Developing Intelligent Health Diagnosis System for Korean Traditional Medicine: Database Construction and Neural Inferences

Kwang Baek Kim[1], Hyun Jun Park[2] and Doo Heon Song[3]

[1] *Department of Computer Engineering, Silla University,*
140 Baegyang-daero(Blvd) 700beon-gil(Rd), Sasang-gu, Busan 617-736, Korea
[2] *Department of Computer Engineering, Pusan National University,*
Jangjeon 2-dong, Geumjeong-gu, Busan 609-735, Korea
[3] *Department of Computer Games, Yong-In Songdam College,*
Mapyeong-dong, Cheoin-gu, Yongin 449-040, Korea
[1]*gbkim@silla.ac.kr,* [2]*hyunjun@pusan.ac.kr,* [3]*dsong@ysc.ac.kr*

Abstract

In this paper, we propose a self-health diagnosis system based on Korea Traditional Medicine that has thousand years of history and popular in Korean general public. The system requires constructing a reliable diseases-symptoms database and classification learning method. Database construction is based on various reports submitted to the government about "Diseases burdensome to Korean Patients" in 2005 and medical contents "Engel Pharm" with 60 diseases. An enhanced ART2 algorithm that has dynamic control of boundary variables to control the number of clusters inside is used to extract top five most probable diseases generated by simple user input. The constructed database and test diagnosis results are verified by Korean Traditional medical doctors and the experiment also shows that the system is easily accessible and reliable in accuracy.

Keywords: Self health diagnosis, Korean traditional medicine, Enhanced ART2

1. Introduction

Korean medicine was originated in ancient and prehistoric times and can be traced back as far as 3000 B.C. It has its own origin and history and the treatment includes herbal medicine, acupuncture, moxibustion, and aromatheraphy. The legendary textbook of Korean traditional medicine, Dong-eui Bo-Gam [1], written in 17th century, was registered to UNESCO Memory of the world in 2009 [2].

The findings revealed that health service utilization among Korean American elderly covered a broad spectrum of health resources, including the use of Western medicine, the traditional Korean medicine (Hanbang), or both Western and traditional clinics [3]. Surely, Korean Traditional Medicine is more popular in Korea and it has its own medical institutes in university system and proven to be equally scientific as western medicine and even cooperated treatment with western medicine is increasing [4] in the last decade.

However, diagnosis based on Korean traditional medicine is not easy to understand by the public as its inference mechanism is largely metaphorical or abstract. There are few internet services for self-diagnosis, causes and treatment information based on Korea Traditional Medicine like [5, 6] for western medicine and if exists, it is only ad-hoc or incomprehensible.

As [2] pointed out, the main difficulties in building informative pre-diagnosis or self-diagnosis system is to build a reliable symptom-disease database and disease classification

system. Recently the government agent [7] published Korean Standard Causes of Death Disease Classification Index (KCD) for such a Korean specialized database build-up.

In this paper, we propose a self-health diagnosis system based on Korea Traditional Medicine. The system needs two major engines - database and disease inference engine. The symptom-disease database was refined from the first version [2] and the classification algorithm is also further developed to increase the classification accuracy based on ART2 [8] but add dynamic control of boundary variables to increase its reliable learning ability [9].

2. Symptom-Disease Database

2.1. Database Collection

We collect 739 diseases and 363 related symptoms based on KCD (Korean Standard Causes of Death Disease Classification Index) which replaces diseases of ICD(International Causes of Death) published by WHO with Korean traditional medicine [2]. However, the database used in this paper is refined with respect to various reports submitted to the government [10] about "Diseases Burdensome to Korean Patients" in 2005 and medical contents "Engel Pharm" thus the kinds of diseases considered in this paper is 60. There are 161 representative symptoms and 22 body parts that such symptoms are frequently appeared in the refined database. Again, collected information is verified by Korean traditional medicine doctors for reliability.

2.2. Database Construction

Figure 1 shows the database construction structure by diagram. The database entry form is the same as it was shown in the previous study [2].
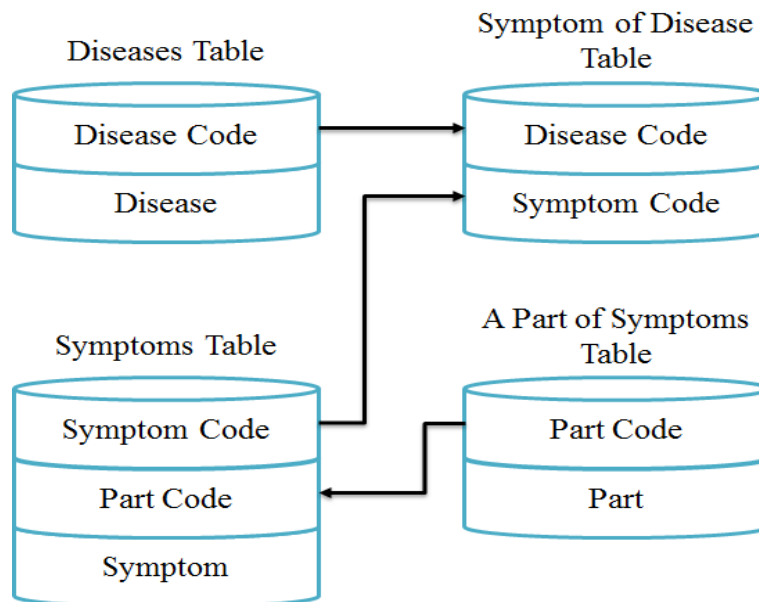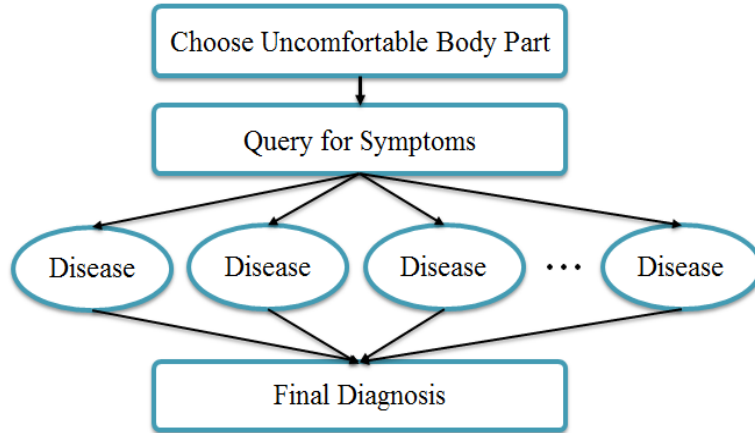

Figure 1. Structure of Database

Figure 2. Behaviour of the Proposed System

3. Proposed Self Health Diagnosis System

On user's side, the proposed system asks for the uncomfortable body parts and symptoms. With that answer, the system computes the similarity for diseases collected in the database (60 in total) and shows top five most probable diseases to the requested user.

For the unsupervised learning occurred in the system, we apply an enhanced ART2 algorithm for the input vector produced by the inquiry that will be explained below.

The final suspicious disease diagnosis follows the procedure depicted in Figure 3 for example. From the user's input (uncomfortable body part) at the bottom, the system computes the association rate with symptoms that are also given by the user. Then it computes the association rate with diseases in the database by our enhanced ART2 algorithm. Top five most probable diseases (the darker the more probable in this figure) are displayed to the user with computed similarity measure.

For the classification process, the proposed system uses an enhanced ART2 learning algorithm. ART2 is an unsupervised learning algorithm that does not have the target value. From an arbitrary input pattern, ART2 is a solution for the plasticity-stability problem without suffering from the local minima problem [8]. ART2 takes the average of all input patterns as the change of connection weights so that the effect is regulated to all clusters. However, it is sensitive to the boundary variables to establish the number of clusters.
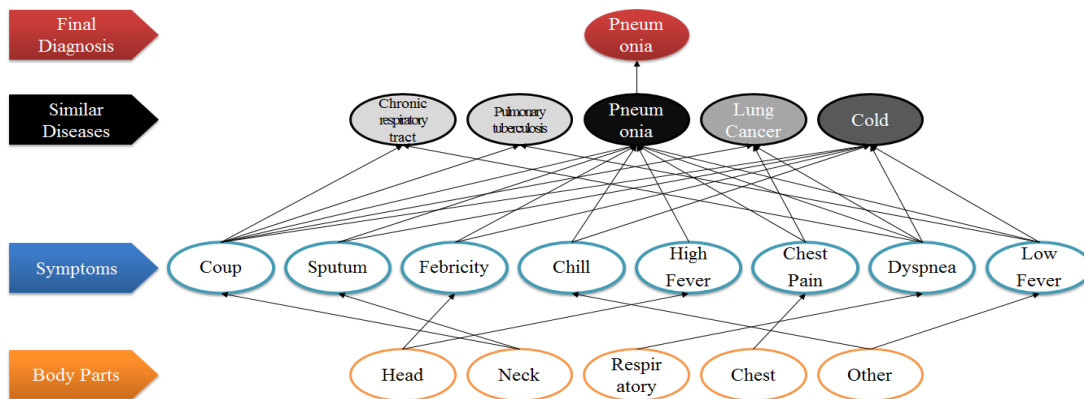
Figure 3. Suspicious Disease Extraction

Boundary variables are static values to decide the threshold of disagreement between stored pattern and input pattern for clustering. In self-health diagnosis, however, those values are usually set by experience. Unfortunately, disease diagnosis problem have many similar patterns thus deciding boundary values are very difficult.

In this paper, we propose a dynamic control of such boundary variables using Yager's intersection operator [11] like equation (1).

$$\mu_{x_1 \cap x_2} = 1 - \min \left\{ 1, \left((1 - \mu_{x_1})^p + (1 - \mu_{x_2})^p \right) \right\} \qquad (1)$$

where it is a monotonous decreasing function with respect to p(0,r).

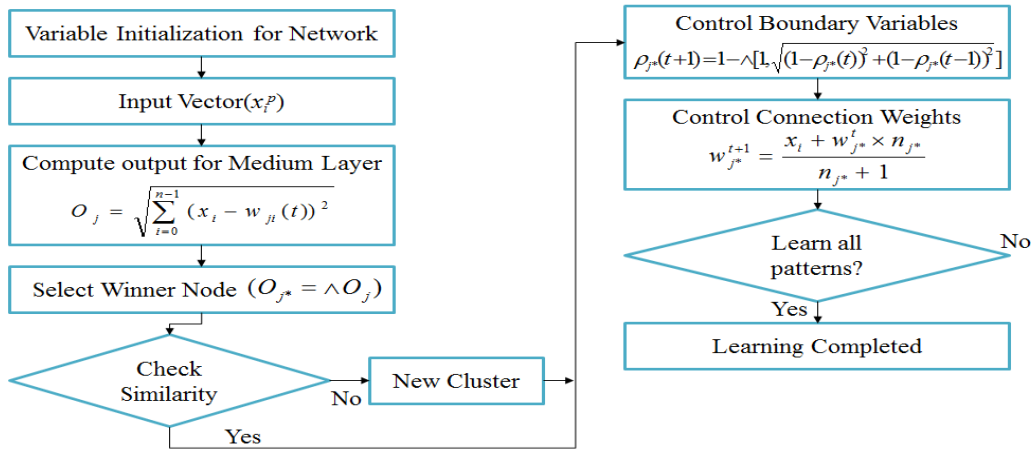The overall learning process by ART2 in this paper is as shown in Figure 4.

Figure 4. Learning Procedure for the Enhanced ART2 Algorithm

4. Experiment and Analysis

The implementation environment is as following; IBM compatible PC with Intel Pentium IV 2 GHz CPU and 1G RAM is used with JDK 1.6 and Oracle 10g are used in implementation and the system is available for on-line environment using JSP.

The system also has the incremental learning system for newly input disease as shown in Figure 5. For new disease, the system administrator gives information about the disease, symptoms associated with that disease, and frequently found body parts and then the incremental learning procedure is performed as shown in Figure 5 (the last button).

Figure 6 shows the structural diagram of the overall system. With user's input - uncomfortable body part and symptoms, the system searches for the corresponding disease and the input vector is given to the ART2 to decide most probable five diseases for that user.
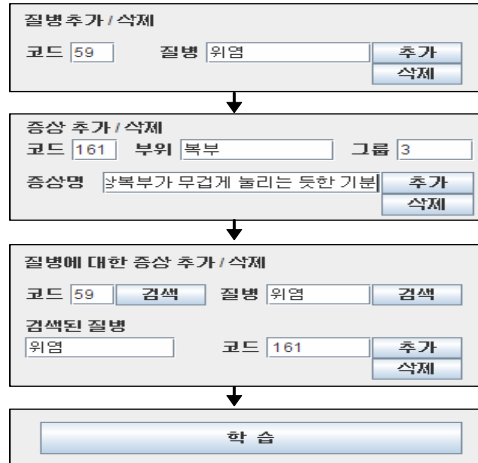
Figure 5. Learning with New Disease Input
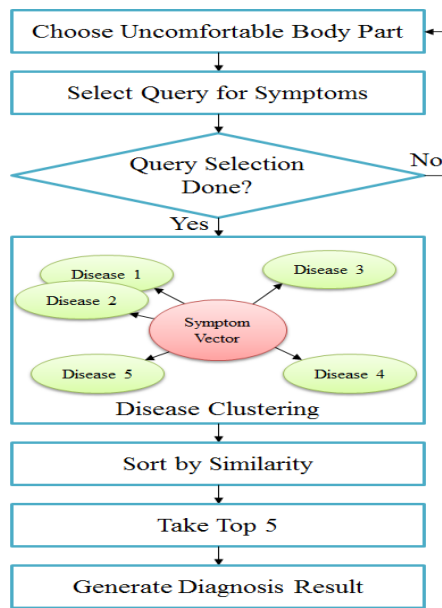
Figure 6. Structural Diagram

In our previous study [12], we used FCM and the enhanced FCM instead of ART2. That pilot system did not have enough information on diseases and symptoms thus the pattern clustering was not accurate enough and often inaccurately gave diagnosis.

Even with enough amount of database used in this paper, 161 symptoms from 60 diseases, one can easily find the power of enhanced ART2 over FCM and enhanced FCM through the result shown in Figure 7. The example shown in this paper has input symptoms as sputum, cough, chill, and dyspnea. Such general symptoms are associated with many diseases and Table 1 shows two related diseases, Pneumonia and Cold, as an example.

Table 1. Symptoms of Pneumonia and Cold

Disease / Body	Pneumonia	Cold
Head	Febricity(Low) High Fever	Febricity(Low) .
Other	Chill	Chill
respiratory organ	Dyspnea	Dyspnea
Chest	pain	.
Neck	Pain (swollen stinging) Cough Sputum	Pain (swollen stinging) Cough Sputum(rhinorrhea)

Although three different systems agree with "Cold" as the most probable disease, the difference is other extracted diseases. In Figure 7 (a), FCM picks lumbar disc as the third probable one which is not so much related to input symptoms. Enhanced FCM (Figure 7 (b)) did not commit such error but the overall confidences were relatively low. However, the proposed system extracts the exactly same result - cold, chronic respiratory tract, pneumonia, tonsillitis, asthma in order - with higher confidence. It is the evidence that enhanced ART2 forms more reliable clusters than enhanced FCM inside the system.

Including this example, many test cases were verified by Korean traditional medical doctor.

(a) FCM   (b) Enhanced FCM   (c) Proposed System
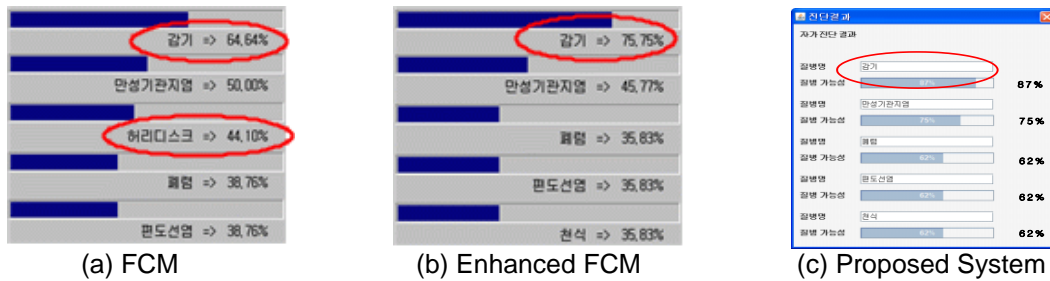
Figure 7. Result from FCM, Enhanced FCM, and the Proposed System

5. Conclusion

In this paper, we propose a self-health diagnosis system for Korea Traditional Medicine. While this traditional medicine is so popular among Korean general public, usually the diagnosis is not quite comprehensible for them since often the diagnosis is too much philosophical. The main reason is that the disease-symptom association is not so much standardized until recently

The self-health diagnosis system consists of two major parts - disease-symptom database and classification/learning algorithm. We construct the database based on KCD (Korean Standard Causes of Death Disease Classification Index) with Korean traditional medicine and used an enhanced ART2 algorithm for classification and learning. Simple user interface is provided and the user picks his/her uncomfortable body part and gives symptoms with respect to that body part. Then the system answers back with five most probable diseases with confidence rate.

From experiment and field expert's review, the proposed system is highly reliable and better in accuracy than the previous FCM based methods.

The system also has fuzzy inference system that compensates the possible misclassification by shortage of information and published for those concerns [2]. Currently this system is available with online access and in the developing phase for the mobile platform.

References

[1] S. K. Bong, "A Study on the Preservation and Utilization of Dongeuibogam", Journal of Korea Institute of Oriental Medicine, vol. 15, no. 1, **(2009)**, pp. 31-42.

[2] K. B. Kim and J. W. Kim, Self health diagnosis system with Korean traditional medicine using fuzzy ART and fuzzy inference rules, Intelligent Information and Database Systems, LNCS, Springer Verlag, vol. 7198, **(2012)**, pp. 326-335.

[3] M. Kim, H. R. Han, K. B. Kim and D.N. Duong, "The use of traditional and Western medicine among Korean American elderly", Journal of Community Health, vol. 27, no. 2, **(2002)**, pp. 109-120.

[4] D. Y. Chung, D. K. Baek, S. I. Hwang, S. H. Shin, D. W. Kim and M. A. Han, "One Case of Systemic Lupus Erythematosus Treated by Integrated Therapy of Western Medicine with Oriental Differential Diagnosis of Symptoms and Signs", Journal of Korean Traditional Internal Medicine, vol. 23, no. 2, **(2002)**, pp. 306-312.

[5] D. L. Berry, L. J. Trigg, W. B. Lober, B. T. Karras, M. L. Galligan, M. Austin-Seymour and S. Martin, "Computerized Symptom and Quality-of-Life Assessment for Patients with Cancer Part I: Development and Pilot Testing", Oncology Nursing Forum, vol. 31, no. 5, **(2004)**, pp. E75-83.

[6] K. H. Mullen, D. L. Berry and B. K. Zierler, "Computerized Symptom and Quality-of-Life Assessment for Patients with Cancer Part II: Acceptability and Usability", Oncology Nursing Forum, vol. 31, no. 5, **(2004)**, pp. E84-89.

[7] National Statistical Office of Republic Korea, http://www.kostat.go.kr.

[8] G. A. Carpenter, and S. Grossberg, "ART2: Self-organization of Stable Category Recognition for Analog Input Patterns", OPTICS, vol. 26, no. 23, **(1997)**, pp. 4919-4930.

[9] K. B. Kim, H. J. Park and D. H. Song, "Self Health Diagnosis System for Korean Traditional Medicine with Enhanced ART2", Proceedings of International Workshop (Bioscience and Medical Research), **(2013)**, vol.33, pp.16-19. December 11-13; Jeju Island, Korea.

[10] Ministry of Health and Welfare, http://www.mohw.go.kr.

[11] H. J. Zimmerman, "Fuzzy Set Theory and Its Applications", Kluwer Academic Publishers, **(1991)**.

[12] B. H. Chung, T. R. Chun, H.C. Kim and K. B. Kim, "Self Health Diagnosis System Using Enhanced FCM Algorithm", Proceedings of KIICE Conference, **(2006)**, pp. 143-149.

Authors

Kwang Baek Kim, received his M.S. and Ph.D. degrees from the Department of Computer Science, Pusan National University, Busan, Korea, in 1993 and 1999, respectively. From 1997 to the present, he is a professor at the Department of Computer Engineering, Silla University, Korea. He is currently an associate editor for Journal of The Korea Society of Computer and Information and The Open Artificial Intelligence Journal (USA). His research interests include fuzzy neural network and applications, bioinformatics, and image processing.

Hyun Jun Park, received his M.S. degrees from the Department of Computer Science, Pusan National University, Busan, Korea, in 2009. From 2009 to the present, he is a doctor course student at the Department of Computer Engineering, Pusan National University, Korea. His research interests include computer vision, image processing, neural network and applications.

Doo Heon Song, received a B.S. degree in Statistics & Computer Science from Seoul National University and M.S. degree Computer Science from the Korea Advanced Institute of Science and Technology in 1983. He received his Ph.D. Certificate in Computer Science from the University of California in 1994. He has been a professor at the Department of Computer Games, Songdam College, Korea, since 1997. His research interests include ITS, machine learning, artificial intelligence, medical image processing, cognitive, and game intelligence.