Deep learning for acupuncture point selection patterns based on veteran doctor experience of Chinese medicine

Conference Paper · October 2012

DOI: 10.1109/BIBMW.2012.6470346

Zhaohui Liang
Guangzhou University of Chinese Medicine

Gang Zhang
GuangDong University of Technology

Jian Yin
Shandong University

Wen-bin Fu
Guangzhou University of Traditional Chinese Medicine

Deep Learning for Acupuncture Point Selection Patterns based on Veteran Doctor Experience of Chinese Medicine

Zhaohui Liang[1,*], Gang Zhang[2,3,*], Ziping Li[1], Jian Yin[2], Wenbin Fu[1,#]
[1] The 2nd Affiliated Hospital, Guangzhou University of Chinese Medicine, Guangzhou, China, 510120
[2] School of Computer Science and Technology, SUN-YAT SEN University, Guangzhou, China, 510275
[3] School of Automation, Guangdong University of Technology, Guangzhou, China, 510006

Corresponding author: Wenbin Fu. e-mail: fuwenbin@139.com

***Abstract*—the inheritance of clinical experience of veteran doctors of Chinese medicine (CM) plays a key role in the development and effectiveness enhancement of Chinese medicine in the history. The clinical experience are classified as the patterns of disease diagnosis and Chinese medical Zheng diagnosis, the identification of core elements of Zheng, the treatment experience and relation of herbal medicine formulae, Zheng and disease, and the common law of diagnosis and treatment in real practice. The source of the experience mainly originates from literature and manuscripts of CM masters, which are being electronically recorded during the last two decades. As a result, it makes feasible to apply data mining to the knowledge discovery through the experience of veteran CM doctors. However, the current focus on this field is limited to the published literature such as journal papers, conference proceedings and textbooks, but the paper based manuscripts personally written by the veteran doctors are usually neglected. In this paper, we established a database for Dr Situ Ling, who is a deceased famous CM acupuncture master in southern China. The study objective is to discover the acupuncture point selection patterns which require profession knowledge and experience from senior CM doctors. It is believed these patterns are deposited as underlying knowledge with various middle level concepts that can be analyzed and discover by a serial of algorithms. Thus in this work, we formularized the patterns of acupuncture point selection as a learning task with deep architecture, which attempts to capture either existent or underlying concepts so as to simulate the planning process of the combined diagnosis of western medicine and Chinese medicine. The Restricted Boltzmann Machines (RBM) was used as the main model for deep learning to process to medical record data with international standard diagnosis (ICD-10) previously made by trained doctors. Then the ICD-10 based diagnosis dataset was introduced into our framework to enhance the concepts diversity. After applying this model, the learning accuracy based on the medical record data base of Dr Situ Ling was raised up to 75%. Thus this model can serve as a solution to discover the acupuncture point selection patterns of CM acupuncture veteran doctors. Furthermore, the data mining study model linked by international diagnosis standard (i.e. ICD-10), point selection patterns, and clinical symptoms will provide useful cues to reveal the essence of Zheng diagnosis through experience of CM veteran doctors.**

Keywords-deep learning;veteran doctor of Chinese medicine; knowledge discover; acupuncture point patterns; ICD-10

I. Introduction

Chinese medicine (CM) is a branch of medical science with the history of thousands of years. Unlike the model of western medicine which is supported by rigorous multi-discipline scientific research and evidence based practice, the foundation of CM mainly originates from the classic canons or masterpieces of ancient China, and the long-term practice in clinic for more than two thousand years. The theories and clinical experience of CM are transmitted from the CM master to their apprentice. The clinical practice was inherited, developed, and respected as experience of veteran CM doctors. The distinguished examples of veteran CM doctor's experience are the Shanghan School which is the inheritance of CM master Zhang Zhongjing and his classic work *Shanghanlun*. And the instance of acupuncture is the inheritance chain from *the Canon of the Yellow Emperor to the Canon of Difficulties*, to the *Jia Yi Canon of Acupuncture*, and then to *the Encyclopedia of Acupuncture.* The CM apprentice doctors learn the medical knowledge and skills through reading the manuscripts of their masters, and improve the therapeutic techniques and methods by their practice. Thus the art of medicine is developed with the cycle of inheritance and development.

The contemporary study of CM also stresses the experience inheritance of veteran CM doctors. In China, the Ministry of Health (MOH), State Administration of Traditional Chinese Medicine (SATCM) initiated the work of inheritance of CM veteran doctors in 1990, and 3 lists of CM veteran doctors were respectively published in 1990, 1997 and 2003. The experience of CM masters whose names are on the list is to be inherited and study[1]. The clinical experience of CM veteran doctors can be summarized based on different themes, i.e. the experience of certain individuals, the experience of disease, the experience of certain herbal medicine formulae, and the experience of clinical decision making methodology [1]. There are several data sources for these studies, which include the classic work of CM, published papers in journals or conference proceedings, textbooks, state and international guidelines and standards. The supporting technologies are electronic data collection and data mining for knowledge discovery during the study.

In the study of literature, systematic review and many text mining methods were applied with satisfactory results. Chen et al introduced the design and development of data warehouse for data mining of veteran CM doctors experience on treatment of major infectious disease. The data were collected from relevant published papers and reports The information was categories to the disease model and treatment model. The disease model contains the disease diagnosis of western medicine and the Zheng diagnosis of CM, and the treatment model contains information of the treatment principles, herbal medicine formula and efficacy assessment. The analytic techniques are business objects (BO), online analytical processing (OLAP), Weka, Oracle Data Miner, SQL server, etc. [2]

For instance, Ma et al. used the Rosetta software to explore the potential linkage between Zheng (i.e.CM syndromes) diagnosis and prescription of herbal medicine formulae for patients with dyspnea. And they concluded that rough concept lattice clustering is effective to discover the hidden rules of CM medication formula and Zheng Diagnosis [3]. Wang et al. applied OLAP to study the types of Zheng diagnosis and corresponding herbal medicine formulae based on the electronic literature data base and text books. The result discovered the consensus on four Zheng types on influenza and their corresponding treatment herbal formulae [4].

The above studies only concentrated on the available electronic data such as the online literature, textbooks and CM standards and guidelines. However, one crude data source is neglected in these studies, i.e. the manuscripts of medical records by the CM veteran doctors. The manuscripts are first hand data source by the CM veteran doctors without the influence of search methodology. This pragmatic setting renders the unexpected cues to the core ideas of Zheng diagnosis and adjustment of the herbal formulae ingredients based on the clinical experience of CM veteran doctors.

In this paper we introduced a deep learning model to analyze the underlying relation of diagnosis of different diseases and acupuncture point selection. The strength of choosing acupuncture as the target CM treatment is that the point location has be standardized by World Health Organization (WHO). However, the diagnosis in the original manuscripts did not accord to a standard coding system. In order to solve this problem, we introduce the International Classification of Diseases 10 (ICD-10) [5] as standard diagnosis. The crude data of the manuscripts were input to an electronic data base (by Microsoft Access), and the western diagnosis was made by trained doctors based on the description of symptoms in the original records. The electronic data was used as the study dataset in this study.

As many research reported that CM doctors make their diagnosis relying on some obvious and underlying concepts [6-8]. The examples of obvious CM concepts are essence, vitality, spirit etc. But the underlying concepts are unable to be directly expressed. In our study data, the diagnosis of each records are transmitted to a confirmed code in ICD-10, which is considered as a concept originated from Western Medicine. Figure 1 illustrates a diagnosis procedure in a concept view.

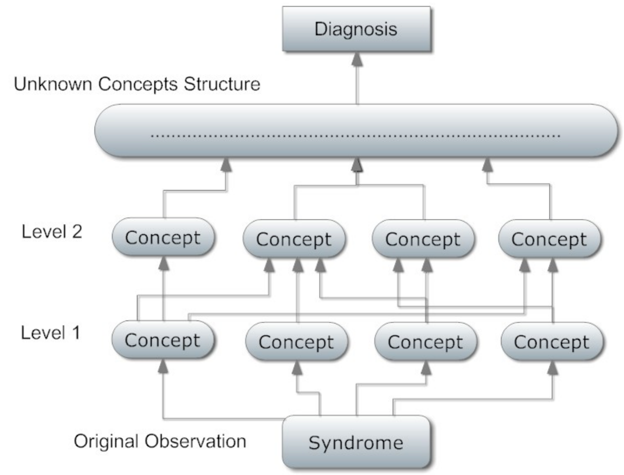

Figure 1. Concepts network for TCM diagnosis

It illustrates more than two concept levels in the diagnosis process and concepts from different levels from a network. Note that doctors have his own network based on his experience and knowledge structure, and different network structures have equal effects in most cases. The ICD-10 diagnosis standard can also be regarded as a certain concept of disease in western medicine.

We attempt to construct a therapy planning recommendation system, which extracts and learns knowledge in historic diagnosis data and outputs a therapy plan for unseen patients. In detail, the problem is confined in optimal acupuncture points selection for a certain disease. The motivation of this work is that we want to construct a learning model with concept levels homogenous to the thinking structure of doctors as illustrated in Figure 1. The objective of the work is to maximize the performance of such model through learning process with previous acupuncture point selection by the CM veteran doctor, Dr Situ Ling. Since we do not have enough information to fully describe the structure of concepts network, an artificial neural network (ANN) like model and a similar training procedure are applied to achieve the optimal solution. However, it can't be solved as an ANN since the concept levels may be much more than three, which causes the failure of BP training algorithm. Thus we switched to deep learning methods.

Deep learning is recently a hot issue in machine learning, which aims at constructing a deep architecture through training to simulate much more complex function than shallow architectures[9]. Motivated by observation of recognition theory, people succeed in recognizing complex objects and events through potentially constructing concepts in their brains. In order to express complicated functions and recognize complex objects and events, models with deep architectures are needed[10]. Deep Belief Network (DBN) [11], Restricted Boltzmann Machine (RBM) [12] and Stacked Autoassociators [13] are the three famous learning models in deep learning.

This paper is organized as following. Section II presents formal problem definition and Restricted Boltzmann

Machines (RBM) model for deep learning. Section III reports evaluation settings and results of the proposed model. And finally we conclude the paper in Section IV.

II. ALGORITHM

A. Problem definition

First of all, we give a formal definition of the problem. We consider the problem of acupuncture points selection patterns based on CM diagnosis for some target diseases basically originate from the dataset of Dr Situ Ling's medical record. During the diagnosis procedure of veteran CM doctor, the doctor inspected his patient and obtained some observation recorded in plain text, which founded the basic of CM diagnosis. To further improve the accuracy of diagnosis, the WHO ICD-10 standard is introduced in the dataset. The ICD-10 diagnosis was made by a trained doctor who reviewed the manuscripts of Dr Situ Ling and made diagnosis based on her professional knowledge. Figure 2 sketches the main framework of this work.

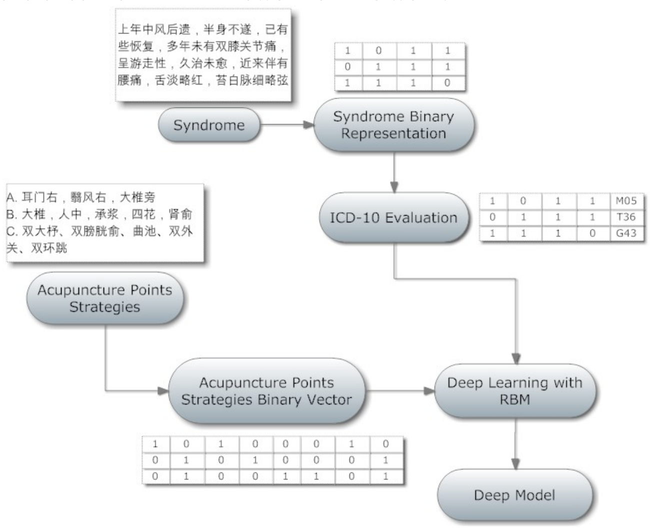

Figure 2. Main algorithm framework

Formally, let $D = \{(X_1, L_1, Y_1), \cdots, (X_n, L_n, Y_n)\}$ be a set of historic diagnosis records, in which $X_i \in \{0,1\}^d, (i = 1, \cdots, n)$ are binary syndrome descriptors indicating whether concerned syndromes occur in observation of a certain patient. $L_i \in R, i = 1, \cdots, n$ are labels of ICD-10 diagnostic codes. $Y_i \in \{0,1\}^d, i = 1, \cdots, n$ are binary vectors, each bit of which stands for whether a concerned acupuncture point is selected for treatment.

The goal of the model is to output a set of acupuncture points consistent with the CM veteran doctor's experience and expertise when the unseen CM Zheng diagnosis description is given.

B. Restricted Boltzmann Machines

Restricted Boltzmann Machines (RBM)[12] is a special type of Boltzmann Machine which is a particular type of energy-based model (EBM). RBM can be viewed as a block-wise optimization deep belief networks (DBN). Eq. (1) shows the energy function of RBM.

$$\begin{aligned} Eng(x,h) &= -b'x - c'h - h'Wx \\ &= -\beta(x) + \sum_i \gamma_i(x, h_i) \end{aligned} \tag{1}$$

where x are observed variables, h are hidden variables and W is weight matrix of arcs between x and h. Then we have joint distribution of x and h, as shown in Eq. (2).

$$P(x,h) = \frac{e^{-Eng(x,h)}}{Z} \tag{2}$$

where $Z = \sum_x e^{-Eng(x)}$, $Eng(x) = \sum_i f_i(x)$. f_i are expert functions. In RBM, f_i only take values $\{-1, +1\}$. According to[12], we further define free energy function as Eq. (3).

$$EngF(x) = -log \sum_h e^{-Eng(x,h)} \tag{3}$$

thus we can express the probability of x by $EngF$.

$$\begin{aligned} P(x) &= \frac{1}{Z} e^{-EngF(x)} = \frac{1}{Z} \sum_h e^{-Eng(x,h)} \\ &= \frac{1}{Z} \sum_{h_1} \sum_{h_2} \cdots \sum_{h_k} e^{\beta(x) - \sum_i \gamma_i(x,h_i)} \\ &= \frac{1}{Z} \sum_{h_1} \sum_{h_2} \cdots \sum_{h_k} e^{\beta(x)} \prod_i e^{-\gamma_i(x,h_i)} \\ &= \frac{e^{\beta(x)}}{Z} \sum_{h_1} e^{\gamma_1(x,h_1)} \sum_{h_2} e^{\gamma_2(x,h_2)} \cdots \sum_{h_k} e^{\gamma_k(x,h_k)} \\ &= \frac{e^{\beta(x)}}{Z} \prod_i \sum_{h_i} e^{-\gamma_i(x,h_i)} \end{aligned} \tag{4}$$

Eq. (4) setups the connection between probability density of x and the energy function defined in Eq. (3) through hidden variables h_i. In our work, hidden variables or their combination reflect underlying concepts during diagnosis by experienced veteran CM doctors. To further disclose the relationship between observation x and hidden concepts h, we can express the conditional probability form of Eq. (4) by applying $\beta(x) = b'x$ and $\gamma_i(x, h_i) = h_i W_i x$.

$$p(h_i = 1|x) = \frac{e^{c_i + W_i x}}{1 + e^{c_i + W_i x}} = sigm(c_i + W_i x) \tag{5}$$

$$p(x_j = 1|h) = \frac{e^{b_j + W'_j h}}{1 + e^{b_j + W'_j h}} = sigm(b_j + W'_j h) \tag{6}$$

An RBM update procedure has been proposed in[9]. To make the paper self-contained, we summarize the RBM update procedure. The main idea is to construct log-likelihood gradient of $p(x)$ with respect to W, c_i and b_i.

$$-\frac{\partial \log p(x)}{\partial W_{ij}} = E_x[p(h_i|x) \cdot x_j] - x_j^{(i)} \cdot sigm(W_i \cdot x^{(i)} + c_i) \tag{7}$$

$$-\frac{\partial \log p(x)}{\partial c_i} = E_x[p(h_i|x)] - x_j^{(i)} \cdot sigm(W_i \cdot x^{(i)}) \quad (8)$$

$$-\frac{\partial \log p(x)}{\partial b_i} = E_x[p(x_j|h)] - x^{(i)} \quad (9)$$

where W_i is the ith row of W, and $x^{(k)}$ stands for values of x in the kth loops of update procedure. Based on the principle of maximal likelihood optimization, a gradient descendant update procedure has been proposed in [9] to find the optimal W, b and c. Gibbs sampling method[14-15] is adopted to run a Markov chain to convergence when sampling $p(x)$. Eq. (10) and (11) describes the updating process.

$$h^{(k+1)} \leftarrow sigm(W'x^{(k)} + c) \quad (10)$$

$$x^{(k+1)} \leftarrow sigm(W'h^{(k+1)} + b) \quad (11)$$

where $h^{(k)}$ stands for values of all hidden codes in the kth iteration, and $x^{(k)}$ stands for input values of the kth iteration. Figure 3 graphically illustrates the relationship between input x and hidden variables h during Gibbs sampling.

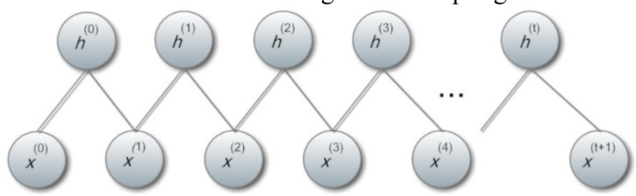

Figure 3. Relationship between x and h during Gibbs sampling

Finally we conclude the RBM update procedure as the end of this section, as Algorithm 1 describes.

Algorithm 1 RBM update

Require:

$x^{(1)}$: sampling from training dataset

eps: learning rate for stochastic gradient descent

W: an p by q weight matrix (between inputs and hidden units), where p and q stand for number of inputs and hidden units

b: bias vector for hidden units c: bias vector for inputs

Ensure:

updated W, b and c

for all hidden units h_i **do**
  $Q(h_i^{(1)} = 1|x_1) = sigm(b_i + \sum_j W_{ij}x_{1j})$
  sample $h_i^{(1)}$ from Q
end for
for all inputs x_i **do**
  $P(x_j^{(2)} = 1|h_1) = sigm(c_j + \sum_i W_{ij}h_{1i})$
  sample $x_j^{(2)}$ from P
end for
for all hidden units h_i **do**
  $Q(h_i^{(2)} = 1|x_2) = sigm(b_i + \sum_j W_{ij}x_{2j})$
end for
$W + e(h_1x_1' - Q(h_2 = 1|x_2)x'2) \rightarrow W$
$b + e(h_1 - Q(h_2 = 1|x_2)) \rightarrow b$
$c + e(x_1 - x_2) \rightarrow c$

In Algorithm 1 there are 3 for loops, in which the first and the third uses h to evaluate x with maximal likelihood meaning, while the second makes use of x in evaluating h. In these loops, when a stop criterion is met, e.g. max number of iteration, or difference between two iteration is smaller than a preset threshold, the algorithm stops. The main idea of Algorithm 1 is that sampling with distribution P, Q and updating parameters with newly calculated distribution take place by turn, until a stop criteria is met [16].

Initial values of x are directly from training dataset, and values of h are randomly initialized to $\{0,1\}$. Weights matrix W is also randomly initialized to $\{0,1\}$. Some previous work reported that combined with some prior knowledge in initialization of h and W yields better convergence than random initialization, which improves the effect of the target model. For simplicity and well-focused on our task, we don't introduce such methods in this work.

III. EVALUATION

A. Dataset

We evaluate the proposed algorithm framework in a real dataset originated from the medical record manuscripts of CM veteran doctor, Dr Situ Ling . Table I lists the names of the diseases as well as corresponding numbers of samples considered in this work.

TABLE I. DATASET DESCRIPTION

Name	Samples	Name	Samples
Arthralgia Syndrome	72	Acne	14
Epilepsy	7	Tinnitus & deafness	13
Abdominal pain	17	Allergic rhinitis	57
Neck & shoulder pain	19	Cervical spondylosis	8
Cough	17	Facial paralysis	16
Traumatic brain injury	10	Migraine	8
Ankylosing Spondylitis	8	Insomnia	10
Headache	24	Flaccidity Syndrome	21
Stomachache	24	Asthma	54
Palpitation	8	Lumbocrural pain	60
Urticaria & Rubella	7		

For training the dataset, the names of CM Zheng and acupuncture points are reformed from plain text to binary vectors description by key words matching, as shown in Figure 2. We manually select 76 key words for the names of Zheng and 51 names of acupuncture points to construct binary vectors description. Thus we have 76-ary Zheng binary vectors and 51-ary acupuncture points binary vectors, in which bit of value 1 stands for the corresponding Zheng or acupuncture point exists and 0 otherwise.

A ICD-10 label of the disease name of western medicine is also attached to each sample as a complement to CM Zheng diagnosis. In our training dataset, a certain disease can be labeled with different ICD-10 labels. Table II illustrates 51 different attached ICD-10 codes of diseases concerned in this paper.

TABLE II. ICD-10 LABELS OF CENCERNED DISEASES

Arthralgia Syndrome	M05, M06, M13, M17, M19, M48, M54, M54.3, M54.5, M99, M79.7
Acne	L70.0
Epilepsy	G40
Tinnitus & deafness	H65, H81.0, S02.1, H90, T36
Abdominal pain	K29, R10.0, K91.8, N34.1
Allergic rhinitis	J30
Neck & shoulder pain	M75.5, M54.2, M75.6, M75.9
Cervical spondylosis	M54.2
Cough	J40, J00, J42, J43.9
Facial paralysis	G51.0
Traumatic brain injury	S06
Migraine	G43
Ankylosing Spondylitis	M45
Insomnia	G47
Headache	G44
Flaccidity Syndrome	G83
Stomachache	K31.8, R10.0, K29
Asthma	J45
Palpitation	I47
Lumbocrural pain	M54.3, M19, M87, M77, M99, M54.4, M54.5, M17
Urticaria & Rubella	L50, B06

B. Settings

The proposed algorithm is evaluated with a matching rate between the output of the model and ground truth acupuncture points therapeutic plan. Accuracy is defined to evaluate the performance of the model as Eq. (12).

$$Loss(y, y') = \frac{H(y, y')}{|y|} \tag{12}$$

In Eq. (12), y and y' stand for the output of the model and the true Acupuncture Points therapeutic plan in test dataset. Function H calculates the Hamming distance between y and y', which evaluates the number of different bits of y and y'. $|y|$ stands for the length of binary vector y.

The dataset contains 474 records totally, each of which is attached with a type of disease. We want to show empirically that how much training data is sufficient to obtain a good model. The evaluation is performed with a ratio p between training and testing data, ranging from 10% to 40% with a step 15%. For diseases that only have no more than 15 samples, we apply Leave-One-Out (LOO) test for such cases. For simplicity, we only consider bi-classification in one-against-rest manner.

C. Evaluation results

Table III lists the evaluation result of the proposed method with settings mentioned in the above subsection.

In Table III evaluation accuracy is determined by $1 - Loss(y, y')$. We observed that the average accuracy is over 75% which is clinically acceptable. To further show the effectiveness of the proposed model with respect to deep learning, we implement artificial neural network (ANN), a traditional shallow architecture learning model as comparison. Five types of diseases are manually selected for comparison.

TABLE III. EVALUATION RESULT

Name	**10%**	**25%**	**40%**	**LOO**
Arthralgia Syndrome	72.8%	75.7%	77.0%	
Acne	78.4%	71.1%	76.0%	Yes
Epilepsy	80.6%	79.8%	79.3%	Yes
Tinnitus & deafness	74.3%	75.8%	82.4%	Yes
Abdominal pain	79.8%	84.0%	77.1%	
Allergic rhinitis	78.4%	75.8%	77.2%	
Neck & shoulder pain	70.1%	78.9%	78.3%	
Cervical spondylosis	77.5%	72.2%	79.2%	Yes
Cough	74.8%	75.5%	74.8%	
Facial paralysis	81.5%	81.7%	81.9%	
Traumatic brain injury	70.0%	80.8%	81.2%	Yes
Migraine	75.8%	82.4%	85.9%	Yes
Ankylosing Spondylitis	75.3%	75.0%	78.7%	Yes
Insomnia	75.8%	79.8%	79.9%	Yes
Headache	75.3%	74.2%	74.1%	
Flaccidity Syndrome	75.8%	77.6%	78.9%	
Stomachache	79.6%	81.9%	82.2%	
Asthma	74.0%	78.5%	79.5%	
Palpitation	79.8%	80.2%	84.2%	Yes
Lumbocrural pain	75.9%	76.1%	78.1%	
Urticaria & Rubella	70.5%	74.8%	80.4%	Yes
Avg.	76%	76.9%	78.5%	

Finally, to show the positive effect of considering the ICD-10 diagnostic information during the learning procedure, we trained a RBM without feeding ICD-10 information and compare to the proposed model. The average accuracy are less than 72% with the same settings as those of RBM with ICD-10 information, which shows the effectiveness of the proposed algorithm.

IV. CONCLUSION

We introduced a learning model with deep architecture to optimize the acupuncture points selection as the stimulation of the clinical thinking of the CM master Dr Situ Ling. Analogous to the thinking structure of human being, the proposed model well obeys the principles of deep learning and the evaluation result is competitive. Meanwhile, we have introduced ICD-10 standard diagnosis of western medicine as a potential concept in our model, which has been proved effective empirically. We applied Restricted Boltzmann Machines to train the deep network which is computational acceptable for a small to middle scale dataset. Since the diagnosis of CM Zheng implies many known and unknown concepts as factors, models with deep architectures may be more robust than shallow ones. Thus we believe it provides a new solution to explore the underlying patterns and rules which are deposited in the clinical experience of CM veteran doctors. Future work includes adapting current deep models to explore Zheng diagnosis and herbal formulae combination patterns based on records of CM veteran doctors experience.

ACKNOWLEDGMENT

Zhaohui Liang and Gang Zhang contributed equally in this study. This work is supported by National Natural Science Foundation of China (No. 81173348 & 81274003), the National Basic Research Program of China (973 Program)

(No. 2010CB530500 & 2010CB530503), the Open Research Fund of Zhejiang First-foremost Key Subject-Acupuncture & Moxibustion (No. ZTK2010B04), Research grant of Guangdong Administration of Chinese Medicine (No. 2010144), Creative Foundation of Guangzhou University of Chinese Medicine (No.10CX24), and the 2012 College Student Career and Innovation Training Plan Project (No. 16).

REFERENCES

[1] Y.H. Wang, Z.C. Jiang, Y.J. Yan, B.Y. Liu, J.G. Zhu, L. Tian, R.L. Gao, P. Li and Y.Z. Jiao, "An idea about study of clinical experience of famous veteran doctors in traditional Chinese Medicine on the baisis of information and data mining technology," World Science and Technology / Modernization of Traditional Chinese Medicine and Materia Medica, vol. 7, Jan. 2005, pp. 98-105.

[2] J.W. Chen, J. Zhang, D.W. Wu, Y. Luo and J.Q. Li, "Development of experience data warehouse for well-known veteran TCM physicians in the treatment of major infectious diseases and construction its intelligent mining platform," Chinese Journal of Medical Library and Information Science, vol. 4, Apr. 2011, pp. 42-47.

[3] H.X. Ma, Z.Y. Lu, B.X. Liu, D.G Zhao, G.Q Wang, R.X. Yu, W, Wang and J.T. Bao, "Rosetta application on traditional Chinese medicine dyspnea syndrome with experience of medication," Chinese Journal of Experimental Traditional Medical Formulae, vol. 18, Jul. 2012, pp. 5-9.

[4] D.W. Wang, Z.T. Zhou and Yi Luo, "Data mining on the clinical experience of Zheng diagnosis and treatment for influenza of contemporary CM veteran doctors," Shenzhen Journal of Integrated Traditional Chinese and Western Medicine, vol. 21, Jun. 2011, pp. 154–156.

[5] WHO, "International Classification of Diseases (ICD)," available on: http://www.who.int/classifications/icd/en/.

[6] R.S Zhang, Y.H. Wang, X.Z. Zhou, B.Y. Liu, N.L. Yao and P. Li, "The funtional design of the essential elements of famous CM veteran doctors and intelegent data mining platform," World Science and Technology / Modernization of Traditional Chinese Medicine and Materia Medica, vol. 10, Jan. 2008, pp. 45-63.

[7] H. Hu, Z.D. Huang, X.J. Xie, J. L and L.F. Wang, "The application of data mining technology to the summary of famous CM veteran doctors," Liaoning Journal of Traditional Chinese Medicine, vol. 37, Nov 2010, pp. 2144-2147.

[8] Z.P. Tan, Y. Luo, J.Q. Li and D.W. Zheng, "Data mining on the Zheng diagnosis and treatment experience for viral hepatitis of famous CM veteran doctors, " Liaoning Journal of Traditional Chinese Medicine, vol.38, Jul. 2011, pp. 1276-1277.

[9] Y. Bengio, "Learning deep architectures for ai," Found. Trends Mach. Learn., vol. 2, no. 1, pp. 1–127, 2009.

[10] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in Proceedings of the 26th Annual International Conference on Machine Learning, ser. ICML '09. New York, NY, USA: ACM, 2009, pp. 609–616.

[11] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," Neural Comput., vol. 18, no. 7, pp. 1527–1554, Jul. 2006.

[12] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," Neural Comput., vol. 14, no. 8, pp. 1771–1800, Aug. 2002.

[13] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in Advances in Neural Information Processing Systems 19, B. Sch¨olkopf, J. Platt, and T. Hoffman, Eds. Cambridge, MA: MIT Press, 2007, pp. 153–160.

[14] J. Weston, F. Ratle, and R. Collobert, "Deep learning via semisupervised embedding," in Proceedings of the 25th international conference on Machine learning, ser. ICML '08. New York, NY, USA: ACM, 2008, pp. 1168–1175.

[15] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Unsupervised learning of hierarchical representations with convolutional deep belief networks," Commun. ACM, vol. 54, no. 10, pp. 95–103, Oct. 2011.

[16] D. Erhan, Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent, and S. Bengio, "Why does unsupervised pre-training help deep learning?" J. Mach. Learn. Res., vol. 11, pp. 625–660, Mar. 2010.