Diagnosis in Traditional Chinese Medicine Using Artificial Neural Networks: State-of-the-art and Perspectives

Minghui Shi
Institute of Artificial Intelligence, Xiamen University, Xiamen, Fujian, 361005, China
smh@xmu.edu.cn

Changle Zhou
Institute of Artificial Intelligence, Xiamen University, Xiamen, Fujian, 361005, China
dozero@xmu.edu.cn

Abstract

Traditional Chinese Medicine (TCM), one of China's splendid cultural heritages, is the science dealing with human physiology, pathology, diagnosis, treatment and prevention of diseases. With the development of modern science, people come to consider the way of the modernization of TCM. Recently, many researchers mainly in China attempt to realize Diagnosis in TCM based on Artificial Neural Networks (DTCMANN). This paper aims at providing an overview of recent DTCMANN studies in TCM field, and focuses on the introduction and summarization of the existing research work about DTCMANN. A review of five major situations, where DTCMANN approaches are applied, has been presented. For each situation, the DTCMANN approaches employed are outlined, as well as the corresponding results. Current research on DTCMANN shows that it is both feasible and promising, and that it is still nearly a piece of virgin soil. The future research direction of DTCMANN is also pointed out based on a discussion of the existing research work.

1. Introduction

Traditional Chinese medicine (TCM) has a history of more than five thousand years. It has a complete theory about the occurrence, development and treatment of diseases, and plays an indispensable role in the health care for Chinese people. The characteristics of TCM can be summarized as the holistic concept and treatment by differentiation of syndromes, which are essentially different from that based on reductionism in western medicine. With the development of modern science, people come to consider the way of the modernization of TCM. In 1980s, for instance, many expert systems about TCM had been developed in China. Those expert systems, however, cannot show perfect performance since they have some inherit limitations such as lacking enough knowledge, no capability of learning and inefficient inference mechanism. So the research for expert systems about TCM at that time decreased due to the temporal technologies. Meanwhile, with the recent development of intelligent computing technologies such as fuzzy logic, neural networks, genetic algorithm, and rough sets theory, etc., it becomes to be possible to develop new technologies for intelligent diagnosis in TCM and go beyond the limitations.

Artificial neural networks (ANNs) have been widely and successfully employed in various theoretical research and practical applications. The capability to learn by example, generalization and parallel computing are the main characteristics of ANNs. Recently, many researchers in China attempt to realize diagnosis in TCM based on ANN (DTCMANN). Considering the rapidly increasing researches carried out on DTCMANN, it is necessary and helpful to provide an overview of recent approaches to DTCMANN. Since TCM is quite different from western medicine both in practice and in theory, it is necessary to take an insight into DTCMANN involving high domain-specificity of diagnostic technology. Motivated by these needs, this review paper focuses on the introduction and summarization of existing research work about DTCMANN, and explores its future research direction.

The rest of this paper is organized as follows. First, the standard feedforward neural network is briefly reviewed in section 2. Then, the general approach to DTDMANN is presented in section 3. Subsequently the recent DTCMANN research works in various situations of diagnosis in TCM are reviewed in section 4, and the future direction of DTCMANN is discussed in section 5. Finally, we conclude in section 6.

2. Standard feedforward neural network

In this section, we briefly review the standard feedforward neural network (FNN). The general structure of FNN has one input layer, several middle layers and one output layer. [1] establishes that an

Third International Conference on Natural Computation (ICNC 2007)
0-7695-2875-9/07 $25.00 © 2007 IEEE

three-layer FNN that has only one hidden layer, with a sufficient number of neurons, acts as an universal approximation of nonlinear mappings. Three-layer FNNs are widely used in practical applications.

Let us glance at the reasoning process (i.e., computing process) of a three-layer FNN and consider the three-layer FNN with n_I input units, n_H hidden units, and n_O output units (see Fig. 1).

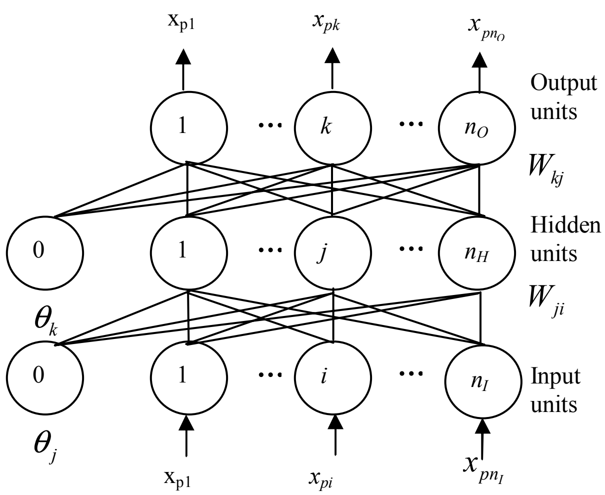

Fig. 1. Architecture of three-layer neural network

When an n_I-dimensional input vector $x_p = (x_{p1}, x_{p2}, \ldots, x_{pn_I})$ is presented to the neural network, the input-output relation of each unit can be written as follows [2]:

Input units:

$$o_{pi} = x_{pi}, \qquad i = 1, 2, \ldots, n_I. \tag{1}$$

Hidden units:

$$o_{pj} = f_1(net_{pj}), \qquad j = 1, 2, \ldots, n_H, \tag{2}$$

$$net_{pj} = \sum_{i=1}^{n_I} w_{ji} o_{pi} + \theta_j, \quad j = 1, 2, \ldots, n_H. \tag{3}$$

Output units:

$$o_{pk} = f_2(net_{pk}), \qquad k = 1, 2, \ldots, n_O, \tag{4}$$

$$net_{pk} = \sum_{j=1}^{n_H} w_{kj} o_{pj} + \theta_k, \quad k = 1, 2, \ldots, n_O. \tag{5}$$

Here w_{ji} and w_{kj} are connection weights, θ_j and θ_k are biases, and f_1 is the transfer function in the hidden layer, f_2 is the transfer function in the output layer. f_1 *and* f_2 are always sigmoid function.

In learning process, some samples are provided to the neural network. The cost function to be minimized for each given sample in learning process is:

$$e_p = \frac{1}{2}\sum_{k=1}^{n_o}(e_k)^2 = \frac{1}{2}\sum_{k=1}^{n_o}(t_{pk} - o_{pk})^2 \tag{6}$$

where, e_p, e_k, t_{pk} and o_{pk} denote training error for the sample, raining error, target output and actual output of the k-th output neuron, respectively.

In the well-known back-propagation (BP) learning algorithm [3], the weights w_{ji} , w_{kj} and biases θ_j , θ_k are updated to decrease the cost function. After training had been completed, the ANN, when given an input, would be expected to produce a corresponding output that accords with samples.

3. General approach to DTCMANN

The point of departure for DTCMANN is: the process of diagnosis in TCM is, in essence, a mapping that searches the corresponding diagnostic conclusions, which may be viewed as image, with respect to the symptoms suffered by patients, which may be viewed as original image. Since ANNs may be viewed as mapping tools [4], they may be applied to the implementation of diagnosis in TCM (see Fig. 2).

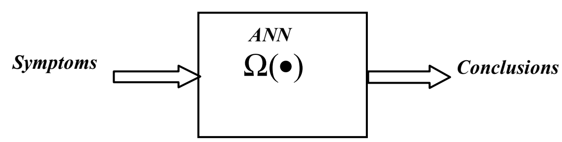

Fig. 2. ANN as a mapping tool in TCM

Recently, there have been some researches published on DTCMANN in China. Most of them assume that the space of symptoms is a two-value space, and use variants x_i (i=1,2,…,n) denote symptoms, so a group of symptoms can be denoted by a set { x_i | i=1,2,…,n } , So the set of symptoms of each patient corresponds to one vector X: (x_1, x_2, … , x_n), where x_i (i=1,2,…,n) is equal to 1 when the i-th symptom occurs, otherwise is equal to 0, and can be viewed as a point in a n-dimensional space. In a similar way, the set of all possible conclusions may be denoted by a set { y_i | i=1,2,…,m } , So the set of conclusions for each patient can be represented by one vector Y: (y_1, y_2, … , y_m), where y_i (i=1,2,…,m) is equal to 1 when the i-th conclusion holds, otherwise is equal to 0, and can be viewed as a point in a m-dimensional space. From this perspective, the model for DTCMANN can be viewed as a mapping from n-dimensional space to m-dimensional space.

ANNs currently employed in DTCMANN are usually three-layer FNNs using BP learning algorithm.

Third International Conference on Natural Computation (ICNC 2007)
0-7695-2875-9/07 $25.00 © 2007 IEEE

The number and meaning of units in each layer are determined as follows. Each unit in the input layer corresponds to each symptom, and each unit in the output layer corresponds to each possible conclusion. The number of units in the middle layer cannot be precisely determined by theory, and always can be determined by experiments.

The general approach to DTCMANN includes two processes: *training process* and *diagnosis process*, and can be summarized as follows. In the *training process*, there are always three steps: first, collect successful clinical cases. Then, transform each case into each pair "symptoms/conclusions", and quantified symptoms and conclusions to attain samples represented by "X/Y", where X and Y are n-dimensional vector and m-dimensional two-value vectors, respectively. Finally, train the ANN using some samples, and test it using the other samples.

There are four steps in the *diagnosis process*: first, collect symptoms of clinical patients. Second, quantify the symptoms collected, and attain an n-dimensional two-value vector X'. Third, taking X' as input, the trained ANN produces an output. Finally, draw a conclusion based on the output.

4. Reasearches in different situations of diagnosis in TCM

Based on the general approach described in section 3, DTCMANN has been applied to various situations of diagnosis in TCM. The current research published may be summarized as follows: discriminating several diseases, discriminating several diseases that belong to a common class, discriminating confusable syndromes, discriminating syndromes that belong to a common disease, discriminating diseases and syndromes that have similar symptoms, and estimating the phases of diseases. In this section, the methods as well as corresponding results will be outlined.

4.1. Discriminating from several diseases

In 1991, [5] pointed out that DTCMANN was feasible and demonstrated a successful experiment. They adopt 200 clinical cases provided by Beijing Xiyuan Hospital, which is attached to the institute of TCM research in China. These clinical cases include four diseases: *Chest Bi*, *Asthma*, *Wind Stroke* and *Metrorrhagia*. From these cases, [5] selected 202 typical symptoms. The symptoms in each case are denoted by an 202-dimensional vector X: $(x_1, x_2, \dots, x_{202})$, where x_i (i=1,2,…,202) is equal to 1 when the symptom corresponding to x_i occurs in the case, otherwise, is equal to 0. The clinical conclusion in each case is denoted by an target vector T=(t_1, t_2, t_3, t_4) , where t_i (i=1,2,3,4) is equal to 1 when the disease corresponding to t_i holds in the case, otherwise, is equal to 0. That is, there are four target patterns for these cases such as (1,0,0,0), (0,1,0,0), (0,0,1,0) and (0,0,0,1). Thus, 200 samples "X/T" are attained. Half of them (i.e., 100) act as training samples to train the ANN, and the rest are testing samples to test the trained ANN. The structure of the ANN adopted by [5] is as follows: 202 units in the input layer respectively corresponding to 202 symptoms, 202 units in the middle layer, and 4 units in the output layer respectively corresponding to 4 diseases. The learning algorithm of the ANN is BP algorithm. The decision strategy adopted by [5] is *maximality principle*, that is, the disease corresponding to the unit with the maximal value in the output layer is deemed to the final diagnostic conclusion. The diagnostic conclusion drawn by ANN may be compared with that in the corresponding testing case. The results of the experiment show that the accurate rate is 96%.

4.2. Discriminating several diseases that belong to a common class

In clinical diagnosis, it is always necessary to estimate the exact disease among several diseases that belong to a common class. For example, *chronic hepatitis*, *hepatocirrhosis* and *liver cancer* belong to *liver disease*. [6] attempted to use ANN to discriminate the three diseases. They collect fifteen typical cases, nine of which are used as testing samples, the rest of which are used as testing samples. From these cases, [6] select five main symptoms: *age* (x_1), *course of disease* (x_2), *edema in upper abdomen* (x_3), *liver sticking out* (x_4) and *swollen spleen* (x_5). The method of quantification for the symptoms is interval-method. For instance, *age* (x_1) is quantified as: 0 ($0<age\leqslant 20$), 1 ($20<age\leqslant 30$), 2 ($30<age\leqslant 40$), 3 ($40<age\leqslant 50$), 4 ($50<age\leqslant 60$), 5 ($age>60$). The structure of the ANN adopted by [6] is as follows: 5 units in the input layer respectively corresponding to 5 symptoms, 9 units in the middle layer, and 1 unit in the output layer. The learning algorithm of the ANN is BP algorithm. The decision strategy adopted by [6] is: 1) if the value of the unit in the output layer belongs to the interval (-0.5, 0.5), then it is *chronic hepatitis*; 2) if belongs to the interval (0.5, 1.5), then *hepatocirrhosis* 3) if belongs to the interval (1.5, 2.5), then *liver cancer* 4)

Third International Conference on Natural Computation (ICNC 2007)
0-7695-2875-9/07 $25.00 © 2007 IEEE

otherwise, no decision. The results of the experiment show that this approach is feasible. [6] points out, however, that the choice of cases has effect on the results of ANN, and that the performance of ANN could be improved by using more typical cases.

4.3. Discriminating confusable syndromes

Syndrome is a unique concept in TCM. It is a summarization about location, cause, nature, and degree of some disease. Some syndromes have similar symptoms and are difficult to be differentiated. [7] attempted to apply ANN to discriminating four confusable syndromes: *Wondering Bi Syndrome*, *Bi Syndrome Marked by Severe Pain*, *Bi Syndrome Marked by Localized Pain*, and *Bi Syndrome Due to Heat*. They employ 80 cases that have been used in *Computer System for Bi Syndrome based on the Ideal of Professor Sheng Guorong*. Half of them (i.e., 40) act as training samples, the rest testing samples. 155 symptoms in these cases are considered, and are quantified by using the same method as [5] just described. The structure of the ANN adopted by [7] is: 155 units in the input layer corresponding to 155 symptoms, 155 units in the middle layer, and 4 units in the output layer respectively corresponding to 4 confusable syndromes. The learning algorithm of the ANN is BP algorithm. The decision strategy adopted by [7] is *maximality principle*. The results of the experiment show that the accurate rate is 95% for testing samples, and 92.5% for testing samples.

4.4. Discriminating syndromes that belong to a common disease

In TCM diagnosis, it is often necessary to discriminate syndromes that belong to the same disease so as to grasp the nature of the disease. For instance, *depression disease* includes five typical syndromes in TCM: *Liver Melancholia and Stagnation of Liver Qi*, *Liver Melancholia and Deficiency of Speech Qi*, *Liver Melancholia and Phlegm-Ying*, *Deficiency of Speech Qi and Heart Qi*, and *Deficiency of Liver-Ying and Kidney-Ying*. [8] exploited ANN to discriminate the five syndromes. They employ 480 cases about patients suffering *depression disease*, among which there were 94 cases for *Liver Melancholia and Stagnation of Liver Qi*, 97 cases for *Liver Melancholia and Deficiency of Speech Qi*, 95 cases for *Liver Melancholia and Phlegm-Ying* , 94 cases for *Deficiency of Speech Qi and Heart Qi*, and 100 cases for *Deficiency of Liver-Ying and Kidney-Ying*. The 480 cases are all used to be training samples, and another 20 cases randomly selected act as testing samples. 75 symptoms in these cases were considered, and were quantified by using the same method as [5]. The structure of the ANN adopted by [8] is as follows: 75 units in the input layer corresponding to 75 symptoms, 80 units in the middle layer, and 5 units in the output layer respectively corresponding to 5 syndromes. The learning algorithm of the ANN is BP algorithm. The decision strategy adopted by [8] is *maximality principle*. The results on the testing samples are satisfactory. [8] believes that it is valuable to improve the convergence speed of ANN.

4.5. Estimating the phases of diseases

In clinical diagnosis, it is important to judge the phase of diseases. [9] developed an ANN system to predicate the phase of the syndrome called *Cold Intermingled with Heat Syndrome*, which belonged to *Rheumatoid Arthritis Disease*. They adopt 50 clinical cases provided by the TCM Hospital of Guangdong Province of China. All cases are used to be training samples, and 40 cases randomly selected from these cases are used to be testing samples. From these cases, 45 symptoms are selected, and the phases of *Cold Intermingled with Heat Syndrome* are divided into three phases called *first phase*, *second phase* and *third phase*, respectively. Each syndrome in the cases is quantified as follows: 0 (don't occur), 1 (slight degree), 2 (middle degree), 3 (heavy degree). Target pattern drawn from each case belongs to one of the three patterns: (1,0,0), (0,1,0), (0,0,1), which respectively corresponded to *first phase*, *second phase* and *third phase*. The structure of the ANN is as follows: 45 units in the input layer corresponding to 45 symptoms, 50 units in the middle layer, and 3 units in the output layer respectively corresponding to *first phase*, *second phase* and *third phase*. The learning algorithm of the ANN is BP algorithm. The decision strategy adopted by [9] is *maximality principle*. The results of the experiment show that the accurate rate is 98%.

5. Discussion and future direction

Researches in the last decades have showed that DTCMANN is feasible and can be applied into various situations in TCM diagnosis. Existing DTCMANN approaches are, however, cut-and-try,

Third International Conference on Natural Computation (ICNC 2007)
0-7695-2875-9/07 $25.00 © 2007 IEEE

and have not been applied to practical diagnosis. Existing work has also showed that DTCMANN has some inherit limitations such as slow convergence speed, request for large number of cases, and absence of explanation. To overcome the limitations, it may be a better research direction to combine ANN with other reasoning mechanisms. In fact, a few researchers have worked along the direction. For example, [10] believed that it is feasible for diagnosis in TCM to combine ANN with fuzzy control technology. [11] explored the method for diagnosis in TCM by combing ANN with symbolic reasoning. [12] examined the approach to diagnosis in TCM by combing ANN with case-based reasoning. [13] preprocessed the clinical cases, before applying them to ANN, by using a method based on rough set theory. These researches have shown that hybrid approaches in DTCMANN has become an important research direction.

6. Conclusion

It is significant to develop TCM by means of theory and technology of modern science. In the last decades, The research work about DTCMANN have involved in various situations of diagnosis in TCM, but mainly stay on the level of experiment, and have not been applied to practical systems. The structures of ANNs adopted in the current approaches to DTCMANN are always three-layer feedforward neural networks, and the learning algorithms used are always BP algorithm. In fact, other structures and learning algorithms for ANN should be considered. Furthermore, it also becomes obvious that combing ANN with other theories and technologies such as symbolic reasoning, fuzzy set theory, rough set theory, etc., will be a promising research direction for DTCMANN.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (Grant No. 60672018)，the National High Technology Research and Development Program of China (863 Program) (Grant No. 2006AA01Z129)，and the Innovation Foundation for Science and Technology of Xiamen University of China (Grant No. XDKJCX20063011).

References

[1] K. Hornik, M. Stinchcombe, H. White, "Multilayer feedforward networks are universal approximators", *Neural Networks* ,vol. 2, no. 5, 1989, pp. 359-366.

[2] H. Ishibuchi, K. Morioka, I. B. Turksen, "Learning by fuzzified neural networks", *International Journal of Approximate Reasoning* , vol. 13, 1995, pp. 327-358.

[3] D. E. Rumelhart, G. E. Hinton, R. J. Williams, "Learning representations by back-propagating errors", *Nature*, vol. 323 , 1986, pp. 533-536.

[4] F. M. Ham, I. Kostanic, *Principles of Neurocomputing for Science & Engineering*. McGraw-Hill Companies, Inc., 2001, pp. 96-97.

[5] W. L. Chen, C. N. Wang, R. J. Guo, "Model for neural network and its application to diagnosis in traditional Chinese medicine"，*Journal of Chinese Medicine*. Vol. 71, no. 2, 1991, pp. 111-113 (in Chinese).

[6] X. G. Wu, W. X. Lu, "An connectionist expert system for liver disease", *Chinese Journal of Biomedical Engineering* , vol. 15, no. 1, 1996 , pp. 62-66 (in Chinese).

[7] W. J. Lin, "Applying BP networks to classification of syndromes of Bi syndrome in TCM". *Journal of Fujian College of TCM* ,vol. 7, no.4, 1997 , pp. 41-43 (in Chinese).

[8] X. P. Fan, Z. Peng, S. Y. Yang, et al., "Study on depression classified system based on multilayer perceptrons artificial neural network", *Computer Engineering and Application*, vol. 13, 2004, pp. 205-208 (in Chinese).

[9] Z. J. Zhou, Z. Y. Mao, Z. Z. Deng,"Neural network approach to medical grading of rheumatoid arthritis", *Journal of Biomedical Engineering*, vol. 16, no. 4, 1999, pp. 379~482 (in Chinese).

[10] Y. J. Bai, H. B. Sheng, Q. G. Meng, et al., "Exploration of methods using neural networks for studying of syndromes in TCM", *Chinese Archives of Traditional Chinese Medicine* , vol. 22, no.12, 2004, pp. 2221-2223 (in Chinese).

[11] W. C. Wang, W. X. Lu, "ECG classification system based on symbolic neural network". *Chinese Journal of Biomedical Engineering*, vol. 15, no. 3, 1996, pp. 202-207 (in Chinese).

[12] W. D. Zhao, Z. H. Sheng, X. H. Du, "Case-based reasoning medical diagnosis based on the strength neural network", *Journal of Southeast University (Natural Science Edition)* , vol. 130, no. 3, 2000, pp. 46-50 (in Chinese).

[13] Z. G. Qin, Z. Y. Mao, "Rough neural network and its application in Chinese Medicine intelligent diagnose system", *Computer Engineering and Application* , vol. 18, 2001, pp. 34-35 (in Chinese).