Application of Disease Pattern Analysis with Lifestyle Risk Factor for Healthcare Promotion Service

Young Sung Cho[1]
[1]Database and Bioinformatics Laboratory in
Department of Computer Science
Chungbuk National University
Cheongju, Korea
E-mail: youngscho@empas.com

Seon-phil Jeong[2*]
[2*]Department of Computer Science
Computer Science and Technology, Programme,
DST, BNU-HKBU United International College,
Zhuhai, China
E-mail: spjeong@uic.edu.hk

***Abstract*— Lately, the Critical Pathway(CP) of Electronic Medical Record(EMR) is used to the guideline for a treatment in the public hospital. We propose a healthcare promotion service using disease pattern with lifestyle risk factors. We classify a medical historical patient data with disease codes with lifestyle risk factors (hypertension, diabetes, smoking, overweight, excessive alcohol intake, and low physical activity) to make the lifestyle risk factors through the classification. We finally make the clusters of disease code with lifestyle risk factors using the medical historical data based on EMR's electronic discharge summary data. As the result of that, we do a healthcare recommending service based on the disease pattern with lifestyle risk. We can build a medical help desk of a public hospital to support people as we check into the public hospital; how to get the procedure of curing, the desired curing clinical method for the healthcare promotion service by each disease code, and how to be better our healthcare. We evaluate the performance of the proposed system by experimenting with the datasets collected at the medical center to measure performance and report some experimental results.**

***Keywords*-EMR; SVM; Classification; Clustering**

I. INTRODUCTION

Today, the development of life sciences and changes in the way of life seem to lead to an extension of life and lead to the chronic and metabolic trends of disease. The annual mortality rate from gastric disease causes serious health problems for Koreans. Diabetes is particularly difficult to recover once it is attacked, but treatment and healthcare can help alleviate diabetes and prevent complications for maintaining health and life. So, patients with diabetes should take diabetes education that can manage effective lifestyle risk factor according to the knowledge based on healthcare recommending service to treat successful diabetes, in order to improve their lifestyle and adapt to their real life. so patients can constantly receive the right treatment. Changes in lifestyle directly are related to the root cause of our disease. The objective of the study was to evaluate a proposing system as the healthcare recommending service for medical information service to reflect the degree of common lifestyle risk factors in the patient and to make the cluster deteriorated with lifestyle habits. The CP is a set of patient care guidelines that describe the goals of the patient care plan and the ideal sequence and timing of physician and staff actions to achieve this goal with optimal efficiency[1,2]. The use of the CP is intended to allow for the standardization of patient management, as well as improve the quality and efficiency of patient care. The CP is used to treat many medical conditions through continuous therapy [3,4]. The healthcare recommending services are required at the medical center for the patient's healthcare promotion services, healthcare curing services, and clinical care planning. The medical center needs a healthcare recommending service for medical information service as a function of the help desk to recover the patient's healthcare. It uses EMR's electronic discharge summary data, and then it presents a way to treat the disease or to improve a patient's healthcare based on successful healthcare treatment historical records. Active computer treatment information is increasingly interested in some predictions for medical diagnosis and healthcare recommending services. There are generally three methods, such as logistic regression, classification tree and neural network to attract attention in particular. There are important classification problems that are studied in various research fields such as machine learning, data mining, and statistical pattern recognition[5][6][7]. We propose health promotion service using disease pattern with lifestyle risk factor in health care plan. We combine patient data with disease codes that have lifestyle risk factors (hypertension, diabetes, smoking, overweight, excessive alcohol consumption, reduced physical activity) and classify medical records data using SVMs to determine the importance of lifestyle risk factors to emphasize the lifestyle risk factors. Finally, we provide health promotion services by analyzing disease patterns by using various input characteristics vector of Korean patients and disease code based on medical historical data based on electronic discharge summary data. This suggestion helps patients find ways to take healthcare recommending services and helps target patients in the healthcare center easily according to the healthcare recommending service. Therefore, patients and medical centers share medical services. The proposing service applies the technique of clustering disease pattern analysis to the healthcare recommendation service. It is important that this

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 April 2019 doi:10.20944/preprints201904.0095.v2

research is continued to restore the patient's condition at all times[8]. We make the solution of disease pattern with the risk rate based on lifestyle risk factors for medical information service. The tendency to accumulate risk is important for health promotion. Information on high-risk populations will help establish future prevention strategies. We conduct experiments with datasets from the medical healthcare center to measure its performance of the healthcare recommending service using disease pattern with lifestyle risk factors based on the electronic discharge summary data. We report some of the experimental results. The rest of this paper is organized as follows. Section 2 briefly describes the literature related to the study. Section 3 illustrates the proposed personalized healthcare recommending service in detail. We present experimental results in Section 4 and conclude in Section 5.

II. RELATIVE WORKS

A. EMR

In recent years, most hospitals have adopted electronic medical record systems instead of writing paper without loss of process structure, scope and information. We must comply with institutional, professional, or government regulations documents to increase electronic medical records. Much public organization for the healthcare industry can use patients' medical information as sharing the medical data. It includes Electronic Health Record (EHR) and Clinical Data Store (CDR). In addition, we have common medical information with various medical institutions. In recent years, hospitals have adopted an artificial neural network format of the electrical medical record system that computerizes medical records instead of writing records on paper without loss of process structure, scope and information [7].

B. Clustering

Clustering divides objects in a given group into a number of clusters, in order that the objects in a particular cluster. The objects are similar to the objects of the other clusters [9]. It groups physical or abstract objects into similar objects of a class. It classifies or segments the data into groups as the natural data structure. It belongs to the undirect data mining tools group[10][11]. It maximizes the similarity inside an object group and minimizes the similarity between the object groups. This partition is the simplest clustering algorithm.

C. SVM

It is designed for binary classification with a classifier even though it is considered easier to use than a neural network. Those are a useful technique to classify the data. The task usually takes training set and testing sets. It is necessary for us to have training data with relevance features of causations and irrelevance features of causations to do learning marked by users. It searches the database for more images related to the query. It has two kinds of task. One is interactive learning and the other is retrieval process to find correct features of causations increasingly.

III. OUR PROPOSAL FOR APPLICATION OF DISEASE PATTERN ANALYSIS

A. Application for Clustering Disease Code With Lifestyle Risk Factors

We generate cluster with the risk rate based on the lifestyle risk factors. The system analyzes the medical historical data, which is recorded by EMR data according to the healthcare recommending service. The prototyping application is run and the prototyping shows the result to classify disease pattern with lifestyle risk factors. From the result of the patient's membership data, we find 4 levels based on the rate of the patient's risk to recommend the services. We show output percent statistics for patient ownership based on weighted medical membership data based on lifestyle risk factors. We can use social variables such as age, gender, blood, region, and patient lifestyle risk factors(hypertension, diabetes, smoking, overweight, excessive alcohol intake, and low physical activity) as input vectors for pre-processing to make the risk rate of lifestyle risk factors through the classification for healthcare recommending service.

It is depicted in the result, that level 1 is the risk rate(1.0%) without risk factor, level 2 is the range of the risk rate(1.1%) with 1 time of risk factor, level 3 is the range of the risk rate(1.2%) with 2 times of risk factors and level 3 is the range of the risk rate(1.3%) with 3 times more of the risk factors.

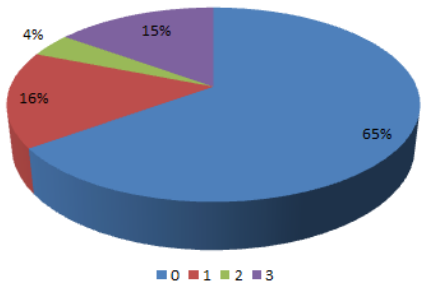

Figure 1. The result of the statistics of risk factors for possession of patients

The system reflects clusters with neighboring patient groups using the new clustering weighting according to lifestyle risk factors, that is, weights that classify them into generated classification codes through the algorithm of classification and patient's disease code based on medical historical data, the whole clinical datasets. Then, the system provides

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 April 2019 doi:10.20944/preprints201904.0095.v2

how to get the method improved the performance for a healthcare recommending service using the disease pattern with the lifestyle risk factors for the healthcare promotion service. The learning algorithm for clustering of medical history record. The representative pattern of bits includes patient's data tendency factors(age, gender, blood type, hypertension, diabetes, smoking, overweight, excessive alcohol intake, and low physical activity). The learning agent system scans pattern of bits using user information, using the algorithm inputs and learns input patterns. It also shows the procedural steps for learning algorithm as follows.

TABLE 1. THE PROCEDURAL STEPS FOR LEARNING ALGORITHM

Step 1: Calculate the distance between the input pattern and connection weight vector of each neuron.
Step 2: Provide the neurons of which the calculated distance are shortest and their neighboring neurons with the privilege.
Step 3: Modify connection weight vector so that the neurons with short distance may be closer to input patterns.
Step 4: Recognition starts at the time when learning finishes.
Step 5: Search the neuron from which input pattern is closest to among neurons in the learned competition layer.
Step 6: Recognize the searched neuron by comparing the input pattern with its output value.

The learning algorithm is based on the statistical model, and it works well on large networks to increase accuracy [13]. We use medical datasets using electronic discharge summary data for healthcare services in Korean hospital patients. We performed clustering to analyze disease codes to reflect the risk rate of lifestyle risk factors. We prototyped the application and got the results of the application. The result is classified disease patterns by lifestyle risk factors. This system provides a way to receive healthcare services by using the input vector as a clinical factor (age, sex, symptom, signal, system review, allergy, surgical history, family history, etc.). The system uses a new clustering of disease codes with risk rates based on lifestyle risk factors to create clusters with neighbors, that is classified by patient's disease code with the risk rate for risk factor using electronic discharge summary data. The system takes the preprocessing task that is able to use the whole medical datasets by the healthcare recommending service of the disease code with the risk rate and then makes the cluster of the medical datasets sorted by category of disease code. That is the cluster from patient data named by patient DB, including the same neighborhood patient data. Of course, the system uses the whole medical datasets. After all, the system provides how to get the healthcare promotion service by the desired treatment of disease code at a risk factor rate.

B. The healthcare recommending service for medical information service

We use disease patterns that reflect risk rate based on the lifestyle risk factors for health promotion services. The healthcare recommending service is a medical center help desk function. It is based on a successful treatment history record according to the HEALTHCARE RECOMMENDING SERVICE. It presents the best treatment of the disease or a plan to restore healthcare by reflecting the patient's the lifestyle risk rate. The system searches patient's disease code with the risk rate in user information. It scans the rate of curing disease code with the risk rate based on the lifestyle risk factors; the highest the desired curing rate of the disease code in the cluster. This system takes the task of clustering method using disease code with the risk rate based on the lifestyle risk factors using the medical history records. It provides the healthcare recommending service for the healthcare promotion service to the best healthcare recommending service with the medical treatment list.

IV. EXPERIMENTAL RESULT

A. EXPERIMENTAL DATA FOR EVALUATION

We do the experimental dataset with 250 patients who have had a clinic for treating a disease at a medical center, the data of 7 medical doctors, the data of 29 categories of disease codes about disease category used in the medical center. We use the 497 medical records for the clinic to conduct the experiments of the proposed system. It is a system performance evaluation for healthcare recommending services and can be assessed for precision, recall and F-measure in the clusters. There are 2 types of datasets collected from medical center. One is training dataset for nine months, the other is testing dataset. We conduct experiments with dataset of the medical center to measure its performance for healthcare recommending service. We report some experimental results.

B. EXPERIMENT AND EVALUATION

For doing clustering, the experimental datasets are transformed into an equivalent set of the simple sentence in temporal order by the formation. We make the evaluation, which is precision, recall and F-measure for proposing the system. We conduct experiments with data set of the medical center to measure its performance. We report some of the experimental results as follows.

TABLE 2. THE RESULT OF THE HEALTHCARE RECOMMENDING SERVICE

Cluster	**Proposing(SVM)**			**Previous(SOM)**			**Existing**		
	Precision1	*Recall1*	*F-measure1*	*Precision2*	*Recall2*	*F-measure2*	*Precision3*	*Recall3*	*F-measure3*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 April 2019 doi:10.20944/preprints201904.0095.v2

C1	35.56	30.94	32.25	36.25	27.98	30.79	36.25	10.19	15.63
C2	49.45	49.14	44.88	49.42	45.76	42.80	49.44	23.81	28.87
C3	38.34	51.86	42.51	39.20	46.75	41.28	39.20	18.33	24.32
C4	38.82	81.48	50.79	38.90	80.69	50.56	38.90	38.89	36.71
C5	43.04	70.34	50.23	45.31	61.37	49.45	45.31	28.70	33.08
C6	40.00	52.63	45.45	40.00	52.63	45.45	37.11	51.11	38.89

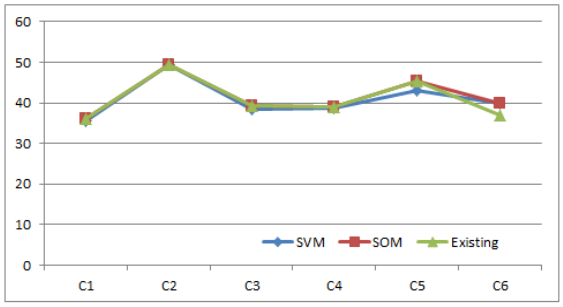

Figure 2. Precision, the result of healthcare recommending service

.

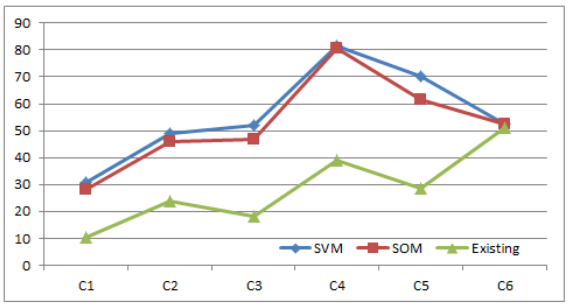

Figure 3. Recall, the result of healthcare recommending service

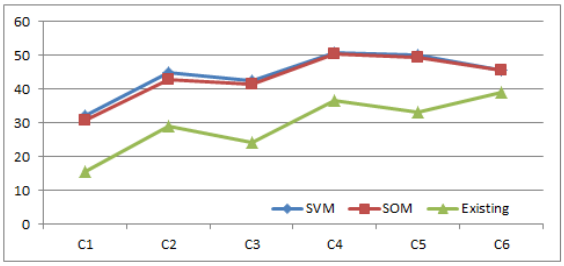

Figure 4. F-measure, the result of healthcare recommending service

The proposing system's overall performance evaluation presents the result of evaluation of healthcare recommending service for the healthcare promotion service with patients in Korea with the clinical rate on Table 2. It presents the evaluation metrics for recommendation system as seen in Table 2. The numbers of cluster (C1~C6) have finished grouping the medical historical data based on disease code (C1: gastric, C2: lung, C3: liver, C4: pancreas, C5: large intestine, C6: breast). The F-measure of proposing system is 14.77% , which is better than the existing system even if the precision of proposing system is lower 0.17 % than existing system, the F-measure of proposing system is 0.96 % , which is higher, tiny difference than the previous system even if the precision is lower 0.64 %, than existing system. As a result, we obtained recommender system with the desired curing clinical rate the for the healthcare promotion service withpatients in Korea. The result of the proposal using disease code with the risk rate based on the lifestyle risk factors is improved, which is better than the existing system in the performance, which is 3.54 better than the previous system in F-measure.

V. CONCLUSION

Nowadays, the hospital almost uses the EMR system without any loss of process structure, scope and content of information[11]. We used disease pattern analysis based on the medical historical data to consider the patient's the lifestyle risk factors. We finally made cluster based on the medical historical data. For doing the analysis of disease pattern based on the medical historical data, we segmented the patient's membership data to make 4 each level by the risk factors of the patient for the healthcare recommending service. We could install a medical help desk of hospital information to help people as they check into the hospital; how to get the method improved the performance of healthcare recommending service for medical information service by each disease code, and how to recover their health for the diagnosis. We evaluated the proposing system on the data set collected in a medical center to measure its performance. We reported some of the experimental results. The tendency for risk factors to cumulate has critical implications for healthcare promotion. Information on high-risk groups will help in planning future preventive strategies. Finally, it is meaningful to present a method of healthcare promotion service using disease pattern with lifestyle risk factors for the healthcare recommending service. We expect to give help in preventing disease and estimating prognosis by discovering useful knowledge that is provided by this service.

References

1. Pearson SD; Goulart-Fisher D; Lee TH. Critical pathways as a strategy for improving care: problems and potential, Ann Intern Med, 1995, 123(12), 941-948.
2. Coffey, R. J.; Richards, J. S; Remmert, C. S; LeRoy, S. S; Schoville, R. R.; & Baldwin, P. J; An introduction to critical paths. Quality Management in Healthcare, 2005, 14(1), 46-55.

 doi:10.20944/preprints201904.0095.v2

3. Hindle, D; Yazbeck, A. M. Clinical pathways in 17 European Union countries: a purposive survey, Australian Health Review, 2005, 29(1), 94-104.
4. Hand, D. J. Construction and assessment of classification rules. Chichester: Wiley. 1997.
5. Michie, D; Spiegelhalter, D. J; Taylor, C. C. Machine learning. Neural and Statistical Classification, 1994, 13.
6. Weiss, S. M; Kulikowski, C. A. Computer systems that learn: classification and prediction methods from statistics, neural nets, machine learning, and expert systems. San Mateo, Calif: M. Kaufmann Publishers, 1991.
7. Cho, Y. S; Ryu, K. H. Predictive Pattern Analysis using SOM in medical data sets for Medical Treatment Service. In 2014 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, IEEE. 2014, 1-5.
8. K. Vivekanandan; P. Krishnakumari. Discrete wavelet transformation of an image based on genetic-algorithm clustering," Indian Journal of Science and Technology. 2008, 1(3), 1-5.
9. Hand, D. J; Mannila, H; Smyth, P. Principles of data mining (adaptive computation and machine learning). Cambridge, MA: MIT Press, 2001.
10. Collier, K; Carey, B; Grusy, E; Marjaniemi, C; Sautter, D. A perspective on data mining. Centre for Data Insight, Northern Arizona University, USA, 1998, 2-4.
11. Cho, Y. S; Moon, S. C; Ryu, K. S; Ryu, K. H. A study on clinical and healthcare recommending service based on cardiovascula disease pattern analysis. International Journal of Bio-Science and Bio-Technology, 2016, 8(2), 287-294.
12. Chang, C. C; Lin, C. J. LIBSVM: a library for support vector machines. ACM transactions on intelligent systems and technology (TIST), 2011, 2(3), 27.
13. Herlocker, J. L; Konstan, J. A; Borchers, A; Riedl, J. An algorithmic framework for performing collaborative filtering. In 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 1999, 230-237. Association for Computing Machinery, Inc.