Requires little data preparation	Unstable for small variations in the data might result in the generation of a completely different tree	classification, and regression problems [32]
RF	Simple structure; Easy to implement; Higher efficiency [37]	Unable to optimize its own parameters; Overfitting can easily occur when the amount of data is large [45]	Classification and regression problems [44]

Table 2 Applications of machine learning in TCM research

Chinese medicine field	Correlation algorithm	Performance of the algorithm
Natural product development	Deep neural network [47, 50], RF [53, 54, 58, 59, 93], SVM [51, 53, 54, 57, 59, 93], DT [59, 93], neural network [53, 59]	RF was better than SVM, neural network and DT in screening hepatotoxic compounds [59]. RF model is more accurate than SVM and DT in identifying molecular characteristics of natural product compounds with the meridians of TCM [93]
Disease diagnosis	SVM [10, 61, 66, 81–83], DT [68, 81–83], neural network [45, 61–63, 65, 82, 83], RF [61, 64, 67, 82, 83], CNN [64, 67, 70–78, 81], RNN [80], LSTM [64], MLP [68], residual CNN [81]	Compared with SVM, DT, traditional neural network and RF, CNN achieves higher accuracy in tongue shape classification [71]; Residual CNN model with highest accuracy and sensitivity in sounds auscultation [81]; SVM usually achieves better performance than other traditional algorithms in pulse diagnosis [82, 83]; CNN perform better than traditional methods in pulse diagnosis, with accuracy above 90% [71]
Disease treatment and effect evaluation	SVM [84, 87, 88], RF [84], GAN [85], CNN [86, 88], DT [88]	CNN performs better compared with SVM and DT [88]
Prediction of biomarkers in TCM	RF [89], SVM [90, 92], Residual CNN [48], DT [92], neural network [92]	SVM model is better (accuracy:98.45%) than DT, neural network in predicting tumor necrosis factor-alpha converting enzyme inhibitors [92]

new model. The effective scientific characterization of natural products is the basis for using machine learning [48]. Chemical descriptors and fingerprints are often used to quantify the natural products' effective chemical entities physicochemical characteristics and the related biological target molecules. Chemical descriptors characterized molecules properties by experimental quantification or theoretics which represent its chemical, physical, or topological characteristics. While chemical fingerprints are more complex for encoding as binary bit strings. Chemical fingerprints can reflect the active constituents in substances and can effectively characterize the quality of TCM's materials [48]. Both molecular descriptors and fingerprints perform crucial functions in machine learning-based applications for drug discovery processes such as target molecule ranking, similarity-based compound search, and virtual screening [49]. For example, machine learning could successfully identify the antibiotic precursor halicin with different structures [50]. The researchers used a deep neural network model that translated the graphical representation of a molecule into a continuous vector through a directed bond-based message passing approach to train the dataset and then used computer simulations to screen compounds that were obtained by vitro screening to finally obtain halicin. Chen [51] used the SVM algorithm to establish a mathematical discriminant model to distinguish the cold and hot nature of natural products. Machine learning can collect and process data based on the medicinal properties, chemical compositions, and function of natural products, allowing automatic discrimination and prediction. Chuang [52] comprehensively discussed how artificial intelligence can address the limitations of molecular descriptors and fingerprints and thereby improve the predictive modeling of compound bioactivities. Yang [53] utilized RF, neural networks, and SVM to identify new compounds in TCM prescriptions for Alzheimer's disease. They utilized data mining to collect Alzheimer's disease-related and unrelated compounds from the literature databases. Then, RF, gradient boosting machine and neural networks were utilized to determine the importance of each feature, and important features were selected by molecular descriptors for feature extraction. The selected features were input to the SVM algorithm to identify the new compounds in TCM prescriptions. Yu [54] used RF to obtain the feature descriptors of natural product compounds, SVM to predict hit molecules based on the feature descriptors screened by RF, and molecular docking to perform virtual screening. They successfully identified 4',5,7-trimethoxyflavone as a potential platelet-derived growth factor receptor α (PDGFRA) inhibitor.

Although the majority of natural products appear inherently safe, clinicians and researchers should also

pay attention to the potential for drug-induced injury. The liver which is the major organ of drug metabolism is more likely to show drug-induced injuries than other organs, and these injuries may lead to hepatitis, liver fibrosis, liver failure, and even death [55]. The kidney is also highly susceptible to drug-induced toxic insults that are a common cause of acute kidney injury [56]. With advancements in machine learning, researchers have turned their attention to the use of machine learning applications for evaluating drug-induced injuries. Hu [57] used SVM and in vitro screening to predict and validate the risk of idiosyncratic drug-induced liver injuries caused by the natural products in *Polygonum multiflorum* Thunb, and provided a powerful tool to screen large datasets for toxicants. He [58] established a large-scale dataset focused on TCM-induced hepatoprotection to train machine learning models such as RF and voting models. Their work helped screen potential hepatoprotectants from natural products. Chen [59] developed a method for screening hepatotoxic compounds in TCM and Western medicine combinations on the basis of chemical structures by using SVM, neural networks, DT, and RF. Their results showed that RF yielded a classification accuracy of 0.838, which was better than other machine learning methods.

Applications of machine learning in disease diagnosis

With the application of AI technology in TCM, AI-assisted disease diagnosis has emerged as a promising research field. With TCM symptoms corresponding to features in the machine learning literature, syndrome elements serve as classes or labels [60], and machine learning has been used in disease diagnosis models (Table 2). Wang [10] used an optimized SVM algorithm to construct a serology-based lung cancer diagnosis model, analyzed the potential therapeutic mechanisms of wogonin in lung cancer, explored the relationship between serological markers and wogonin targets, and constructed a signal pathway regulated by wogonin. Shi [61] developed a new fatigue classification method by integrating pulse data and tongue images with machine learning algorithms and using machine learning models, including SVM, RF, and neural networks, to diagnose disease-related fatigue and non-disease-related fatigue. Senoner [62] achieved good results when using the neural network algorithm with electrocardiogram data to assist the diagnosis of preexcitation syndrome. Using the neural network model based on blood pressure data, Sun [63] established a TCM syndrome diagnosis model of coronary heart disease. Zhang [64] developed a TCM assistive diagnostic system by utilizing bidirectional LSTM with RF for named entity recognition, a CNN for text

processing for disease diagnosis, and an integrated learning model for syndrome prediction. Zhao [65] utilized an adaptive resonant neural network for quantitative diagnosis of TCM syndrome types.

Clinical information for TCM diagnosis is collected by the diagnostic methods of looking, listening and smelling, asking, and touching. Intelligent auxiliary diagnosis methods based on these four diagnostic methods in TCM are constantly developing with the accumulation of clinical diagnosis and treatment records, experimental records, TCM databases, books, medical literature, and the other knowledge. Diagnosis based on visual examination is an important method to obtain disease information, and tongue and eye diagnoses are its main components. TCM tongue diagnosis involves interpretation of tongue images obtained by doctors on the basis of the theory of TCM after observing the tongue coating, quality, shape, and other tongue-related information. Tongue diagnosis provides much information about the state of the body, and the diagnostic points of tongue diagnosis include the color and state of the tongue coating, color, texture, shape, and characteristics of the sublingual vein and tongue body parts, etc., which are important features in tongue diagnosis data collection. Traditional machine learning algorithms such as SVM [66], DT, neural networks [45], and RF [67] have been previously used as intelligent auxiliary diagnosis algorithms for tongue diagnosis. The accuracy of SVM in processing hyperspectral red-green-blue tongue images based on tissue type combination can be as high as 93.11% [66]. Liu [68] selected 22 kinds of tongue features in 311 participants to establish the training data set, and used DT (accuracy rate, 66.9%) and MLP (accuracy rate, 64.3%) to classify the tongue images corresponding to kidney deficiency. Qi [67] used the open-source Weka software to classify the color of 728 tongue images, and obtained an RF prediction accuracy of 84.94%. Yan [69] used deep learning and RF to classify normal, mild, and severe teeth-marked tongues. Lu [70] utilized Ridge-CNN to classify sublingual varices of TCM with an accuracy rate of 87.5%.

Although SVM, RF, and MLP have been shown to be effective for simple image classification, they are not satisfactory for complex tasks. With advancements in machine learning, deep learning is being gradually applied for complex task processing for intelligent tongue diagnosis. CNN models are good at image classification, and have been shown to be better than other traditional algorithms in this field [71]. The AlexNet, GoogLeNet, ResNet, and DenseNet network structures with CNN as the model algorithm have been applied for tongue image classification. Huo [72] used a CNN model with an AlexNet network structure and achieved higher accuracy

for tongue shape classification as well as reduced training time for the CNN model. Xiao [73] used the improved AlexNet network structure to build a tongue coating color classification model. Using the GoogLeNet network, Christian [74] proposed the inception module to optimize training from another perspective, extracting more features with the same amount of computation. With advancements in CNN, ResNet solved the problem of difficult training, high error rates, and a rapid decline in accuracy after the CNN depth increases. Shao [75] first separated the tongue and tongue coating, and then used the separated images as input to classify the tongue and tongue coating using ResNet-50. Residual connections make the CNN deeper, stronger, and more efficient. DenseNet further expands network connectivity to ensure maximum information flow between layers. Using the AlexNet network structure, Chen [76] introduced the dense connection method in DenseNet and proposed the tongue-coating classification model TonNet.

Eye evaluations can realize intelligent auxiliary diagnosis through fundus image analysis. Retinal vessels are the only visible blood vessels that can be evaluated by simple fundus photography, and this approach provides a convenient method to evaluate cardiovascular status. One study found that retinal features were associated with stroke, and the researchers used the CNN model of the ResNet50 network structure to conduct a stroke risk assessment using retinal images [77]. Sun [78] used CNN to extract features and identify syndromes of yin deficiency, and achieved good results.

Traditional auscultation mainly involves listening to sounds. With advancements in medical treatment, auscultation using equipment is now also utilized in TCM. The combination of an electronic stethoscope with artificial intelligence technology can allow digital acquisition of heart sounds, providing an objective basis for heart sound auscultation. Traditional machine learning methods for heart sound auscultation usually involve segmentation, feature extraction, and classification. Although traditional machine learning methods allows rapid model training, they usually require complex preprocessing and post-processing steps. However, advancements in deep learning, especially in the CNN model, have yielded favorable results for intelligent diagnosis based on heart sounds. The intelligent heart sound auscultation process includes signal acquisition, signal preprocessing, heart sound feature extraction, and model training [79]. Fernando [80] proposed a heart sound segmentation method based on the combination of RNN with attention mechanism, which can effectively learn features from irregular and noisy heart sounds. Liu [81] used gradient-enhanced DT, SVM, CNN, and residual convolutional recurrent networks to analyze heart sound signals. The

results showed that the residual convolutional recurrent network model has the highest recognition accuracy and sensitivity for the four types of coronary heart disease heart sounds.

Pulse diagnosis is one of the most important diagnostic methods in TCM. Doctors use three fingers to touch the wrist at three specific positions, namely, inch, off, and ruler, to examine the pulse and determine the health of patients. With advancements in sensors, detectors, and sensor technologies, digital palpation data can now be obtained from the same location, enabling AI technology to process palpation data and make diagnoses [82]. At present, most of the artificial intelligence technologies used in pulse diagnosis are limited to classical machine learning algorithms and their improved versions, including SVM, RF, DT, and neural network. Each learning algorithm shows unique advantages in pulse diagnosis learning classification. SVM usually achieves better performance than other traditional algorithms [82, 83]. As a widely used and adaptable deep learning method, CNN model algorithms have been proposed for TCM pulse diagnosis. CNN is good at mining local features and classifying and extracting global features. Moreover, CNN has been shown to perform better than traditional methods in AI-assisted pulse diagnosis, with accuracy above 90% [71].

Applications of machine learning in disease treatment and effect evaluation

Machine learning has recently been successfully applied in disease treatment and effect evaluation of TCM, such as in prescription recommendation, transition prediction, and treatment prognosis (Table 2). The treatment prognosis model has received increasing attention in the context of clinical diagnosis and treatment decision-making [84]. Zhang [85] utilized transformer and GAN to develop an auxiliary tool to prescribe TCM prescriptions based on the patient's clinical electronic health records. In their approach, transformer was used for TCM prescription generation, while the GAN model aims to augment the training set to further enhance the overall system performance by reducing overfitting effect. Dong [86] proposed a TCM prescription recommendation based on subnetwork term mapping and deep learning. They used TCM clinical case data to construct a natural product-symptom-related knowledge graph, constructed a symptom network by combining a meta path method and knowledge graph, proposed a subnetwork-based symptom term mapping method, utilized CNN as the train model, and finally output the prediction probability of each natural product to obtain the recommended prescription [86]. Dengzhan Shengmai capsule is a patented TCM preparation for the secondary prevention of stroke.

Lu [87] utilized SVM to classify the network matrix of the Dengzhan Shengmai capsule group at baseline versus after treatment. SVM classification revealed significant white matter network alterations after treatment in the drug groups, with an accuracy of 68.18%. Tang [84] used RF, SVM, logistic regression, and extreme gradient boosting to predict whether colorectal cancer recurrence and metastasis with TCM factors would occur within 3 years and 5 years after radical surgery. The results showed that the four methods all showed certain predictive ability (area under the curve values > 0.70). Liu [88] proposed a graph CNN model to predict formula efficacy. The performance of graph CNN for multi-classification of tonic formulae showed the best result in comparison with SVM, naive Bayes, logistic regression, DT, and K-nearest neighbor.

Applications of machine learning in prediction of biomarkers in TCM

The continuous advancement of information technology and biotechnology has yielded substantial biomarker data for TCM investigations using machine learning. Zhang [89] used RF and least absolute shrinkage and selection operator (LASSO) regression to identify important characteristic genes of oxidative stress. The receiver operating characteristic results demonstrated that the model was better in prediction efficiency with an AUC of 0.873. They also found that Nobiletin, which targets PLA2G4, may indicate a third pathway for the treatment of acute myeloid leukemia. Zhang [90] utilized SVM and LASSO to screen the underlying feature biomarkers in four RNA microarray datasets of myocardial infarction. These two machine learning methods yielded 10 and 14 genes, respectively. IL1B and TLR2 were the intersection biomarkers obtained by SVM and LASSO. On the basis of these biomarkers, several natural products such as *dan shen* and *san qi*, were identified as the potential TCM preparations for the treatment of myocardial infarction. By utilizing machine learning (residual CNN and partial least squares discriminant analysis), fingerprint, and network pharmacology, Li [48] screened the potential biomarkers in different parts of *Wolfiporia cocos*. Yuan [91] used RF to construct a drug-target prediction model to predict the key targets of *Corydalis Rhizoma* in the treatment of cardiovascular and cerebrovascular diseases. Cong [92] utilized SVM, DT, and back-propagation neural network to predict novel and selective tumor necrosis factor-alpha converting enzyme inhibitors. In their work, the SVM model showed the best overall prediction accuracy (98.45%) (Table 2).

Research foundation and future research direction
Advantages of algorithm ensemble and establishment
of syndrome element diagnosis model of coronary heart
disease

Our team proposed that algorithm ensemble is more suitable for TCM data models. Although an algorithm model can realize the construction of a prediction model based on the data for certain characteristics, the lack of universality and transferability limits the applicability of such models in systematic research on the diagnosis and treatment process of symptom-syndrome-treatment-prescription-natural products. Therefore, considering the current status of artificial intelligence in TCM, unification of different diagnosis and treatment rules on the basis of syndrome elements through a combination of multiple algorithms may facilitate accurate calculation of the entire diagnosis and treatment process in symptom-syndrome-treatment-prescription-natural products. The general structure of an ensemble learning algorithm consists of a set of “individual learners”, which are used to create an ensemble using a certain policy [94]. Using the ensemble principle, the advantages of each model can be extracted to integrate the optimized fusion model. Voting and stacking are common strategies for algorithm ensembles. A voting ensemble is a simple and effective fusion method that can weight the prediction results of a single model to improve model diversity while ensuring performance. The stacking ensemble is a more powerful learning-based ensemble strategy that uses the initial dataset to train a “component learner” and then generates a new data set for training a “meta-learner”. The training results show that the prediction performance of the integrated model on the training set and the test set is generally better than that of the single model. Fu [95] combined CNN with traditional algorithm models to analyze tongue-coating properties and found that the performance of the integrated model was improved. On the basis of the

voting and stacking strategy, Yang [96] performed rule integration-model fusion on three machine learning models of SVM, RF, and neural network, and achieved good performance. Ge [97] proposed an ensemble algorithm that integrated the attention mechanism and LSTM, and showed that this ensemble algorithm can effectively select salient locations with higher accuracy and less computation.

The diagnosis and treatment theory in TCM shows the characteristics of diversified intersection. TCM treatment based on syndrome differentiation involves eight principal forms of differentiation, qi and blood fluid syndrome differentiation, zang-fu differentiation, and meridian syndrome differentiation. On the other hand, the compatibility rules include monarch and minister compatibility, flavor compatibility, and component compatibility. In addition, differences in the diagnosis and treatment rules among TCM sects and TCM physicians have made it difficult to reserve large high-quality TCM data of the same rule system. To address this problem, our team proposed syndrome elements that could link the symptoms, treatment, natural products, and prescriptions, and thereby unify different diagnosis and treatment rules [98]. The team used the improved transformer algorithm to construct a diagnostic model of coronary heart disease syndrome elements. This transformer model integrated the Seq2Seq module of RNN, LSTM and self-attention mechanism. At the same time, the multi-head attention mechanism, compound word vector, and random inactivation are used to study the syndrome elements of coronary heart disease (Fig. 6).

Proposal of an application that combining machine
learning with TCM theory and natural product
computational biology

The effective chemical components of natural products have attracted attention worldwide. Rapid screening and

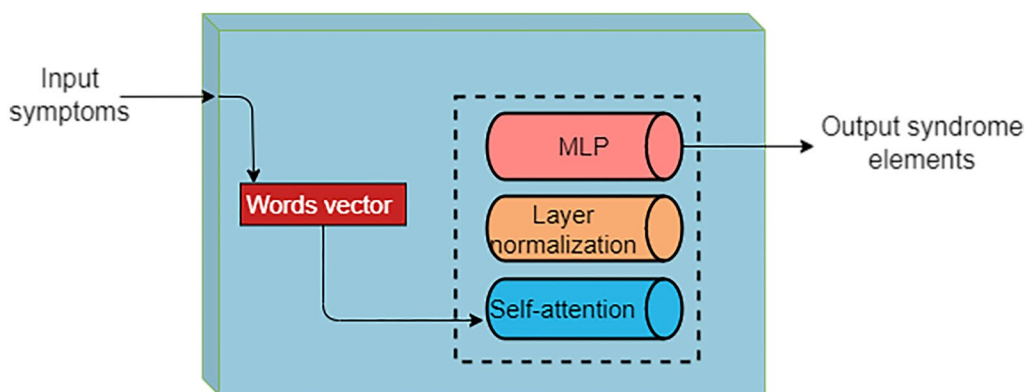


Fig. 6 The diagnostic model of coronary heart disease syndrome elements by our team

identification of potential candidate compounds are very vital to determine the mechanisms underlying the therapeutic effects of drugs and can greatly ameliorate the development of new drugs. Since the successful development of artemisinin, the expectations for discovery of novel drugs with high efficacy and minimal adverse effects from TCM have increased. In this regard, a combination of the knowledge of effective chemical components of natural products with computational biology and machine learning under the optimization of TCM syndrome differentiation theory can facilitate disease treatment to realize effective development of machine learning in TCM clinical practice. The team proposed machine learning using a combination of computational biology of natural products and TCM theory. In this approach, computational biology is first used to study the pharmacology of natural products. Using the key molecular targets of disease as the research aspect, computational biology is used to simulate and screen the effective components of natural products for the key molecular target of disease. Based on molecular mechanisms identified in natural products screening, machine learning is performed on the selected natural products to establish a prediction model of molecular characteristics of natural products compounds and attributes, TCM syndrome differentiation, and meridians. For disease clinical data collection, the establishment of patient syndrome model in combination with a TCM theory screening model can allow optimization of the final therapeutic drugs.

The study used computational biology methods to analyze and screen natural products. Computational biology can unravel the seemingly impenetrable complexity of biological systems by an integrated approach which employed high-performance computers, state-of-the-art software and algorithms, mathematical modeling, and statistical analyses [99]. Molecular dynamics, which is a branch of computational biology, can simulate the molecular mechanisms of effective chemical molecules of natural products acting on the body. Using this approach, the effective chemical molecules of natural products can be screened out to future identify natural products which are effective to remedy disease. Fang [100] summarized various cheminformatics, bioinformatics, and systems biology resources used to reconstruct drug-target networks for natural product medicine. Fu [101] developed a data-clustering method using a collection of 2,012 compounds associated with natural products and found that the cold and hot properties of natural products can be related to the physicochemical and target pathways of their constituent compounds. Wang [93] used DT, SVM, and RF algorithms for the first time to link the molecular characteristics of natural products compounds with the meridians of TCM. They identified the molecular

characteristics of 646 natural products and their active constituents, including structure-based fingerprints and absorption, distribution, metabolism, and excretion characteristics. The meridian properties of TCM were predicted by machine learning methods, with the highest accuracy of 0.83, and RF showed the best accuracy.

Syndrome differentiation and treatment form the core of TCM theory. The development of intelligent diagnosis based on machine learning provides a dialectical basis for TCM syndrome differentiation. The use of machine learning in the research and development of TCM provides a basis for syndrome differentiation, while the unified diagnosis and treatment rules based on syndrome elements provide a direction for syndrome differentiation. Machine learning with molecular basis underlying syndrome elements may better classify diseases and improve clinical treatment effectiveness. Now, a web platform named SoFDA (<http://www.tcmip.cn/Syndrome/front/#/>) which is a network-based evaluation tool of multi-way associations among diseases, syndrome differentiation, and prescriptions could facilitate the understanding of syndrome differentiation and natural products from the perspective of molecular biology, enriched gene ontology terms or signaling pathways associated with syndrome differentiation [102]. Syndrome differentiation and treatment theory also conduct the relationship between natural products and syndrome elements which based on the function of natural products and their matched syndrome elements. By contacting natural products, syndrome elements, and molecules, promising research that combining machine learning with TCM theory and natural product computational biology were proposed. The stem diagram of these proposal is shown in Fig. 7. After using machine learning algorithms to intelligently diagnose disease and syndrome differentiation, the natural products screened by computational simulation can be used to realize intelligent diagnosis and treatment according to the results of intelligent differentiation of TCM. This process will use computational biology, machine learning, and TCM theory to achieve intelligent diagnosis and treatment of TCM, and is a potential research direction for TCM machine learning.

Conclusion

In summary, this study reviewed the applications of machine learning in TCM research, including the principles of deep learning and traditional machine learning algorithms, the application of machine learning algorithms in TCM research, and an analysis of promising research directions. Machine learning has been applied in natural product research, TCM disease diagnosis, disease treatment and effect evaluation, and prediction of

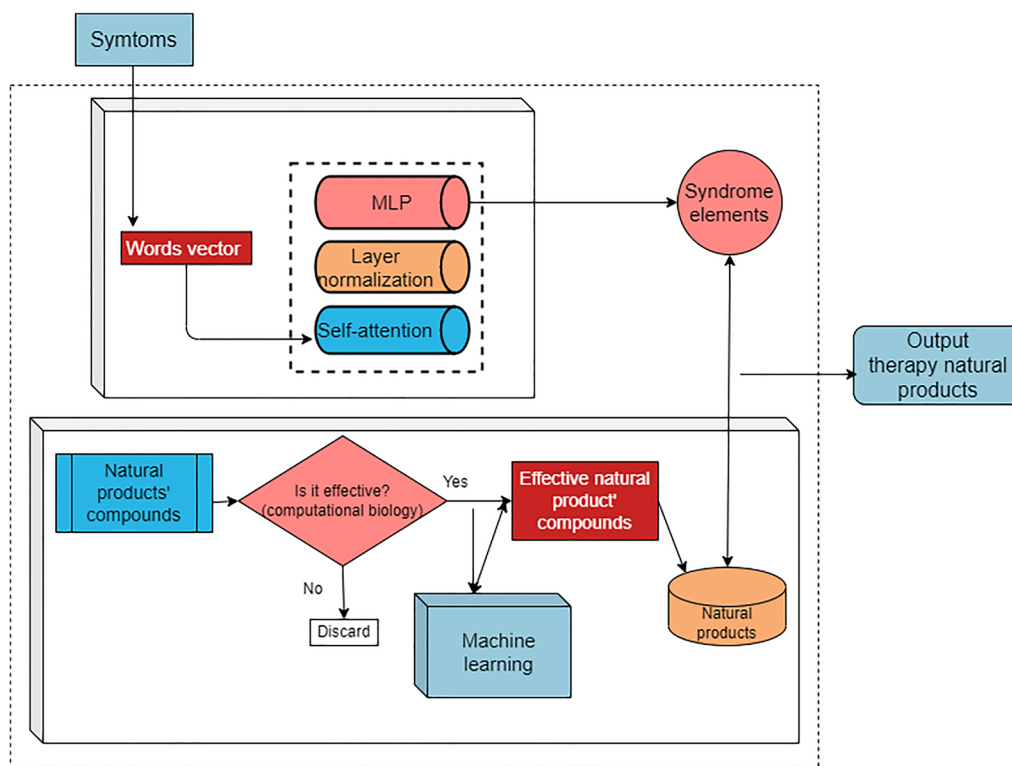


Fig. 7 The perspective study of integrated methods including machine learning, TCM theory, natural product ingredients, and computational biology

biomarkers in TCM. Traditional machine learning algorithms such as SVM, RF, DT and MLP are widely used in TCM learning. With advancements in machine learning, deep learning has found additional applications in TCM. Deep learning shows higher prediction performance than traditional machine learning algorithms. Although the clinical diagnosis and treatment process in TCM produces large amounts of data, the diversity in diagnosis and treatment models has resulted in a lack of uniform standards. Thus, the use of syndrome elements as a unified standard is important for addressing the difficulties in developing artificial intelligence-based techniques for TCM. The multi-algorithm rule integration proposed herein is more suitable for a TCM data model. Natural products contain many chemical components that influence the therapeutic effects. By utilizing computational simulation, these medicines can be screened at the molecule level. TCM theory serves as the guideline for the use of natural products, and the integration of machine learning with TCM theory, natural product, and computational simulation can yield an intelligent artificial intelligence-driven diagnosis and treatment model based on the effective chemical components of natural products under the guidance of TCM theory. Thus, the combination of machine learning with our understanding of

effective chemical components of TCM and TCM theory offers a practical direction for the use of artificial intelligence in TCM, which can be expected to have far-reaching implications.

Although the development of machine learning in TCM is a promising study, the challenges and difficulties cannot be ignored. Effectiveness and safety are issues that need to be paid attention to in the development of artificial intelligence in TCM. The accurate application of machine learning in TCM theory which is the guiding program of TCM is related to the clinical effectiveness of TCM artificial intelligence research, and is a difficulty and challenge that TCM artificial intelligence needs to solve. The clinical effectiveness of intelligent diagnosis and treatment of the chemical molecular mechanism of natural products under the guidance of correct TCM theory is a worth working. Safety in TCM intelligent diagnosis and treatment is another key point that influence its development. The potential for drug-induced injury should be taken into account in TCM artificial intelligence research.

Acknowledgements

Not applicable.

Author contributions

SYM, JLL and WHL conceived and designed the study. YML and XSH critiqued the study. PRQ and ZLJ performed the arrangement of pictures and tables. SYM wrote the manuscript. JL and JW supervised the study. All authors read and approved the final manuscript.

Funding

This study was supported by National Natural Science Foundation of China (Key Program) (NO. 82230124); Traditional Chinese Medicine Inheritance and innovation "Ten million" talent project - Qihuang Project Chief Scientist Project (NO. 0201000401); State Administration of Traditional Chinese Medicine 2nd National Traditional Chinese Medicine Inheritance Studio Construction Project (Official Letter of the State Office of Traditional Chinese Medicine (2022) No. 245); National Natural Science Foundation of China (General Program) (No.81974556).

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors agree to the publishing of this article in Chinese Medicine.

Competing interests

The authors declare no competing interests.

Received: 30 January 2023 Accepted: 28 March 2023

Published online: 20 April 2023

References

- Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920–30. <https://doi.org/10.1161/CIRCULATIONAHA.115.001593>.
- Mao S, Sejdic E. A review of recurrent neural network-based methods in computational physiology. *IEEE Trans Neural Netw Learn Syst*. 2022. <https://doi.org/10.1109/TNNLS.2022.3145365>.
- Seetharam K, Kagiya N, Sengupta PP. Application of mobile health, telemedicine and artificial intelligence to echocardiography. *Echo Res Pract*. 2019;6(2):R41–52. <https://doi.org/10.1530/ERP-18-0081>.
- Calderaro J, Seraphin TP, Luedde T, Simon TG. Artificial intelligence for the prevention and clinical management of hepatocellular carcinoma. *J Hepatol*. 2022;76(6):1348–61. <https://doi.org/10.1016/j.jhep.2022.01.014>.
- Razzaq M, Clement F, Yvinec R. An overview of deep learning applications in precocious puberty and thyroid dysfunction. *Front Endocrinol (Lausanne)*. 2022;13:959546. <https://doi.org/10.3389/fendo.2022.959546>.
- Zhang DY, Cheng YB, Guo QH, Shan XL, Wei FF, Lu F, et al. Treatment of masked hypertension with a chinese herbal formula: a randomized, placebo-controlled trial. *Circulation*. 2020;142(19):1821–30. <https://doi.org/10.1161/CIRCULATIONAHA.120.046685>.
- Tang JL, Liu BY, Ma KW. Traditional chinese medicine. *Lancet*. 2008;372(9654):1938–40. [https://doi.org/10.1016/S0140-6736\(08\)61354-9](https://doi.org/10.1016/S0140-6736(08)61354-9).
- Wu C, Chen J, Lai-Han LE, Chang H, Wang X. Editorial: artificial intelligence in traditional medicine. *Front Pharmacol*. 2022;13:933133. <https://doi.org/10.3389/fphar.2022.933133>.
- Chu H, Moon S, Park J, Bak S, Ko Y, Youn BY. The use of artificial intelligence in complementary and alternative medicine: a systematic scoping review. *Front Pharmacol*. 2022;13:826044. <https://doi.org/10.3389/fphar.2022.826044>.
- Wang S, Hou Y, Li X, Meng X, Zhang Y, Wang X. Practical implementation of artificial intelligence-based deep learning and cloud computing on the application of traditional medicine and western medicine in the diagnosis and treatment of rheumatoid arthritis. *Front Pharmacol*. 2021;12:765435. <https://doi.org/10.3389/fphar.2021.765435>.
- Guo Y, Chen J, Du Q, Van Den Hengel A, Shi Q, Tan M. Multi-way back-propagation for training compact deep neural networks. *Neural Netw*. 2020;126:250–61. <https://doi.org/10.1016/j.neunet.2020.03.001>.
- Ozawa S, Toh SL, Abe S, Pang S, Kasabov N. Incremental learning of feature space and classifier for face recognition. *Neural Netw*. 2005;18(5–6):575–84. <https://doi.org/10.1016/j.neunet.2005.06.016>.
- Ibtehaz N, Rahman MS. Multiresunet: rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural Netw*. 2020;121:74–87. <https://doi.org/10.1016/j.neunet.2019.08.025>.
- Schrauwen B, D'Haene M, Verstraeten D, Campenhout JV. Compact hardware liquid state machines on FPGA for real-time speech recognition. *Neural Netw*. 2008;21(2–3):511–23. <https://doi.org/10.1016/j.neunet.2007.12.009>.
- Gross A, Murthy D. Modeling virtual organizations with latent dirichlet allocation: a case for natural language processing. *Neural Netw*. 2014;58:38–49. <https://doi.org/10.1016/j.neunet.2014.05.008>.
- Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. Convolutional neural networks for radiologic images: a radiologist's guide. *Radiology*. 2019;290(3):590–606. <https://doi.org/10.1148/radiol.2018180547>.
- Wang S, Yang DM, Rong R, Zhan X, Xiao G. Pathology image analysis using segmentation deep learning algorithms. *Am J Pathol*. 2019;189(9):1686–98. <https://doi.org/10.1016/j.ajpath.2019.05.007>.
- Chatzikonstantinou C, Konstantinidis D, Dimitropoulos K, Daras P. Recurrent neural network pruning using dynamical systems and iterative fine-tuning. *Neural Netw*. 2021;143:475–88. <https://doi.org/10.1016/j.neunet.2021.07.001>.
- Wang J, Wang J, Fang W, Niu H. Financial time series prediction using elman recurrent random neural networks. *Comput Intell Neurosci*. 2016;2016:4742515. <https://doi.org/10.1155/2016/4742515>.
- Gunturkun R. Using elman recurrent neural networks with conjugate gradient algorithm in determining the anesthetic the amount of anesthetic medicine to be applied. *J Med Syst*. 2010;34(4):479–84. <https://doi.org/10.1007/s10916-009-9260-2>.
- Tang Q, Wu B. Multilayer game collaborative optimization based on elman neural network system diagnosis in shared manufacturing mode. *Comput Intell Neurosci*. 2022;2022:6135970. <https://doi.org/10.1155/2022/6135970>.
- Le VT, Tran-Trung K, Hoang VT. A comprehensive review of recent deep learning techniques for human activity recognition. *Comput Intell Neurosci*. 2022;2022:8323962. <https://doi.org/10.1155/2022/8323962>.
- Tariverdi A, Venkiteswaran VK, Richter M, Elle OJ, Torresen J, Mathiasen K, et al. A recurrent neural-network-based real-time dynamic model for soft continuum manipulators. *Front Robot AI*. 2021;8:631303. <https://doi.org/10.3389/frobt.2021.631303>.
- Landi F, Baraldi L, Cornia M, Cucchiara R. Working memory connections for LSTM. *Neural Netw*. 2021. <https://doi.org/10.1016/j.neunet.2021.08.030>.
- Hong H, Li X, Wang M. Gane: a generative adversarial network embedding. *IEEE Trans Neural Netw Learn Syst*. 2020;31(7):2325–35. <https://doi.org/10.1109/TNNLS.2019.2921841>.
- Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med Image Anal*. 2019;58:101552. <https://doi.org/10.1016/j.media.2019.101552>.
- Castro W, Oblitas J, Santa-Cruz R, Avila-George H. Multilayer perceptron architecture optimization using parallel computing techniques. *PLoS ONE*. 2017;12(12):e189369. <https://doi.org/10.1371/journal.pone.0189369>.
- Huang Y, Lu Y, Taubmann O, Lauritsch G, Maier A. Traditional machine learning for limited angle tomography. *Int J Comput Assist Radiol Surg*. 2019;14(1):11–9. <https://doi.org/10.1007/s11548-018-1851-2>.
- Noble WS. What is a support vector machine? *Nat Biotechnol*. 2006;24(12):1565–7. <https://doi.org/10.1038/nbt1206-1565>.
- Nedaie A, Najafi AA. Support vector machine with dirichlet feature mapping. *Neural Netw*. 2018;98:87–101. <https://doi.org/10.1016/j.neunet.2017.11.006>.
- Heikamp K, Bajorath J. Support vector machines for drug discovery. *Expert Opin Drug Discov*. 2014;9(1):93–104. <https://doi.org/10.1517/17460441.2014.866943>.

32. Cheng X, Manandhar I, Aryal S, Joe B. Application of artificial intelligence in cardiovascular medicine. *Compr Physiol*. 2021;11(4):2455–66. <https://doi.org/10.1002/cphy.c200034>.
33. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. *Radiographics*. 2017;37(2):505–15. <https://doi.org/10.1148/rg.2017160130>.
34. Zhao C, Li GZ, Wang C, Niu J. Advances in patient classification for traditional Chinese medicine: a machine learning perspective. *Evid Based Complement Alternat Med*. 2015;2015:376716. <https://doi.org/10.1155/2015/376716>.
35. Jones FC, Plewes R, Murison L, MacDougall MJ, Sinclair S, Davies C, et al. Random forests as cumulative effects models: a case study of lakes and rivers in muskoka, canada. *J Environ Manage*. 2017;201:407–24. <https://doi.org/10.1016/j.jenvman.2017.06.011>.
36. Galicia A, Talavera-Llames R, Troncoso A, Koprinska I, Martínez-Álvarez F. Multi-step forecasting for big data time series based on ensemble learning. *Knowl-Based Syst*. 2019;163:830–41. <https://doi.org/10.1016/j.knsys.2017.06.011>.
37. Savargiv M, Masoumi B, Keyvanpour MR. A new random forest algorithm based on learning automata. *Comput Intell Neurosci*. 2021;2021:5572781. <https://doi.org/10.1155/2021/5572781>.
38. Zhang Y, Miao D, Wang J, Zhang Z. A cost-sensitive three-way combination technique for ensemble learning in sentiment classification. *Int J Approx Reason*. 2019;105:85–97. <https://doi.org/10.1016/j.ijar.2018.10.019>.
39. Yin Y, He C, Xu B, Li Z. Coronary plaque characterization from optical coherence tomography imaging with a two-pathway cascade convolutional neural network architecture. *Front Cardiovasc Med*. 2021;8:670502. <https://doi.org/10.3389/fcvm.2021.670502>.
40. Guo Y, Chen Y, Tan M, Jia K, Chen J, Wang J. Content-aware convolutional neural networks. *Neural Netw*. 2021;143:657–68. <https://doi.org/10.1016/j.neunet.2021.06.030>.
41. Mieloszyk RJ, Bhargava P. Convolutional neural networks: the possibilities are almost endless. *Curr Probl Diagn Radiol*. 2018;47(3):129–30. <https://doi.org/10.1067/j.cpradiol.2018.01.008>.
42. Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput*. 2019;31(7):1235–70. https://doi.org/10.1162/neco_a_01199.
43. Chen H, He Y. Machine learning approaches in traditional chinese medicine: a systematic review. *Am J Chin Med*. 2022;50(1):91–131. <https://doi.org/10.1142/S0192415X22500045>.
44. Yao C, Spurlock DM, Armentano LE, Page CJ, VandeHaar MJ, Bickhart DM, et al. Random forests approach for identifying additive and epistatic single nucleotide polymorphisms associated with residual feed intake in dairy cattle. *J Dairy Sci*. 2013;96(10):6716–29. <https://doi.org/10.3168/jds.2012-6237>.
45. Li D, Hu J, Zhang L, Li L, Yin Q, Shi J, et al. Deep learning and machine intelligence: new computational modeling techniques for discovery of the combination rules and pharmacodynamic characteristics of traditional chinese medicine. *Eur J Pharmacol*. 2022;933:175260. <https://doi.org/10.1016/j.ejphar.2022.175260>.
46. Bi L, Kim J, Kumar A, Feng D, Fulham M. Synthesis of positron emission tomography (pet) images via multi-channel generative adversarial networks (GANS). Cham: Springer International Publishing; 2017. p. 43–51.
47. Porras G, Chassagne F, Lyles JT, Marquez L, Dettweiler M, Salam AM, et al. Ethnobotany and the role of plant natural products in antibiotic drug discovery. *Chem Rev*. 2021;121(6):3495–560. <https://doi.org/10.1021/acs.chemrev.0c00922>.
48. Li L, Zuo Z, Wang Y. Practical qualitative evaluation and screening of potential biomarkers for different parts of wolfiporia cocos using machine learning and network pharmacology. *Front Microbiol*. 2022;13:931967. <https://doi.org/10.3389/fmicb.2022.931967>.
49. Vatanserver S, Schlessinger A, Wacker D, Kaniskan HU, Jin J, Zhou MM, et al. Artificial intelligence and machine learning-aided drug discovery in central nervous system diseases: state-of-the-arts and future directions. *Med Res Rev*. 2021;41(3):1427–73. <https://doi.org/10.1002/med.21764>.
50. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A deep learning approach to antibiotic discovery. *Cell*. 2020;180(4):688–702. <https://doi.org/10.1016/j.cell.2020.01.021>.
51. Chen YX, Li F, Sun ZY, Zhou ZL, Wang W. Support vector machines analysis of free lipid compositions on cold or heat property of traditional chinese medicines. *Liaoning J Traditional Chin Med*. 2011;38(01):127–9. <https://doi.org/10.1016/j.jcell.2020.01.021>.
52. Chuang KV, Gunsalus LM, Keiser MJ. Learning molecular representations for medicinal chemistry. *J Med Chem*. 2020;63(16):8705–22. <https://doi.org/10.1021/acs.jmedchem.0c00385>.
53. Yang B, Bao W, Hong S. Alzheimer-compound identification based on data fusion and forgeNet_SVM. *Front Aging Neurosci*. 2022;14:931729. <https://doi.org/10.3389/fnagi.2022.931729>.
54. Yu X, Zhu X, Zhang L, Qin JJ, Feng C, Li Q. In silico screening and validation of PDGFRA inhibitors enhancing radioiodine sensitivity in thyroid cancer. *Front Pharmacol*. 2022;13:883581. <https://doi.org/10.3389/fphar.2022.883581>.
55. He S, Zhang C, Zhou P, Zhang X, Ye T, Wang R, et al. Herb-induced liver injury: phylogenetic relationship, structure-toxicity relationship, and herb-ingredient network analysis. *Int J Mol Sci*. 2019;20(15):3633. <https://doi.org/10.3390/ijms20153633>.
56. Brown AC. Kidney toxicity related to herbs and dietary supplements: online table of case reports. Part 3 of 5 series. *Food Chem Toxicol*. 2017;107(Pt A):502–19. <https://doi.org/10.1016/j.fct.2016.07.024>.
57. Hu X, Du T, Dai S, Wei F, Chen X, Ma S. Identification of intrinsic hepatotoxic compounds in polygonum multiflorum thunb. Using machine-learning methods. *J Ethnopharmacol*. 2022;298:115620. <https://doi.org/10.1016/j.jep.2022.115620>.
58. He S, Yi Y, Hou D, Fu X, Zhang J, Ru X, et al. Identification of hepatoprotective traditional chinese medicines based on the structure-activity relationship, molecular network, and machine learning techniques. *Front Pharmacol*. 2022;13:969979. <https://doi.org/10.3389/fphar.2022.969979>.
59. Chen Z, Zhao M, You L, Zheng R, Jiang Y, Zhang X, et al. Developing an artificial intelligence method for screening hepatotoxic compounds in traditional chinese medicine and western medicine combination. *Chin Med*. 2022;17(1):58. <https://doi.org/10.1186/s13020-022-00617-4>.
60. Wang H, Liu X, Lv B, Yang F, Hong Y. Reliable multi-label learning via conformal predictor and random forest for syndrome differentiation of chronic fatigue in traditional chinese medicine. *PLoS ONE*. 2014;9(6):e99565. <https://doi.org/10.1371/journal.pone.0099565>.
61. Shi Y, Yao X, Xu J, Hu X, Tu L, Lan F, et al. A new approach of fatigue classification based on data of tongue and pulse with machine learning. *Front Physiol*. 2021;12:708742. <https://doi.org/10.3389/fphys.2021.708742>.
62. Senoner T, Pfeifer B, Barbieri F, Adukauskaite A, Dichtl W, Bauer A, et al. Identifying the location of an accessory pathway in pre-excitation syndromes using an artificial intelligence-based algorithm. *J Clin Med*. 2021;10(19):4394. <https://doi.org/10.3390/jcm10194394>.
63. Sun GX, Yao XY, Yuan ZK, Zuo HN, Hao WH. The realization of the bp neural network model based on the matlab coronary heart disease of TCM syndrome. *Chin Archives Traditional Chin*. 2011;29(08):1774–6. <https://doi.org/10.3390/jcm10194394>.
64. Zhang H, Ni W, Li J, Zhang J. Artificial intelligence-based traditional chinese medicine assistive diagnostic system: validation study. *JMIR Med Inform*. 2020;8(6):e17608. <https://doi.org/10.2196/17608>.
65. Zhao Y, Huang Y. Quantitative diagnosis of TCM syndrome types based on adaptive resonant neural network. *Comput Intell Neurosci*. 2022;2022:2485089. <https://doi.org/10.1155/2022/2485089>.
66. Zhi L, Zhang D, Yan JQ, Li QL, Tang QL. Classification of hyperspectral medical tongue images for tongue diagnosis. *Comput Med Imaging Graph*. 2007;31(8):672–8. <https://doi.org/10.1016/j.compmedimag.2007.07.008>.
67. Qi Z, Tu LP, Chen JB, Hu XJ, Xu JT, Zhang ZF. The classification of tongue colors with standardized acquisition and ICC profile correction in traditional Chinese medicine. *Biomed Res Int*. 2016;2016:3510807. <https://doi.org/10.1155/2016/3510807>.
68. Liu C. A study of tongue features in children with tic disorders of kidney emotion deficiency based on decision tree and neural network. *Shandong university of Chinese medicine*. 2020.
69. Yan JJ, Li XD, Guo R, Yan HX, Wang YL. Research on classification of dentate tongue based on deep learning and random forest. *Chin Archives Traditional Chin Med*. 2022;40(02):19–22.

70. Lu PH, Chiang CC, Yu WH, Yu MC, Hwang FN. Machine learning-based technique for the severity classification of sublingual varices according to traditional Chinese medicine. *Comput Math Methods Med*. 2022;2022:3545712. <https://doi.org/10.1155/2022/3545712>.
71. Wang Y, Shi X, Li L, Efferth T, Shang D. The impact of artificial intelligence on traditional Chinese medicine. *Am J Chin Med*. 2021;49(6):1297–314. <https://doi.org/10.1142/S0192415X21500622>.
72. Huo C, Zheng H, Su H, Sun Z, Cai Y, Xu Y. Tongue shape classification integrating image preprocessing and convolution neural network. In: 2017 2nd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS). 2017:42–6.
73. Xiao QX, Zhang J, Zhang H, Li XG, Zhou L. Tongue coating color classification based on shallow convolutional neural network. *Meas Control Technol*. 2019;38(03):26–31.
74. Christian S, Wei L, Jia YQ, Pierre S, Scott R, Dragomir A et al. Going deeper with convolutions. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015; pp. 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
75. Shao YW. Research on intelligent tongue diagnosis based on deep learning. Xiamen University; 2018.
76. Chen HZ. Research on application of tongue recognition model based on convolutional neural network. Yanshan University, 2019.
77. Qu Y, Zhuo Y, Lee J, Huang X, Yang Z, Yu H, et al. Ischemic and haemorrhagic stroke risk estimation using a machine-learning-based retinal image analysis. *Front Neurol*. 2022;13:916966. <https://doi.org/10.3389/fneur.2022.916966>.
78. Sun XH, Fu ZT, Yan L, Zhou ZJ. Application research of efficientNet on eye recognition of yin deficiency syndrome. *Inform Traditional Chin Med*. 2020;37(03):29–34 <https://doi.org/10.3389/fneur.2022.916966>.
79. Xu WZ, Yu K, Xu JJ, Ye JJ, Li HM, Shu Q. Artificial intelligence technology in cardiac auscultation screening for congenital heart disease: present and future. *J Zhejiang Univ (Med Sci)*. 2020;49(05):548–55.
80. Fernando T, Ghaemmaghami H, Denman S, Sridharan S, Hussain N, Fookes C. Heart sound segmentation using bidirectional LSTMs with attention. *IEEE J Biomed Health Inform*. 2020;24(6):1601–9. <https://doi.org/10.1109/JBHI.2019.2949516>.
81. Liu J. Exploration and application of artificial intelligence technology in screening of heart sounds with auscultation in children with congenital heart defect. Chongqing Medical University, 2021.
82. Luo ZY, Cui J, Hu XJ, Tu LP, Liu HD, Jiao W, et al. A study of machine-learning classifiers for hypertension based on radial pulse wave. *Biomed Res Int*. 2018;2018:2964816. <https://doi.org/10.1155/2018/2964816>.
83. Lee BJ, Jeon YJ, Ku B, Kim JU, Bae JH, Kim JY. Association of hypertension with physical factors of wrist pulse waves using a computational approach: a pilot study. *BMC Complement Altern Med*. 2015;15:222. <https://doi.org/10.1186/s12906-015-0756-7>.
84. Tang M, Gao L, He B, Yang Y. Machine learning based prognostic model of Chinese medicine affecting the recurrence and metastasis of i–iii stage colorectal cancer: a retrospective study in China. *Front Oncol*. 2022;12:1044344. <https://doi.org/10.3389/fonc.2022.1044344>.
85. Zhang H, Zhang J, Ni W, Jiang Y, Liu K, Sun D, et al. Transformer- and generative adversarial network-based inpatient traditional Chinese medicine prescription recommendation: development study. *JMIR Med Inform*. 2022;10(5):e35239. <https://doi.org/10.2196/35239>.
86. Dong X, Zheng Y, Shu Z, Chang K, Xia J, Zhu Q, et al. TCMPr: TCM prescription recommendation based on subnetwork term mapping and deep learning. *Biomed Res Int*. 2022;2022:4845726. <https://doi.org/10.1155/2022/4845726>.
87. Lu H, Zhang J, Liang Y, Qiao Y, Yang C, He X, et al. Network topology and machine learning analyses reveal microstructural white matter changes underlying Chinese medicine dengzhan shengmai treatment on patients with vascular cognitive impairment. *Pharmacol Res*. 2020;156:104773. <https://doi.org/10.1016/j.phrs.2020.104773>.
88. Liu J, Huang Q, Yang X, Ding C. Hpe-gcn: predicting efficacy of tonic formulae via graph convolutional networks integrating traditionally defined herbal properties. *Methods*. 2022;204:101–9. <https://doi.org/10.1016/j.jymeth.2022.05.003>.
89. Zhang J, Chen Z, Wang F, Xi Y, Hu Y, Guo J. Machine learning assistants construct oxidative stress-related gene signature and discover potential therapy targets for acute myeloid leukemia. *Oxid Med Cell Longev*. 2022;2022:1507690. <https://doi.org/10.1155/2022/1507690>.
90. Zhang Q, Guo Y, Zhang B, Liu H, Peng Y, Wang D, et al. Identification of hub biomarkers of myocardial infarction by single-cell sequencing, bioinformatics, and machine learning. *Front Cardiovasc Med*. 2022;9:939972. <https://doi.org/10.3389/fcvm.2022.939972>.
91. Yuan J, Wang ZZ, Song LJ, Xue Y, Zhang WJ. Study on prediction of compound-target-disease network of corydalis yanhusuo based on supervised learning. *Hainan Med J*. 2020;31(13):1638–43 <https://doi.org/10.3389/fneur.2022.916966>.
92. Cong Y, Yang XG, Lv W, Xue Y. Prediction of novel and selective TNF-alpha converting enzyme (TACE) inhibitors and characterization of correlative molecular descriptors by machine learning approaches. *J Mol Graph Model*. 2009;28(3):236–44. <https://doi.org/10.1016/j.jmglm.2009.08.001>.
93. Wang Y, Jafari M, Tang Y, Tang J. Predicting meridian in Chinese traditional medicine using machine learning approaches. *Plos Comput Biol*. 2019;15(11):e1007249. <https://doi.org/10.1371/journal.pcbi.1007249>.
94. Zhang F, Li J, Wang Y, Guo L, Wu D, Wu H, et al. Ensemble learning based on policy optimization neural networks for capability assessment. *Sens (Basel)*. 2021;21(17):5802. <https://doi.org/10.3390/s21175802>.
95. Fu SY, Zheng H, Yang ZJ, Yan B, Su HY, Liu YP. Computerized tongue coating nature diagnosis using convolutional neural network. In 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), 2017; pp.730–4. <https://doi.org/10.1109/ICBDA.2017.8078732>
96. Yang R, Zhao G, Yan B. Discovery of novel c-JUN n-terminal kinase 1 inhibitors from natural products: integrating artificial intelligence with structure-based virtual screening and biological evaluation. *Molecules*. 2022;27(19):6249. <https://doi.org/10.3390/molecules27196249>.
97. Ge H, Yan Z, Yu W, Sun L. An attention mechanism based convolutional LSTM network for video action recognition. *Multimed Tools Appl*. 2019;78(14):20533–56. <https://doi.org/10.1007/s11042-019-7404-z>.
98. Jie W, Lian D, Hongzheng L, Jinlei L, Hengwen C. Construction of an artificial intelligence traditional Chinese medicine diagnosis and treatment model based on syndrome elements and small-sample data. *Engineering-Proc*. 2022;8(01):29–32.
99. Computational biology. Codon Publications: Brisbane (AU); 2019.
100. Fang J, Liu C, Wang Q, Lin P, Cheng F. In silico polypharmacology of natural products. *Brief Bioinform*. 2018;19(6):1153–71. <https://doi.org/10.1093/bib/bbx045>.
101. Fu X, Mervin LH, Li X, Yu H, Li J, Mohamad ZS, et al. Toward understanding the cold, hot, and neutral nature of Chinese medicines using in silico mode-of-action analysis. *J Chem Inf Model*. 2017;57(3):468–83. <https://doi.org/10.1021/acs.jcim.6b00725>.
102. Zhang YQ, Wang N, Du X, Chen T, Yu ZC, Qin YW, et al. SoFDA: an integrated web platform from syndrome ontology to network-based evaluation of disease-syndrome-formula associations for precision medicine. *Sci Bull (Beijing)*. 2022;67(11):1097–101. <https://doi.org/10.1016/j.scib.2022.03.013>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

