RESEARCH ARTICLE

Advanced Science Letters
Vol. 23, 11879–11883, 2017

Classification of Lung Disease Syndromes in Traditional Chinese Medicine Based on Learning Vector Quantization

I. G. P. Asto Buditjahjanto[1,*], Naim Rochmawati[1], and R. Hapsari Peni[2]

[1]*Informatics Engineering Department, Universitas Negeri Surabaya, Kampus Ketintang, Jalan Ketintang, 60231, Surabaya, Indonesia*
[2]*Electrical Engineering Department, Universitas Negeri Surabaya, Kampus Ketintang, Jalan Ketintang, 60231, Surabaya, Indonesia*

In Indonesia, the development of treatment methods of Traditional Chinese Medicine (TCM) has been growing rapidly. This is indicated that TCM as an alternative treatment other than conventional medicine to cure the disease. The problem arises when the students or people who want to learn TCM experiencing difficulty in determining the classification of syndrome. While conventionally, the determination of the classification of syndrome requires sufficient experience for the students or people that begin to learn TCM in order to be able to determine the classification of syndrome. The purpose of this research is to create software application that can help make a decision in determining the classification of lung disease syndrome in TCM. This software application can be used as a learning medium for students and people who learn about the classification of lung disease syndrome in TCM. This research uses Learning Vector Quantization (LVQ) method. As it is known, LVQ has superiority in classifying a data set into several clusters according to weight training on LVQ. The classification of lung syndrome-based LVQ is built by using input as many as 46 symptoms of lung syndrome, and the output as many as 5 types of lung syndrome. The simulation results show that LVQ can be used to classify the type of lung syndrome very well.

Keywords: Learning Vector Quantization (LVQ), Database, Decision Support System, Traditional Chinese Medicine (TCM).

1. INTRODUCTION

Currently, the development of treatment methods of Traditional Chinese Medicine (TCM) has been growing rapidly in Indonesia. This is indicated by large number of people use TCM as an alternative treatment beside conventional medicine to cure the disease. The problem arises when the students or people who want to learn TCM experiencing difficulty in determining the classification of syndrome. To diagnose a syndrome, it is needed some patients' physical status. They collected through four types of diagnostic methods such as: inspection, auscultation and olfaction, inquiry, and palpation. Then, a clinical TCM practitioner analyze and synthesize the information, symptoms, and patients' physical status to differentiate a disease.[1] The classification of the disease syndrome in TCM is one way to determine what action should be carried out in the treatment of patients. Incorrect in classifying a disease syndrome, causing the patient's recovery will be hampered.

While the determination of classification syndrome traditionally, it requires sufficient experience for people that learn TCM to be able to determine a disease syndrome. This research uses Learning Vector Quantization (LVQ) method. As it is known, LVQ has superiority in classifying a data set into several clusters according to weight training on LVQ.[2,3] LVQ has been widely used in helping human needs such as for classification textual document making it easier for searching for documents,[4] to classify dimensional hyperspectral imagery[5] and in the industrial world one of which is used to diagnosis of bottle filling plant on real time.[6] Other fields that often use LVQ is the field of medicine. LVQ can assist to classify of mass spectrometry data of mass cancer,[7] to classify of medical image such as MRI brain image to determine the presence of tumors,[8] to help find abnormalities in the urinary system,[9] to help find the presence of gallbladder stone,[10] to determine the presence of breast cancer,[11] to assist in classification of EEG[12] and also to classify of type of meningioas.[13] Indonesia as a tropical country has high humidity. This condition will affect the population's health condition. The tendency of the population to get the lung disease syndrome is great. Therefore this study is only restricted to learn about classification of lung disease syndrome. In this research the symptoms and the syndromes of lung is based on Cheng.[14] The number of types of lung disease syndrome is 5 types and the

*Author to whom correspondence should be addressed.

numbers of symptoms are 46 pieces. The purpose of this research is to create software applications that can help make a decision in determining the classification of lung disease syndrome in TCM. These software applications can be used as a learning medium for students and people who learn about the classification of lung disease syndrome in TCM. This research builds LVQ using a number of data input and output. Data input comes from the preferences of several experts (doctors) toward the symptoms of the patient. While the output data derived from the preferences of the experts toward the syndrome that appear based on these symptoms. Therefore these data are used to train LVQ to get optimum weight so that it can classify lung disease syndrome. Testing is conducted to determine the accuracy and precision in determining lung disease syndrome. Testing is carried out using data of patients' status from the clinic.

2. LEARNING VECTOR QUANTIZATION (LVQ)

This research uses LVQ as a method of classification to determine the type of lung syndrome. The parameters of LVQ are as follow: T is Target; J is the amount of the difference between data and weights; C is the smallest difference in weight class; W is Weight; α is Learning rate; X is input vector.

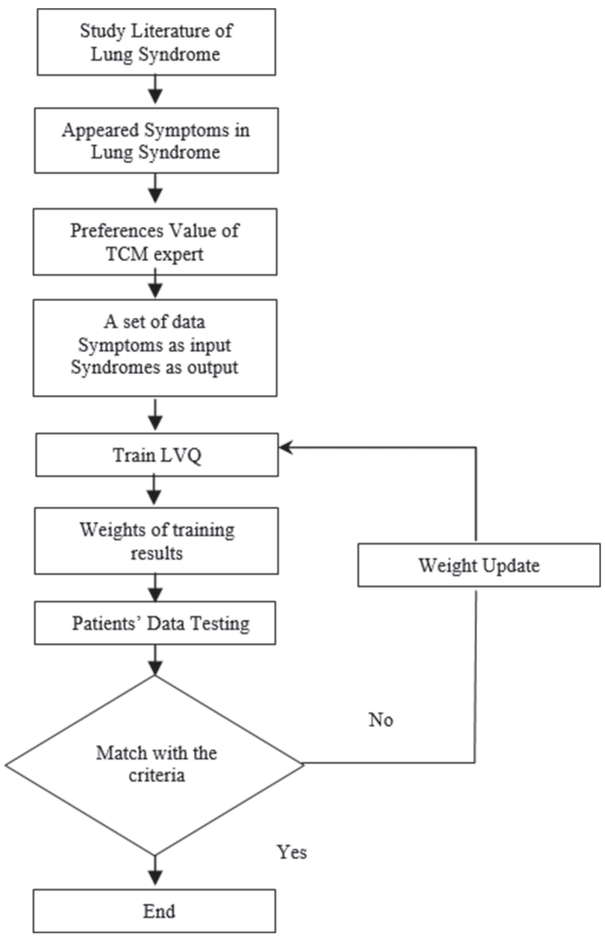

Fig. 1. Diagram block of research for lung syndrome classification.

Table I. Clinical manifestation for lung syndrome type invasion of the lungs by wind cold.

Lung syndrome type	Invasion of the lungs by wind cold	Weight preference (0-1)
Clinical manifestation	Severe chills	(……)
	Slight fever	(……)
	Headache	(……)
	General aching	(……)
	Absence of sweat	(……)
	Nasal discharge	(……)
	Cough with clear thin sputum	(……)
	White tongue coating	(……)
	Superficial and tense pulse	(……)

The LVQ algorithm is as follow:[15]

1. Initiate the reference vector and the learning rate (α).
2. As long as the stop condition is false, do steps 3 and 4. Condition stop if the loop reaches the maximum iteration and change the ratio is smaller than the minimum error.
3. For each training input vectors, do as follow:
   a. Find J so that $\|x - W_j\| \|x - W_j\|$ valued minimum.

$$C_j = \sqrt{\sum_{i=1}^{n} (x_i - W_{ij})^2}$$

   b. Update W_j with:
      i. If $T = C_j$ then $W_j(\text{new}) = W_j(\text{old}) + \alpha[X - W_j(\text{old})]$
      ii. If $T \neq C_j$ then $W_j(\text{new}) = W_j(\text{old}) - \alpha[X - W_j(\text{old})]$.
4. Decrease learning rate (α).
5. Test stopping condition: This could be the learning rate reaching a sufficiently small value or a fixed number of iterations.

Figure 1 shows a flowchart of research that using LVQ in order to assist in the classification of lung syndrome types. The initial stage is studying the literature related to lung syndrome. The next stage is to determine symptoms that appear in lung syndrome. The next step, it determines the weight of each symptom from lung syndromes. This research cooperates with three experts from the Academy of Acupuncture Surabaya (AAS) to determine weight of each symptom based on their preferences. Each expert gives their preferences toward to symptoms as the input that correlate with the type of lung syndromes as output.

Type of lung syndrome in this research consists of five types such as: invasion of the lungs by wind cold, the accumulation of heat in the lungs, phlegm retention of damp in the lungs, lung qi deficiency and lung yin deficiency. Table I shows an example of an instrument to take the preference of the expert toward the symptoms one of type of lung syndrome. Furthermore, the preference of the expert becomes weights to train LVQ. By setting parameter such as learning rate, number of decrease learning, minimum number of learning rate, maximum number of epoch and training LVQ. Afterwards, we train the LVQ until the optimal weight of LVQ is achieved. Testing is the next stage, at this stage, the optimal weights obtained in the preceding stage are tested. Testing is done by entering symptoms of patient data. If the test results show the same results with the syndrome the experts' preference then the weights can be used. But, if the test results do not match with the experts' preference then it required training weight back again.

Table II. A set of symptoms in lung syndromes criteria.

Code	Symptom
C01	Severe chills
C02	Slight fever
C03	Headache
C04	General aching
C05	Absence of sweat
C06	Nasal discharge
C07	Cough with clear thin sputum
C08	White tongue coating
C09	Superficial and tense pulse
C10	High fever
C11	Slight chills
C12	Sweating
C13	Thirst
C14	Dry nose
C15	Yellow nasal discharge
C16	Epistaxis
C17	Sore throat
C18	Cough with yellow and sticky sputum
C19	Constipation
C20	Deep yellow urine
C21	A red tongue with yellow coating
C22	Superficial and rapid pulse
C23	Cough with shortness of breath
C24	Fullness and stuffiness in the chest
C25	Orthopnoea in severe cases
C26	Cough with white and sticky sputum
C27	Sticky tongue coating
C28	A rolling pulse
C29	Feeble cough
C30	Shortness of breath
C31	Dislike of speaking
C32	Weak voice
C33	Pale complexion
C34	General lassitude
C35	Spontaneous sweating
C36	A pale tongue
C37	A thready pulse
C38	Dry cough without sputum
C39	Dry throat
C40	Hoarse voice
C41	Emaciation
C42	Feverish sensation in the palms and soles
C43	Tidal fever
C44	Flushed cheeks
C45	A red tongue with lack of fluid
C46	A thready, rapid pulse

Table II shows the symptoms that exist in the five main types of lung syndrome. The numbers of symptoms are as many as 46 kinds of symptoms. The symptoms become input for LVQ were further tested with weights that have been trained to give a decision in the classification of lung syndrome type.

3. RESULTS AND DISCUSSION

LVQ parameters in this study are as follow: Learning rate = 0.1, Number of Decrease learning rate = 0.1, Minimum number of learning rate = 0.001, Maximum number of epoch = 1000. The simulation program uses matlab programming. Figure 2 shows LVQ network for this research. The testing is done by using the optimal weights, where the simulation results show accurate results. It can be shown by entering symptoms according to the expert preferences, the simulation result shows a syndrome classification in accordance with the preferences expert.

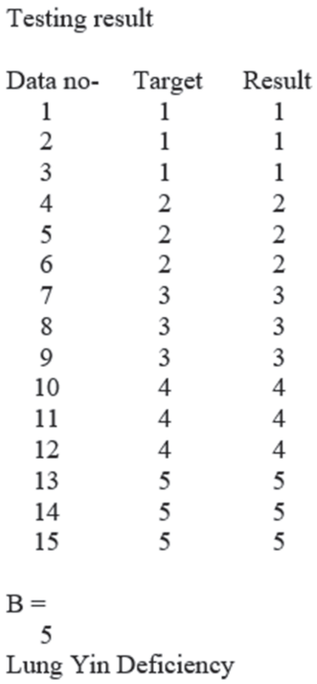

Fig. 2. Simulation result for data testing using optimal weights.

Figure 3 is a GUI program that is used to interact between the user and the computer program to determine the classification of lung syndrome. Figure 3 also shows a simulation with input symptoms such as pale complexion, dry cough without sputum, hoarse voice and a red tongue with lack of fluid. Based on selected symptoms, the simulation results show that the selected lung syndrome is lung yin deficiency syndrome. GUI program can also indicate treatment method, principle of therapy and acupuncture point therapy for selected patients for the syndrome.

LVQ GUI-based program can also be used for learning for students in determining lung syndrome. Students can enter symptoms into the GUI program. These symptoms are obtained by students through interview/anamnesis thereafter written on a sheet of patient data. To check whether the diagnosis result of lung syndrome through interviews with patients whether it is true or false the students enter the symptoms into in computer simulation by ticking the selected symptoms. Furthermore, the simulation results will show lung syndrome from patients. The use of computer simulations only acts as a facilitator that assists students in determining the symptoms of the patients is appropriate or not when explored at the interview. If the symptoms that entered to the computer simulation do not match the patient's condition, then automatically the simulation results are also not appropriate.

Table III shows the test results of the patients' data sheet that diagnosed by AAS students. Patients' data sheet used for testing are as many as 12 pieces. The simulation results show significant results, in which the simulation results nearly equal to the outcomes of students' diagnose. There is only one incorrect diagnosis that is at line number 7, where the results of a student diagnosed is Accumulation of Heat in the Lungs while the results of computer simulation classifies as Retention of phlegm Damp

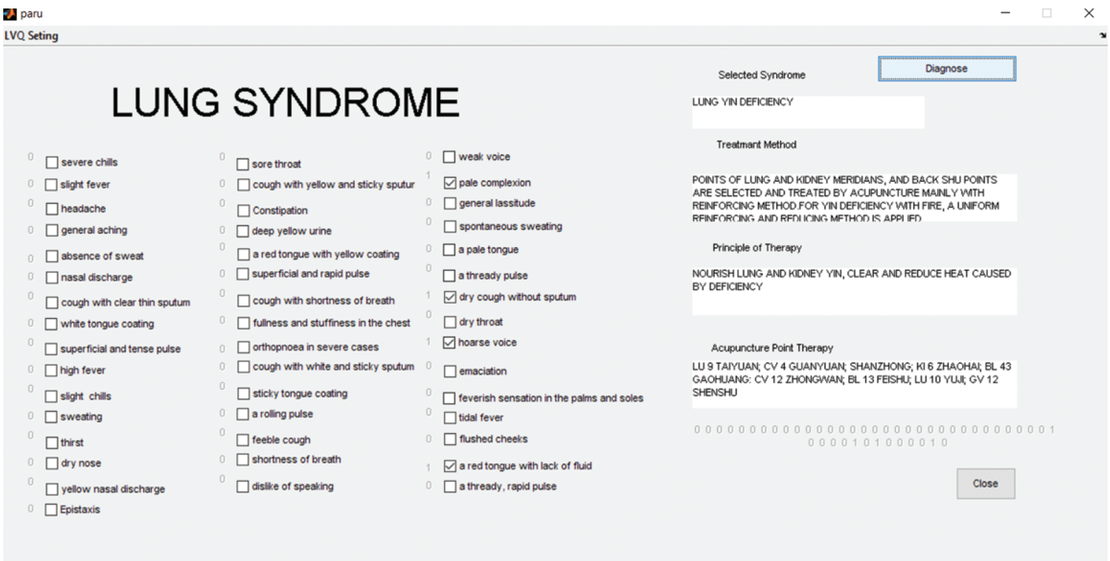

Fig. 3. Simulation result for data testing in a GUI program.

Table III. A set of appeared symptoms in lung syndromes.

No	Symptoms appear in data sheet patients	Syndrome diagnosis		Accuracy	
		Students	Computer simulation	Yes	No
1	Dry cough without sputum, a red tongue with lack of fluid, pale complexion, hoarse voice	Lung Yin deficiency	Lung Yin deficiency	✓	
2	Feeble cough, general lassitude, pale complexion, white tongue coating, pale tongue, shortness of breath, weak voice	Lung Qi deficiency	Lung Qi deficiency	✓	
3	Feeble cough, sweating, general lassitude, pale tongue	Lung Qi deficiency	Lung Qi deficiency	✓	
4	Feeble cough, pale complexion, shortness of breath, weak voice, pale tongue	Lung Qi deficiency	Lung Qi deficiency	✓	
5	Sweating, superficial pulse and tense pulse, yellow nasal discharge, cough with yellow and sticky sputum, yellow tongue coating, sticky tongue.	Accumulation of heat in the lungs	Accumulation of heat in the lungs	✓	
6	Absence of sweat, white tongue coating, sticky tongue coating,	Invasion of the lungs by wind cold	Invasion of the lungs by wind cold	✓	
7	Cough with shortness of breath, superficial and rapid pulse, thirst, general lassitude, dry throat	*Accumulation of heat in the lungs*	*Retention of phlegm damp in the lungs*		✓
8	Cough with yellow and sticky sputum, a red tongue with yellow coating, constipation	Accumulation of heat in the lungs	Accumulation of heat in the lungs	✓	
9	Cough with white and sticky sputum, shortness of breath, sweating, rolling pulse, fullness and stuffiness in the chest, sticky tongue coating, headache	Retention of Phlegm Damp in the Lungs	Retention of Phlegm Damp in the Lungs	✓	
10	Cough with white and sticky sputum, shortness of breath, rolling pulse, sticky tongue coating	Retention of phlegm damp in the lungs	Retention of phlegm damp in the lungs	✓	
11	Sweating, general lassitude, thirst, a thready pulse, a thready, rapid pulse, weak voice	Lung Qi deficiency	Lung Qi deficiency	✓	
12	Shortness of breath, a thready pulse, general lassitude, pale complexion, a pale tongue	Lung Qi deficiency	Lung Qi deficiency	✓	

in the Lungs. It occurs, because students are still in the learning phase in determining the type of lung syndrome.

4. CONCLUSION

Classification of types of lung syndrome can be done by using LVQ method. As it is known, that LVQ method has the advantage in classifying. The classification of lung syndrome-based LVQ is built by using input as many as 46 symptoms of lung syndrome, and the output as many as 5 types of lung syndrome. The simulation results show that LVQ can be used to classify the type of lung syndrome very well.

References and Notes

1. Z. Chen and P. Wang, *Evidence-Based Complementary and Alternative Medicine* 2012, Article ID 783923 **(2012)**.
2. B. Michael, A. Ghosh, and B. Hammer, *Journal of Machine Learning Research* 8, 323 **(2007)**.

3. B. K. Karayiannis Nicolaos and P.-I. Pai, *IEEE Transactions on Neural Networks* 7, 1196 **(1996)**.
4. U. M. Fahad, M. K. Sikander Hayat, *Information Technology Journal* 6, 154 **(2007)**.
5. W. Zhaowen, N. M. Nasrabadi, and T. S. Huang, *IEEE Transactions on Geoscience and Remote Sensing* 52, 4808 **(2014)**.
6. D. Mustafa, O. Yazicioglu, and A. Kentli, *Tehnicki vjesnik/Technical Gazette* 21, 689 **(2014)**.
7. P. Schneider, M. Biehl, F.-M. Schleif, and B. Hammer, Advanced metric adaptation in Generalized LVQ for classification of mass spectrometry data, *Proceedings of 6th International Workshop on Self-Organizing Maps* **(2007)**.
8. Z. Chalabi, N. Berrached, N. Kharchouche, and H. Mouhadjer, *Journal of Applied Sciences* 8, 1149 **(2008)**.
9. D. Frank, S. Müller-Hagedorn, H. M. Liebich, and G. Gauglitz, *Artificial Intelligence in Medicine* 28, 265 **(2003)**.
10. Gupta Dolly, Gour Sundar Mitra Thakur, and Abhishek, *International Journal of Computer Science and Information Technologies* 3, 3934 **(2012)**.
11. R. R. Janghel, Ritu Tiwari, and Anupam Shukla, *International Journal of Computer Applications* 975 **(2013)**.
12. E. Baltas, et al., An LVQ Classifier of EEG Coherence Patterns for Pain Detection **(2002)**.
13. A. Qureshi Hammad, Nasir M. Rajpoot, Khalid Masood, and V. Hans, Classification of Meningiomas Using Discriminant Wavelet Packets and Learning Vector Quantization **(2006)**, http://www.dcs.warwick.ac.uk/~hammad/Papers/miua_06.pdf.
14. C. Xinnong, Z. Bing, and W. Hongcai, Acupuncture Therapeutics, People's Military Medical Press, Singing Dragon, USA **(2011)**.
15. T. Kohonen, The self-organizing map, *Proceedings of the IEEE* **(1990)**, Vol. 78, pp. 1464–1480.

Received: 30 August 2016. Accepted: 30 May 2017.