IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 15, 2019, accepted June 3, 2019, date of publication June 6, 2019, date of current version June 25, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2921318

Intelligent Syndrome Differentiation of Traditional Chinese Medicine by ANN: A Case Study of Chronic Obstructive Pulmonary Disease

QIANG XU[1], WENJUN TANG[2], FEI TENG[3], (Member, IEEE), WEI PENG[1], YIFAN ZHANG[4], WEIHONG LI[5], CHUANBIAO WEN[1], AND JINHONG GUO[6], (Member, IEEE)

[1]College of Medical Information Engineering, Chengdu University of Traditional Chinese Medicine, Chengdu 611130, China
[2]Affiliated Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu University of Traditional Chinese Medicine, Chengdu 611130, China
[3]School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China
[4]Department of Respiratory, Chengdu First People's Hospital, Chengdu 610016, China
[5]School of Basic Medical Science, Chengdu University of Traditional Chinese Medicine, Chengdu 611130, China
[6]School of Information and Communication Engineering, Institute of Medical Equipment, University of Electronic Science and Technology of China, Chengdu 611731, China

Corresponding authors: Weihong Li (lwh@cdutcm.edu.cn), Chuanbiao Wen (wcb@cdutcm.edu.cn), and Jinhong Guo (guojinhong@uestc.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 81804220 and Grant 81804222, in part by the Sichuan Science and Technology Department Project under Grant 2019YFS0148 and Grant 2018SZ0065, and in part by the China Postdoctoral Science Foundation under Grant 2018M643429.

ABSTRACT Traditional Chinese medicine (TCM) is effective in preventing and treating all manner of diseases, which has been incorporated into the latest global medical outline (Ver.2019) by World Health Organization (WHO). As one of the most important characteristics of TCM, syndrome differentiation (SD) provides curative effect assurance. SD is a high-dimensional complex function with symptoms/signs as input and syndrome type as output. Artificial neural network (ANN) provides an all-purpose data-driven solution to fit high-dimensional complex function, making ANN a promising approach for modeling intelligent SD (ISD) for TCM. In this paper, we chose chronic obstructive pulmonary disease (COPD) as an example for investigating ISD for TCM based on ANN. First, we built a full-group ANN model that combines ANN with full-group datasets composed of 18 471 real clinical records. In addition, we built four extra models with ANN and four subgroup datasets. For comparison, we built another four models with four traditional machine-learning algorithms and the full-group datasets. We used accuracy and F1-scores to evaluate the models' performance. With an accuracy of 86.45% and an F1 score of 82.93%, the full-group ANN model outperformed the four comparison models built from traditional machine-learning algorithms, and however, the four subgroup models achieved a better performance than the full-group ANN model. We concluded that the ANN can potentially provide a way for ISD for TCM, and our subgroup modeling suggests ideas for further optimizing the ISD.

INDEX TERMS Artificial neural network, traditional Chinese medicine, COPD, intelligent syndrome differentiation, subgroup modeling.

I. INTRODUCTION

According to the latest global medical outline (Ver.2019) [1] of the world Health Organization (WHO), traditional Chinese medicine (TCM), which is a popular complementary and alternative medicine, is effective in preventing and treating many diseases [2]–[4]. In contrast to the approach of western medicine (WM), TCM provides a unique methodology to manage diseases, especially for disease diagnosis. WM practitioners place more emphasis on the collection of the symptoms/signs than on the classification of diseases. The TCM specialists collect symptoms/signs by means of

The associate editor coordinating the review of this manuscript and approving it for publication was Abdullah Iliyasu.

clinical observation from the patient's self-reporting, which cannot be accurately measured and impressionable to the study sample and environmental factors. Rather than the other way around, TCM has a particularly significant classification task that is called syndrome differentiation (also known as pattern classification or ZHENG differentiation). The syndrome (ZHENG or pattern) is differentiated by a set of symptoms/signs determined by TCM practitioners using four main diagnostic methods: inspection, auscultation/olfaction, interrogation, and palpation, which reflect the pathological and physiological changes of disease occurrence and development. Consequently, syndrome differentiation (SD) is one of the most important characteristics that lead to the curative effect of TCM.

From the perspective of mathematics, the SD procedure can be rated as a classification function with symptoms/signs as input and syndrome types as output. With the development of new computational methods, increasing investigators are using a wealth of approaches to fit this function and establish computation-assisted syndrome differentiation systems. However, some investigators have focused on only one of the four diagnostic methods. For inspection diagnosis, Watsuji *et al.* [5] built a tongue inspection diagnosis system with the fuzzy rules based on the tongue image databases alone. Kanawong *et al.* [6] proposed a ZHENG classification system with several classification algorithms, such as Support Vector Machine, Multilayer Perceptron Networks and Random Forest, but their system was also based on tongue images alone. Similar studies on the relationship between tongue signs and syndromes include the research of Siu *et al.* [7] and others. For auscultation diagnosis, Yan *et al.* [8] used SVM algorithm and the nonstationarity information of the auscultation signals to automatically recognize healthy individuals from Qi-deficiency or Yin-deficiency. Yang and Chiu [9] also used the fractal dimension parameter with logistic regression. For interrogation diagnosis, Li *et al.* [10] investigated the symptom-syndrome interactions for inquiry diagnosis of coronary heart disease (CHD). Liu *et al.* [11] introduced multilabel learning algorithms to make this diagnosis. Similar studies have been carried out by other researchers [12], [13]. In general, the aforementioned studies were oriented toward the relationships between separate diagnosis and syndrome types. However, in TCM, the synthesis of the four diagnostic methods is the most critical principle for syndrome differentiation (SD). Therefore, these studies have been incomplete. A wealth of research [14]–[17] has been done to model intelligent syndrome differentiation (ISD) based on electronic medical records (EMRs), which include all four diagnosisic informations. However, the data from the EMRs, which were manually input by TCM clinicians, are unstructured, so the performance of the final model is inevitably impacted. Moreover, scholars commonly chose the traditional machine-learning algorithms (such as Softmax regression, Navive-Bayes, support vector machine, and K-nearest neighbor) to develop the relationship between symptoms/signs and syndrome types. Nevertheless, in real clinical practice, because one symptom is related to several syndromes, the relationships between symptoms and syndromes are phenomenally complex. In addition, a patient may suffer from several diseases at the same time, so there may be several syndromes. Furthermore, syndromes may change during the TCM treatment of the disease, so the syndrome differentiation can be ever changing [18]. Therefore, in a more comprehensive view, TCM is increasingly regarded as a black box model, and the classification function (or the relationship between symptoms/signs and syndrome types) should be multidimensional [19], [20] and of prime complexity. These approaches are associated with underfitting or overfitting by the traditional machine learning algorithms with respect to real clinical practice.

In recent years, with the increase in computing efficiency and accessible real world big data, the technology of ANN has made remarkable progress. Nowadays we can train the network designed with large size of layers and neurons by the big datasets, which is gradually evolved to the deep neural network (DNN). However, in the AI field, many researchers regard ANN as a black box, and it is hitherto difficult to provide rigorous mathematical arguments to prove that ANN outshines the traditional machine learning algorithms. We can only arrive at the conclusions by the intuitive analysis that compared with traditional machine learning algorithms, ANN, in special, with large size of layers and neurons covers more abstract features extracted from the data, thereby it exerts a better performance. Moreover, the success of ANN in many other scenarios inspired us to merge the merit of ANN into ISD of TCM.

Chronic obstructive pulmonary disease (COPD) is a common disease of the respiratory system with high morbidity and mortality. It is expected to become the 3^{rd} leading cause of death in the world by 2020 [21]. Increasing evidences have indicated that TCM is effective in the prevention and treatment of COPD [22], [23]. In this paper, we chose COPD as an example for investigating the benefits using ANN for intelligent syndrome differentiation in TCM. It must be mentioned that in many cases a more complex network is prone to overfitting or underfitting, especially when the training data is not 'big' enough, so there is a tradeoff between efficiency and complexity of the network. Fueled by the scale of datasets in our experiments (18,471), we designed a network with just 10 hidden layers. The contributions of this paper are as follows:

(1) In TCM science, as a critical procedure for managing diseases, syndrome differentiation (SD) can be considered as a high-dimensional complex function with symptoms/signs as input and syndrome types as output, and in computer science, ANN provides a general method for fitting such high-dimensional complex functions [24]. Consequently, in this paper, We merged the merit of ANN into ISD of TCM and modeled the ISD of COPD by ANN. This model could be applied to similar diseases that share common characteristics with COPD.

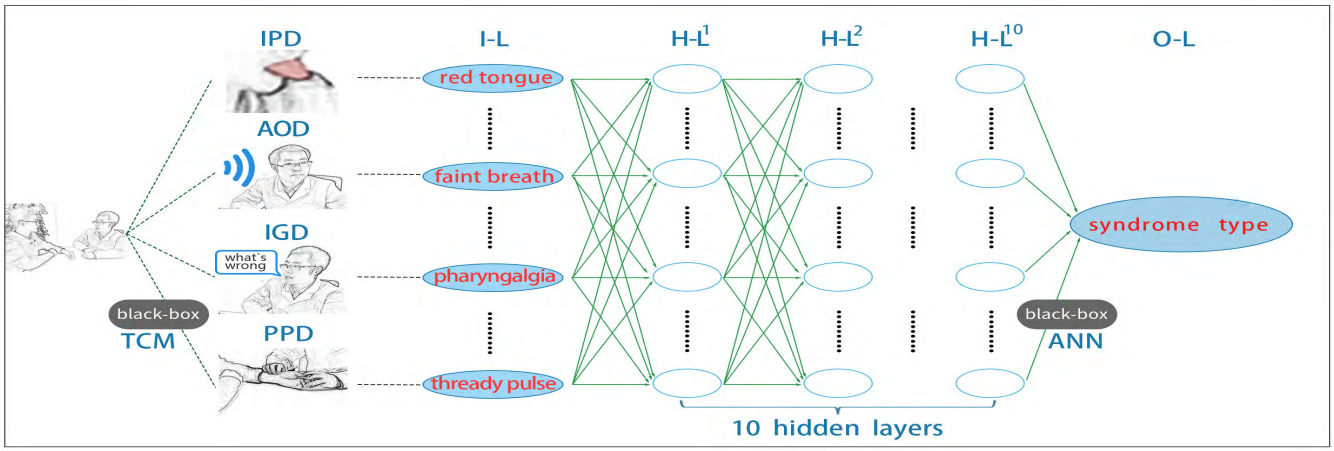


FIGURE 1. The topological architecture of ANN, where'IPD' denotes inspection diagnosis, 'IGD' denotes interrogation diagnosis, 'AOD' denotes auscultation and olfaction diagnosis, 'PPD' denotes palpation diagnosis, 'I-L' denotes input layer, 'H-L' denotes hidden layer, 'O-L' denotes output layer. The similar property of 'black-box' inspire us to merge ANN into TCM.

(2) In some cases, datasets are more critical for modeling than the algorithms. In contrast to previous studies, we collected 18,471 samples in our study from the real clinical practices. Moreover, all the data were evaluated by professional TCM practitioners to be sure that the data were reliable. Additionally, accuracies and F1-scores were used to evaluate the model's performance, and we designed four comparison models using traditional machine learning algorithms to guarantee the objectivity of our study.

(3) In our work, we separated four subgroups from the datasets by a prior grouping method to reduce the exploration space of the network, and we built four subgroup models. The results demonstrated that the subgroup models achieved a better performance than the full-group model.

The remainder of this paper is organized as follows. Section 2 briefly introduces the algorithm and the topological architecture of ANN. The details of our initial experiments, intervening additional experiments, and comparison experiments are described in Section 3. In Section 4, we describe the results of all the experiments. Finally, in Section 5, we present our conclusions.

II. ARCHITECTURE AND ALGORITHMS

The diagnostic process of TCM can be separated into symptoms/signs collection and syndrome differentiation. In clinical practice, TCM specialists will collect symptoms/signs by inspection, auscultation/olfaction, interrogation, and palpation, followed by a comprehensive analysis of the collected symptoms/signs for a purpose of classification with respect to syndrome types. This analysis involves a complex black-box model, which can be fitted by the topological architecture of ANN, as shown in Fig.1. By data-driven means, ANN combines low-level features into more abstract high-level features to create a representation of the distributed features. However, more features require more complex computations and more space for exploration, and the model fails to perform as we expected, especially when the training data is not 'big' enough. Fueled by the scale of datasets in our experiments (18,471), we designed a network with just 10 hidden layers to limit the complexity of the model. In the next section, we will briefly review the algorithms of ANN based on the architecture illustrated in Fig.1.

A. THE FORWARD PROPAGATION ALGORITHMS OF ANN

To obtain the input for the network, one starts at the first layer and the network carries out a series of linear operations and activation operations layer by layer with corresponding weight matrixes and bias vectors until it reaches the last layer, which is the output of the network. The input of the network is defined as follows:

$$\vec{x} = [x_1, x_2, \ldots, x_i, \ldots, x_m], \quad i = 1, 2, 3, \ldots, m \tag{1}$$

Correspondingly, the output of the network is described as:

$$\vec{y} = [y_1, y_2, \ldots, y_k, \ldots, y_n], \quad k = 1, 2, 3, \ldots, n \tag{2}$$

The output of neuron units in the l^{th} hidden layer is expressed by Eq.3, which is calculated from the Eq.4 and Eq.5, where W_{ij}^l represents the weights that combines the i^{th} neuron of the $l-1^{th}$ layer to the j^{th} neuron of the l^{th} layer, b_i^l represents the bias of the l^{th} layer, net_i^l represents the output of i^{th} in the l^{th} layer, and $\sigma(x)$ represents the activation function.

$$h^l = [h_1^l, h_2^l, \ldots, h_j^l, \ldots, h_{sl}^l], \quad j = 1, 2, 3, \ldots, sl \tag{3}$$

$$h_i^l = \sigma(net_i^l) \tag{4}$$

$$net_i^l = \sum_{j=1}^{l} W_{ij}^l h_j^{l-1} + b_i^l \tag{5}$$

B. THE BACK PROPAGATION ALGORITHMS OF ANN

The back propagation algorithm is used to update the weights and biases of all neurons. The updating rules are shown in

Eq.6 and Eq.7, where $E(x)$ refers to the loss function and η represents the learning rate.

$$W_{ij}^{l} = W_{ij}^{l} - \eta \frac{\partial E}{\partial W_{ij}^{l}} \tag{6}$$

$$b_{i}^{l} = b_{i}^{l} - \eta \frac{\partial E}{\partial b_{i}^{l}} \tag{7}$$

The partial derivatives of the weights and biases of the output layer are calculated as follows, where x_k and y_k refer to the expected and actual output of the network, respectively.

$$\frac{\partial E}{\partial W_{kj}^{L}} = -(x_k - y_k)\sigma'(x)|_{x=net_k^L} h_j^{L-1} \tag{8}$$

$$\frac{\partial E}{\partial b_{k}^{L}} = -(x_k - y_k)\sigma'(x)|_{x=net_k^L} \tag{9}$$

Let

$$\delta_k^L = -(x_k - y_k)\sigma'(x)|_{x=net_k^L} \tag{10}$$

Then

$$\frac{\partial E}{\partial W_{kj}^{L}} = \delta_k^L h_j^{L-1} \tag{11}$$

$$\frac{\partial E}{\partial b_{k}^{L}} = \delta_k^L \tag{12}$$

And let

$$\delta_k^{L-1} = \sum_{k=1}^{n} W_{kj}^{L} \delta_k^L \sigma'(x)|_{x=net_k^L} \tag{13}$$

Next, the partial derivatives of the weights and biases of the $l^{th}(2 \leq or \leq l \leq or \leq L-1)$ layer are calculated as follows: where δ_j^{L-1} is calculated as the Eq.16

$$\frac{\partial E}{\partial W_{kj}^{L-1}} = \delta_j^{L-1} h_i^{L-2} \tag{14}$$

$$\frac{\partial E}{\partial b_{j}^{L-1}} = \delta_j^{L-1} \tag{15}$$

$$\delta_j^{L-1} = \sum_{k=1}^{n} (x_k - y_k)\sigma'(x)|_{x=net_k^L} W_{kj}^{L} \sigma'(x)|_{x=net_k^L} \tag{16}$$

III. EXPERIMENTS

A. DATASETS

1) DATA SOURCES

We used the experimental data from 3,237,645 outpatient electronic medical records (EMRs) of the digital diagnosis and treatment platform of TCM, which is managed by Sichuan Provincial Administration of TCM, China. The data are from 33 medical institutions in Sichuan Province, China from October 2015 to January 2019. We screened 18,471 (5.79%) of the records that had complete four-diagnostic symptoms/signs and syndromes for cases diagnosed as COPD, and we defined the 18,471 records as the full-group dataset.

2) DATA CLEANING

The data for the EMRs was manually input by TCM practitioners, which can lead to some problems, such as incorrect entries, missing entries, duplication, and other irregularities, so it was necessary to subject the data to cleaning before modeling. Each record was cross-validated by two primary physicians, and final validation was conducted by the resident physician. The standard for the four-diagnostic symptoms/signs was based on both the National Standard of GB/T 20348-2006 of China and the reference Differential Diagnosis of TCM Symptoms [25]. For example, symptoms in original records such as ''stupo'', ''poor appetite'', and ''low appetite'' were standardized as ''eating less''. We summarized 321 standardized four-diagnostic symptoms from the records. The standard of syndrome types was based on both the Trade Standard of ZY/T001.1.94 of China and the Guidelines for TCM Diagnosis and Treatment of COPD [26]. The syndrome types of the 18,471 records were categorized into nine classes: L1-syndrome of lung qi deficiency, L2-syndrome of lung-spleen deficiency, L3-syndrome of lung-kidney deficiency, L4-syndrome of qi-yin deficiency, L5-syndrome of wind-cold attacking lung, L6-syndrome of external cold and internal drinking, L7-syndrome of lung phlegm-turbidity, L8-syndrome of lung phlegm-heat, and L9-syndrome of lung phlegm-deceiving the mind.

3) DATA NORMALIZATION

Sample input is the symptoms/signs of each sample, which is set as a 1×321vector. Let $x_i = [x_{i,1}, x_{i,2}, x_{i,3,\ldots}, x_{i,321}]$ be the vector, each element satisfies the following rule which is shown in the Eq.17 to take value.Correspondingly. Sample label is the syndrome types of each sample, which is set as a 1×9vector. Let $l_i = [l_{i,1}, l_{i,2}, l_{i,3,\ldots}, l_{i,9}]$ be the label of the sample, each element satisfies the following rule which is shown in the Eq.18 to take value.

$$x_{i,j} = \begin{cases} 1, & the\ j^{th}\ symptoms\ exists \\ 0, & the\ j^{th} symptoms\ not\ exists \end{cases} \quad j = 1, 2, 3, \ldots, 321 \tag{17}$$

$$l_{i,j} = \begin{cases} 1, & the\ lable\ of\ l_i \in the\ j^{th}\ class \\ 0, & otherwise \end{cases} \quad j = 1, 2, 3, \ldots, 9 \tag{18}$$

B. EXPERIMENT BASED ON ANN ALGORITHM AND FULL-GROUP DATASETS

We used the ANN algorithm described in Section 2, and the experimental data was from the full-group datasets described in Section 3-A. 75% of the data was randomly selected as training data, and 25% was used as test data. See Table 2. One 321-dimension layer was included in the network as the input layer. We designed nine fully-connected hidden layers, and the end of the network was one 4-class Softmax layer as

the output layer. In order to keep the optimization algorithm robust and optimize efficiency, the dropout ratio was set at 0.5, and the learning rate was set at 0.01. Meanwhile, the initialization network weights were sampled in a Gaussian distribution $\mathcal{N}(0, 0.01)$, and the bias was initialized to 0. We used a min-batch size of 50, and the loss function was defined as follows:$\mathcal{L} = \frac{1}{2m}\sum_{k=1}^{m}\|l_k - y_k\|$, where m represents the number of the samples, l_k represents the real label value vector of the k^{th} training sample, y_k represents the prediction label value vector of the network.We defined the above model as full-group ANN model which is based on ANN algorithm and full-group datasets.

C. ADDITIONAL EXPERIMENTS BASED ON SUBGROUP DATASETS AND ANN

The ABCD grading system is the latest COPD disease assessment method suggested by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) [27]. The A group patients have fewer symptoms and low risk, the B group patients have more symptoms and low risk, the C patients have less symptoms and high risk, and the D patients have more symptoms and high risk. The patients with COPD are divided into these groups by assessing abnormal pulmonary function, future acute exacerbation risk, and complications. We manually separated patients from full-group datasets into the four subgroups using the diagnostic conclusions in the original data, i.e. A-group datasets, B-group datasets, C-group datasets, and D-group datasets. Table 1 provides the details of the distribution of labels in each subgroup. It should be noted that because many of the records did not include a definite grouping diagnosis by clinicians, the total number of datasets for the four subgroups was 8,173, which was less than the number of full-group datasets. We also chose 75% of the data from each subgroup as training data and 25% as test data (as shown in table.2), and we used the same ANN algorithm, the same network structure, and the same super-parameters that were described in the Section 3-B to train the four subgroups of data. Finally, the corresponding models were defined as A-group ANN model, B-group ANN model, C-group ANN model and D-group ANN model.

D. COMPARISON EXPERIMENTS BASED ON TRADITIONAL MACHINE LEARNING ALGORITHMS AND FULL-GROUP DATASETS

For the purpose of comparison, we chose traditional machine-learning algorithms (such as Softmax regression (SR), Navive-Bayes (NB), K-nearest neighbor (KNN), and support vector machine (SVM)) to manage the classification task. To make the comparison more objective, training data and test data were both from the full-group datasets, 75% of which were training data and the rest were test data. Syndrome differentiation is a multi-classification task, and SVM algorithms are usually used to resolve binary classification problems, so we had to improve the algorithm to deal with the multi-classification task described here. In our improved scheme, we used nine classes in all samples, and we constructed a multi-classifier that consisted of nine binary classifiers. We adopted a one-against-all strategy and used a radial bias function (RBF) kernel. We achieved the best value of a gamma function parameter (38.21) and loss function parameter (58.00) by the grid-search method. We defined the above models as the full-group SR model, the full-group NB model, the full-group SVM model, and the full-group KNN model. They were all based on traditional machine-learning algorithms and the full-group datasets.

TABLE 1. The distribution of labels in each group.

Group	L1	L2	L3	L4	L5	L6	L7	L8	L9
A_group	123	341	256	105	320	896	112	360	13
B_group	28	462	358	351	256	652	253	430	4
C_group	2	236	621	123	332	260	172	153	12
D_group	63	96	123	254	18	110	110	132	36
full_group	1247	3215	1641	1522	2864	3523	1456	3020	253

TABLE 2. The size of the training data and testing data in each group.

Group	Total Size	Training Size	Training Ratio	Test Size	Test Ratio
full_group	18471	13766	75%	4685	25%
A_group	2526	1894	75%	632	25%
B_group	2794	2095	75%	699	25%
C_group	1911	1433	75%	478	25%
D_group	942	707	75%	235	25%

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. THE PERFORMANCE OF THE FULL-GROUP ANN MODEL

For the full-group ANN model, Fig.2 shows the curve of accuracy and loss during the training process. The accuracy was calculated in term of $\frac{Pr}{Tr} \times 100\%$, where Pr refers to the number of correct predictions, and Tr refers to the total number of test samples. We defined the maximum accuracy of the training process as that of the model, and the accuracy of full-group ANN model was 86.45%. In addition, according to Table 1, the distribution of labels was unbalanced in our experiments. To more objectively reflect the performance of all models in this paper, we report the F1-score ($\frac{1}{c}\sum_{k=1}^{c}\frac{2p_k \times r_k}{p_k + r_k}$) [28], [29] of all the experimental results, where p_k and r_k represent the precision and recall of the labels, respectively, and c indicates the number of label categories. Table 3 shows the confusion matrix of the full-group ANN model. We can calculate the precision and recall from the matrix, and the value of the F1-score of the full-group ANN model was 82.93%. From the accuracy (86.45%) and the F1-score (82.93%) of the full-group ANN model, we conclude that ANN can be useful for the construction of an ISD model of TCM that can reach the middle level of TCM physicians. This paper also provides a method for ISD for other diseases identified by TCM.

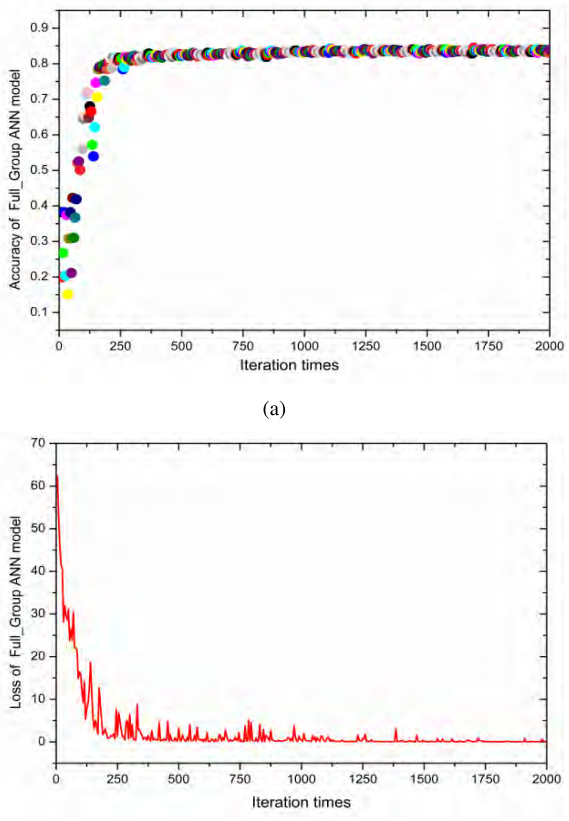


FIGURE 2. The accuracy and loss cure of full-group ANN model.

TABLE 3. The confusion matrix of the full-group ANN model.

lable	L1	L2	L3	L4	L5	L6	L7	L8	L9
L1	535	23	2	2	12	0	0	7	0
L2	23	623	4	1	0	1	1	0	1
L3	41	3	365	0	7	8	2	45	2
L4	2	41	3	364	0	36	0	0	35
L5	0	7	7	0	583	0	7	1	3
L6	23	42	1	1	0	661	0	0	21
L7	1	24	47	0	7	0	223	0	25
L8	10	8	25	12	35	1	0	624	0
L9	7	11	5	1	0	0	1	0	72

B. THE PERFORMANCE OF THE MODELS BASED ON SUBGROUP DATASETS

Tables 4–7 report the confusion matrix of each subgroup model. The precision and recall of each label in each model can be calculated, as shown in the Fig.3. The results show that the precision and recall of the subgroup models had a wider range than the full-group model, which was, in the main, due to more imbalance of the datasets in the subgroup datasets. Nevertheless, the accuracies and the F1-scores shown in table8 show that the subgroup models one-up the full-group model. It is well known that ANN provides a general data-driven method to fit high-dimensional complex functions, but the more complex and the higher dimensional the functions are, the more data and computing space the network

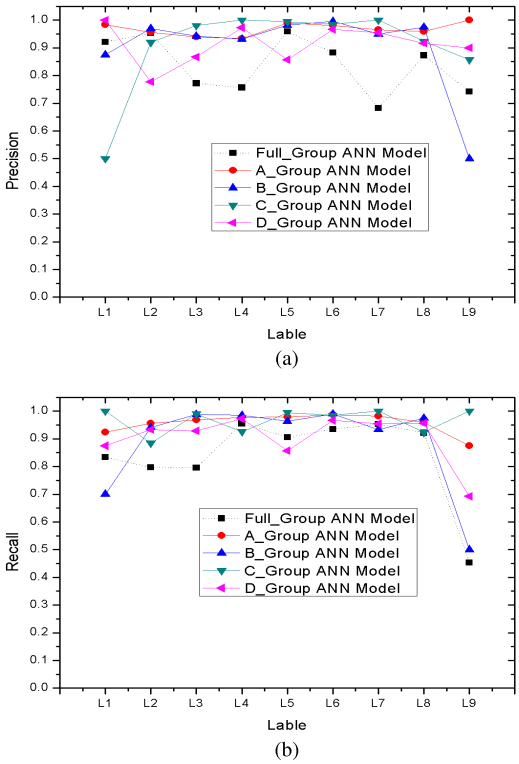


FIGURE 3. The precision and recall of the all labels of subgroup models compared with full-group ANN model.

TABLE 4. The confusion matrix of the A-group ANN model.

lable	L1	L2	L3	L4	L5	L6	L7	L8	L9
L1	61	0	0	0	0	0	0	1	0
L2	1	86	0	1	0	1	0	0	1
L3	2	2	62	0	0	0	0	0	0
L4	1	0	0	42	0	1	1	0	0
L5	0	0	0	0	95	0	0	1	0
L6	1	0	2	0	0	156	0	0	0
L7	0	2	0	0	0	0	56	0	0
L8	0	0	0	0	2	0	0	47	0
L9	0	0	0	0	0	0	0	0	7

requires. Abstractly speaking, we can regard the full-group ANN model based on the full-group datasets as a complex function. Accordingly, the subgroup models trained by the subgroup datasets express the above functions in segments. In a sense, subgroup models can be thought to be piecewise functions, which can reduce the order or dimension of more complex functions and thereby reduce the exploration space for the neural network and improve the model performance. The strategy of submodeling is an important innovation in this paper, which has proved by the additional experiments to be effective for further optimizing the ISD of TCM.

C. THE PERFORMANCE OF THE MODELS BASED ON TRADITIONAL MACHINE LEARNING ALGORITHMS

Table 8 shows the accuracies and the F1 scores of the all traditional machine learning models compared with full-group ANN model. The F1 scores are calculated from

TABLE 5. The confusion matrix of the B-group ANN model.

lable	L1	L2	L3	L4	L5	L6	L7	L8	L9
L1	7	0	0	0	0	0	0	1	0
L2	1	96	0	1	0	1	0	0	0
L3	0	2	81	0	1	0	0	0	2
L4	1	0	0	68	0	1	2	1	0
L5	0	0	0	0	53	0	0	1	0
L6	0	1	0	0	0	198	0	0	0
L7	0	3	0	0	0	0	56	0	0
L8	1	0	1	0	1	0	0	114	0
L9	0	0	0	0	0	0	2	0	2

TABLE 6. The confusion matrix of the C-group ANN model.

lable	L1	L2	L3	L4	L5	L6	L7	L8	L9
L1	1	0	0	0	0	0	0	1	0
L2	0	23	0	1	0	1	0	0	0
L3	0	2	99	0	0	0	0	0	0
L4	0	0	0	25	0	0	0	0	0
L5	0	0	0	0	189	0	0	1	0
L6	0	1	0	0	0	65	0	0	0
L7	0	0	0	0	0	0	36	0	0
L8	0	0	1	0	1	0	0	24	0
L9	0	0	0	1	0	0	0	0	6

TABLE 7. The confusion matrix of the D-group ANN model.

lable	L1	L2	L3	L4	L5	L6	L7	L8	L9
L1	21	0	0	0	0	0	0	0	0
L2	0	14	0	1	0	1	1	0	1
L3	0	1	26	0	0	0	0	0	3
L4	2	0	0	71	0	0	0	0	0
L5	0	0	0	0	6	0	0	1	0
L6	0	0	1	0	0	29	0	0	0
L7	1	0	0	0	0	0	21	0	0
L8	0	0	1	0	1	0	0	22	0
L9	0	0	0	1	0	0	0	0	9

TABLE 8. The performance of subgroup models compared with full-group ANN Model according to Accuracy and F1-score.

Model	Method	Databases	Accuracy	F1-score
full_group ANN Model	ANN	full_group datasets	0.8645	0.8293
A_group ANN Model	ANN	A_group datasets	0.9683	0.9613
B_group ANN Model	ANN	B_group datasets	0.9657	0.8928
C_group ANN Model	ANN	C_group datasets	0.9791	0.9268
D_group ANN Model	ANN	D_group datasets	0.9149	0.9053

the information found in Fig.4, which shows the precision and recall of all the labels of the models. Table 8 shows that the full-group ANN model outperformed the traditional machine-learning models, and Fig.4 shows that the precision and recall of the full-group ANN model is more balanced for

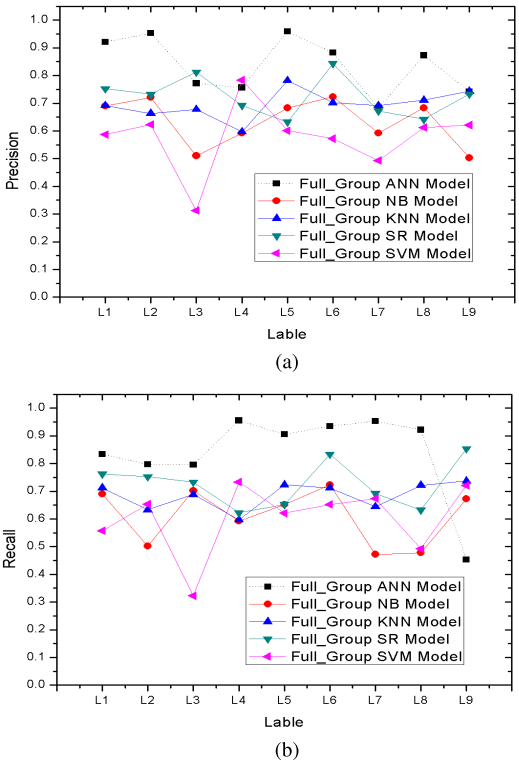


FIGURE 4. The precision and recall of the all labels of traditional machine learning models compared with full-group ANN model.

TABLE 9. The performance of traditional machine-learning models compared with full-group ANN model according to accuracy and F1-score.

Model	Method	Databases	Accuracy	F1-score
full-group ANN Model	ANN	full-group datasets	0.8645	0.8293
full-group SR Model	SR	full-group datasets	0.7521	0.7235
full-group NB Model	NB	full-group datasets	0.6782	0.6132
full-group KNN Model	KNN	full-group datasets	0.7748	0.6902
full-group SVM Model	SVM	full-group datasets	0.5864	0.5875

each label. The comparison experiment conclusively showed that ANN is superior to the traditional machine-learning algorithm in modeling the ISD of TCM.

V. CONCLUSIONS

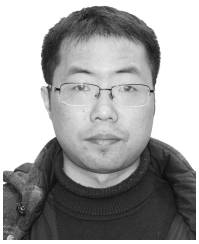
In this paper, ANN was used to model the ISD of TCM for the case study of COPD. We conducted the experiment with a dataset composed of 18,741 real clinical records and a network containing nine hidden layers, and the results showed that the ISD could be used by the middle level of TCM physicians. In addition, we built four subgroup models with ANN and four subgroup datasets, and the subgroup models achieved a better performance than the full-group model. For comparison, we modeled the syndrome differentiation by four traditional machine-learning algorithms to determine which ANN was best for modeling intelligent syndrome

differentiation. There continues to be significant work on the optimization of the model. The following are important for this work.

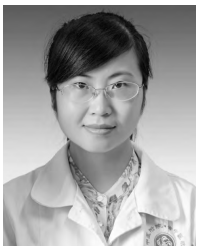
1) The size of the datasets is a critical factor for modeling. The model of ISD can always benefit from a larger dataset. It is believed that with the platform from which data are constantly uploaded, the performance of the model will be phenomenal.
2) The quality of data is the other important factor for modeling. More professional TCM practitioners are expected to manage the data cleaning for building public datasets.
3) In the future, we will be also be committed to improving the algorithm. With the increasing data we can augment the complexity of the network, and ensemble learning is an alternative strategy

REFERENCES

[1] D. Cyranoski, "The big push for chinese medicine for the first time, the world health organization will recognize traditional medicine in its influential global medical compendium," *Nature*, vol. 561, no. 7724, pp. 448–450, 2018.
[2] P. H. Y. Caldwell, L. Lai, X. Li, Y. Fei, H. Cao, J. Sun, N. Liang, J. Liu, and K. M. Scott, "Development and delivery of an evidence-based medicine programme for Traditional Chinese Medicine students using an international collaborative approach: An action research project," *Eur. J. Integr. Med.*, vol. 20, pp. 98–105, Jun. 2018.
[3] Y. Lu, W. Peng, J. Adams, and D. W. Sibbritt, "Treating people with arthritis with traditional chinese medicine (TCM): An examination of the perception of TCM practitioners," *Acupuncture Med.*, vol. 36, no. 4, pp. 228–239, Aug. 2018.
[4] Y. H. Yeh, Y. J. Chou, N. Huang, C. Pu, and P. Chou, "The trends of utilization in traditional Chinese medicine in taiwan from 2000 to 2010: A population-based study," *Medicine*, vol. 95, no. 27, Jul. 2016, Art. no. e4115.
[5] T. Watsuji, S. Arita, S. Shinohara, and T. Kitade, "Medical application of fuzzy theory to the diagnostic system of tongue inspection in traditional Chinese medicine," in *Proc. IEEE Int. Fuzzy Syst. Conf.*, vol. 1, Aug. 2002, pp. 145–148.
[6] R. Kanawong, T. Obafemi-Ajayi, J. Yu, D. Xu, S. Li, and Y. Duan, "Zheng classification in traditional chinese medicine based on modified specular-free tongue images," in *Proc. IEEE Int. Conf. Bioinf. Biomed. Workshops*, Oct. 2012, pp. 288–294.
[7] S. C. Hui, Y. He, and D. T. C. Thach, "Machine learning for tongue diagnosis," in *Proc. 6th Int. Conf. Inf., Commun. Signal Process.*, Dec. 2007, pp. 1–5.
[8] J. Yan, Q. Shen, Y. Wang, F. Li, R. Guo, C. Chen, and Y. Shen, "Detecting non-stationarity for auscultation signal of traditional Chinese medicine," *Wuhan Univ. J. Natural Sci.*, vol. 16, no. 1, pp. 83–87, Feb. 2011.
[9] C.-C. Chiu, M.-T. Yang, and C. S. Lin, "Using fractal dimension analysis on objective auscultation of traditional chinese medical diagnosis," *J. Med. Biol. Eng.*, vol. 22, no. 4, pp. 219–224, Dec. 2002.
[10] G. Z. Li, S. Sun, M. You, Y. L. Wang, and G.-P. Liu, "Inquiry diagnosis of coronary heart disease in Chinese medicine based on symptom-syndrome interactions," *Chin. Med.*, vol. 7, no. 1, p. 9, Dec. 2012.
[11] G.-P. Liu, G.-Z. Li, Y.-L. Wang, and Y.-Q. Wang, "Modelling of inquiry diagnosis for coronary heart disease in traditional Chinese medicine by using multi-label learning," *BMC Complementary Alternative Med.*, vol. 10, no. 1, p. 37, Dec. 2010.
[12] X. Liu, P. Lu, X. Zuo, Y. Gao, Y. Yang, and J. Chen, "A new method of modeling of inquiry diagnosis for coronary heart disease in traditional Chinese medicine," in *Proc. 4th Int. Conf. Biomed. Eng. Informat. (BMEI)*, vol. 3, Oct. 2011, pp. 1631–1634.
[13] H. Shao, G. Li, G. Liu, and Y. Wang, "Symptom selection for multi-label data of inquiry diagnosis in traditional Chinese medicine," *Sci. China Inf. Sci.*, vol. 56, no. 5, pp. 1–13, May 2013.
[14] G.-P. Liu, J.-J. Yan, Y.-Q. Wang, W. Zheng, T. Zhong, X. Lu, and P. Qian, "Deep learning based syndrome diagnosis of chronic gastritis," *Comput. Math. Methods Med.*, vol. 2014, Mar. 2014, Art. no. 938350.
[15] X.-B. Yang, Z.-H. Liang, G. Zhang, Y.-J. Luo, and J. Yin, "A classification algorithm for TCM syndromes based on P-SVM," in *Proc. Int. Conf. Mach. Learn. Cybern.*, vol. 6, Aug. 2005, pp. 3692–3697.
[16] T. Liu, C. Xia, Y. Wang, and J. Xu, "Classifying syndromes in traditional Chinese medicine based on ISOMAP-SVM," in *Proc. 5th Int. Conf. Biomed. Eng. Informat.*, Oct. 2012, pp. 464–468.
[17] W. Bo, Z. Ming-Wei, Z. Bin, W. Wei-Jie, "Data mining application to syndrome differentiation in traditional chinese medicine," in *Proc. 7th Int. Conf. Parallel Distrib. Comput., Appl. Technol.*, Dec. 2006, pp. 128–131.
[18] C. Zhao, G.-Z. Li, C. Wang, and J. Niu, "Advances in patient classification for traditional chinese medicine: A machine learning perspective," *Evidence-Based Complementary Alternative Med.*, vol. 2015, Apr. 2015, Art. no. 376716.
[19] L.-C. Lo, J. Y. Chiang, T.-L. Cheng, and P.-S. Shieh, "Visual agreement analyses of traditional chinese medicine: A multiple-dimensional scaling approach," *Evidence-Based Complementary Alternative Med.*, vol. 2012, Aug. 2012, Art. no. 516473.
[20] S. K. Poon, J. Poon, M. McGrane, X. Zhou, P. Kwan, R. Zhang, and D. M.-Y. Sze, "A novel approach in discovering significant interactions from TCM patient prescription data," *Int. J. Data Mining Bioinf.*, vol. 5, no. 4, pp. 353–368, Jan. 2011.
[21] A. Jarab, E. Alefishat, T. Mukattash, K. Alzoubi, and S. Pinto, "Patients' perspective of the impact of COPD on quality of life: A focus group study for patients with COPD," *Int. J. Clin. Pharmacy*, vol. 40, no. 3, pp. 573–579, Jun. 2018.
[22] S.-Y. Li, J.-S. Li, M.-H. Wang, Y. Xie, X.-Q. Yu, Z.-K. Sun, L.-J. Ma, W. Zhang, H.-L. Zhang, F. Cao, and Y.-C. Pan, "Effects of comprehensive therapy based on traditional Chinese medicine patterns in stable chronic obstructive pulmonary disease: A four-center, open-label, randomized, controlled study," *BMC Complementary Alternative Med.*, vol. 12, no. 1, p. 197, Dec. 2012.
[23] J.-S. Li, S.-Y. Li, Y. Xie, X.-Q. Yu, M.-H. Wang, Z.-K. Sun, L.-J. Ma, X.-H. Jia, H.-L. Zhang, J.-P. Xu, and C.-X. Hou, "The effective evaluation on symptoms and quality of life of chronic obstructive pulmonary disease patients treated by comprehensive therapy based on traditional Chinese medicine patterns," *Complementary Therapies Med.*, vol. 21, no. 6, pp. 595–602, Dec. 2013.
[24] G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks," *Science*, vol. 355, no. 6325, pp. 602–606, Feb. 2017.
[25] J.-D. Zhao, "Differential diagnosis of TCM symptoms," People's Med. Publishing House, Tech. Rep., 1994.
[26] J.-S. Li, S.-Y. Li, and X.-Q. Yu, "Guidelines for tcm diagnosis and treatment of chronic obstructive pulmonary disease," (in Chinese), *J. Traditional Chin. Med.*, vol. 1, no. 35, 2012.
[27] A. Cysneiros, C. Antunes, F. Paula, and F. Froes, "The ABCD of COPD. . . for patients," *Acta Med. Portuguesa*, vol. 31, no. 12, pp. 703–705, Dec. 2018.
[28] A. Fujino, H. Isozaki, and J. Suzuki, "Multi-label text categorization with model combination based on F_1-score maximization," in *Proc. 3rd Int. Joint Conf. Natural Lang. Process.*, vol. 2, 2008, pp. 1–6.
[29] C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and F-Score, with implication for evaluation," in *Proc. Eur. Conf. Inf. Retr.* Springer, Mar. 2005, pp. 345–359.



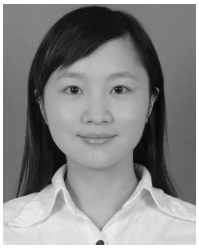
QIANG XU received the B.S., M.S., and Ph.D. degrees from the Chengdu University of Traditional Chinese Medicine, in 2012, 2014, and 2017, respectively, where he is currently a Postdoctoral Fellow. His current research interest focuses on the study of artificial intelligence of the TCM, which is supported by the National Natural Science Foundation of China and the China Postdoctoral Science Foundation.



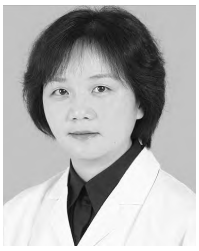
WENJUN TANG received the B.S. and M.S. degrees from the Chengdu University of Traditional Chinese Medicine, in 2005 and 2008, respectively, where she is currently a Resident Physician of TCM. She was involved in data cleaning in this study.

FEI TENG received the B.S. and M.S. degrees from Southwest Jiaotong University, China, in 2006 and 2008, respectively, and the Ph.D. degree from the Ecole Centrale Paris, France, in 2011. She is currently an Associate Professor with the School of Information Science and Technology, Southwest Jiaotong University. Her research interests include medical informatics, industrial big data mining, and cloud computing.

WEI PENG received the bachelor's degree from the School of Pharmacy, Chengdu University of Traditional Chinese Medicine, in 2006, where he is currently with the School of Pharmacy. He has published over 40 research papers (among these, over 30 as first or corresponding authors) with total citation of approximately 600. His current main interests are focused on the intelligent manufacturing and quality evaluation of TCM drugs preparation, intelligent diagnosis of diseases, and new strategies for candidate drugs screening.



YIFAN ZHANG received the B.S. and M.S. degrees from the Beijing University of Traditional Chinese Medicine, in 2010 and 2012, respectively, and the Ph.D. degree from the Chengdu University of Traditional Chinese Medicine, in 2017. She is currently a Resident Physician of TCM with the Department of Respiratory, Chengdu First People's Hospital, China. She was involved in data cleaning in this study.



WEIHONG LI has cultivated over 20 master's and Ph.D. students. She is currently a Professor and the Ph.D. Supervisor with the Chengdu University of Traditional Chinese Medicine, China. She conducted the final validation of the data in this study. She is a member of the Diagnostic Society of Traditional Chinese Medicine and the Geriatrics Committee of the World Federation of Traditional Chinese Medicine and the Academic Leader of the Sichuan Traditional Chinese Medicine Administration.

CHUANBIAO WEN is currently the Director of the Digital Medicine Research Institute and also with the Information and the Educational Technology Center, Chengdu University of Traditional Chinese Medicine, China. His current research interests include medical informatics and adverse herb-herb interaction.

JINHONG GUO received the bachelor's degree in electronic engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2010, and the Ph.D. degree in biomedical engineering from Nanyang Technological University, in 2014. After his doctoral studies, he was a Postdoctoral Fellow of the pillar of engineering design with MIT, SUTD, Singapore, from 2014 to 2015. He was a Visiting Professor with the School of Mechanical Engineering, University of Michigan, Ann Arbor, in 2016. He is currently a Full Professor with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, and with the Chengdu University of Traditional Chinese Medicine, Chengdu. He has published over 70 publications in top journals, such as the IEEE TII, TBME, TBioCAS, *Analytical Chemistry*, and *Biosensors and Bioelectronics*. His current research interest focuses on electrochemical sensors and lab-on-a-chip devices for point of care test toward clinical use. He was a recipient of the China Sichuan Thousand Talents Plan for Scholars Award, in 2015, and the Chengdu Expert in Science and Technology Award, in 2015. He served as a Lead Guest Editor and a Guest Editor for the IEEE TBioCAS, *Micromachines*, *Electrophoresis*, and *Sensors*.