International Journal of Bio-Science and Bio-Technology
Vol.8, No.2 (2016), pp.287-294
http://dx.doi.org/10.14257/ijbsbt.2016.8.2.27

A Study on Clinical and Healthcare Recommending Service based on Cardiovascula Disease Pattern Analysis

Young Sung Cho[1], Song Chul Moon[2], Kwang Sun Ryu[1], Keun Ho Ryu[1,*]

[1] *Database and Bioinformatics Laboratory, School of Electrical and Computer Engineering, Chungbuk National University, Chungdae-ro 1, Seowon-gu, Cheongju, Korea, youngscho@empal.com, {ksryu, khryu}@dblab.chungbuk.ac.kr*
[2] *Department of Computer Science, Namseoul University, Sunghwan-eub Seobuk-gu, Cheonan-city, Korea, moon@namseoul.ac.kr*

Abstract

Recently, the clinical and healthcare recommending service is required in medical center for the clinical diagnosis and plan of treatment in connection with cardiovascula disease. We propose a method of the clinical and healthcare recommending service based on cardiovascula disease pattern analysis for medical treatment service. We use SVM(Support Vector Machine) to segment the clinical historical data, to join patients' clinical test data with input vectors of multi-parametric features, cardiovascula disease code, input factors and finally forms clusters of the clinical historical data based on electronic medical record. Then, we make an application on the clinical and healthcare recommending service for cardiovascula disease treatment information of cardiovascula patients to reduce patients' search effort to get the curing information and the diagnosis for recovering their health, to improve the accuracy for the clinical and healthcare recommending service. We carry out experiments with data set of medical center to measure its performance. We report some of the experimental results.

Keywords*: EMR(Electronic Medical Record), Cluster Analysis, SOM, SVM*

1. Introduction

Nowaday, most hospitals have adopted some form of EMR system that computerizes existing medical records which have been written on a paper without any loss of process structure, scope and content of information. The EMR becames the basic data for the ways to operate the effective controlling program for hospital management with nurses and patients. As new information about the biology of disease emerges, treatments will be developed and modified to increase effectiveness, precision, survivability, and quality of life [1]. A clinical and healthcare recommending service using information technologies, for instance clustering analysis of cardiovascula disease pattern, to meet the needs of patient to recover their health condition, it is important to remain as studied as always in the research. The thriving medical applications of data mining in the fields of medicine and public health has led to the popularity of its use in knowledge discovery in databases (KDD). Data mining has revealed novel biomedical and healthcare acquaintances for clinical decision making that has great potential to improve the treatment quality of hospitals and increase the survival rate of patients [2]. The recommending service is required in medical center for clinical diagnosis of disease, medical treatment service and the plan of medical treatment. It is necessary for them to obtain the service of helping information to recover their healthcare. It is the purpose of the function of help desk in medical center. It suggests how to cure the disease or how to make the medical treatment service to recover their health for patient's medical treatment based on successful clinical treatment history records using electronic discharge summaries data. Medical treatment

ISSN: 2233-7849 IJBSBT

service is increasing interest in some predictive methods for clinical diagnosis. Generally, there are three methods, that have attracted particular attention, logistic regression, classification trees, and SVM. There is an important problem of classification extensively which is studied in several research areas, such as statistical pattern recognition, machine learning and data mining [3-5]. The cardiovascular disease has become one of main diseases that threaten the human healthy especially in many countries. So, we propose a method of the clinical and healthcare recommending service based on cardiovascula disease pattern analysis for medical treatment service. We use SVM to segment the clinical historical data to use patients' clinical test data and we can extract new multi-parametric features from clustering algorithm is some kind of methods of pattern analysis using SVM in clinical data sets. Finally, we make the clusters of the clinical historical data to use the electronic discharge summaries data based on EMR with cardiovascula disease code, multi-parametric features in clinical data sets, input vectors to take cardiovascula disease treatment information of cardiovascula disease patients in order to do recommending service. This proposing method helps patient to find easily how to get the clinical and healthcare recommending service and helps their target patient in the medical center easily. Therefore, patients and medical centers take some benefit from the clinical and healthcare recommending service. We make the solution of medical treatment service via clinical and healthcare recommending service using cardiovascula disease pattern analysis. We carry out experiments with data set of medical center to measure its performance of the clinical and healthcare recommending service using cardiovascula disease pattern analysis for medical treatment service. We report some of the experimental results.

2. Cardiovascula Disease Pattern Analysis For Clinical and Healthcare Recommending Service

Many public organization for the healthcare industry can use patients' clinical information as sharing clinical data. For doing clustering about that, we generate the cluster for evaluation of proposing system for medical treatment service of cardiovascula patients. We expect to give help in preventing cardiovascula disease and estimating prognosis by discovering useful knowledge, that is, we can extract new multi-parametric features in clinical data sets in order to enhance the accuracy and efficiency of the experimental results obtained via a new method reflected by new multi-parametric features as the classifier for the clinical and healthcare recommending service. We use the SVM which is introduced by V. Vapnik. They are well founded in terms of computational learning theory and very open to theoretical understanding and analysis[6]. We use SVM learning to segment the clinical historical data, to join patients' clinical test data with input vectors of multi-parametric features, such as patient-id(varchar), disease code category(varchar), outpatient cure date(date), gender(boolean), age(numeric), hyper blood pressure(boolean), diabets millitus(boolean), smoking(boolean), old myocardial infarction(boolean), ejection fraction(numeric), blood glucose(numeric), total glucose(numeric), triglyceride(numeric), systoric blood pressure(numeric), diastolic blood pressure(numeric), and hyperlipidemia(boolean) etc., cardiovascula disease code, input factors and finally forms clusters of the clinical historical data based on the electronic discharge summaries data. SVMs are a useful technique for data classication. Although SVMs are considered easier to use than Neural Networks, SVMs are originally designed for binary classification, which has classifiers. SVM classifiers can be learned from training data of relevance feaures of causations and irrelevance feaures of causations marked by users. Classication task usually involves separating data into training and testing sets. Given a training set of instance-label pairs ($\mathbf{x}_i$,$\mathbf{y}_i$), i=1, 2, ..., *m* where input pattern $\mathbf{x_i}$ □□R_n and class $\mathbf{y}_i$□□□□1, □1□, SVMs, the aim of the SVM is to

find the optimal hyperplane that will classify each pattern x_i into the correct class y_i, the support vector machines (SVM), require the solution of the following optimization problem:

$$S = \{x_i y_i\}_{i=1}^{n} \ , \ \square(S) = \square\{ x_i, x_j\}.$$

In the linear case, the margin is defined by the distance of the hyperplane to the nearest of the positive and negative examples. The formula for the output of a linear SVM is

$$u = \vec{w} \cdot \vec{x} = b.$$

where *w* is the normal vector to the hyperplane and *x* is the input vector. The separating hyperplane is the plane u=0. The nearest points lie on the planes u = ±1. The margin *m* is thus

$$m = \frac{1}{\|\omega\|_2}.$$

A non-linear SVM maps the training samples of application from the original input space into a higher-dimensional space using a kernel function k ($x_i x_j$) [9]. When applied to two points x_i *and* x_j , k($x_i x_j$), is a generalised form of the inner product $x_i \times x_j$ in Equation (1). It is necessary for us to use the RBF(Radial Basis Function) as kernel fuction in order to do SVM learning as follows:

$$k_{RBF}(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right). \qquad (1)$$

The Lagrangian maximisation problem becomes:

$$\text{Max} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \ \alpha_j y_i y_j k(x_i, x_j). \qquad (2)$$

$$\text{s.t} \ \ \alpha_i \geq 0 \ \ i=1,\ldots,n, \sum_{i=1}^{n} \alpha_i y_i = 0.$$

The Lagrangian α_i can be used to solve the above optimisation problem. We take the proposal reflected new multi-parametric features to clinical data sets to join the electronic discharge summaries data based on EMR to classify cardiovascula disease pattern via a SVM learning for the clinical and healthcare recommending service. Then we create clusters of the clinical historical data based on the electronic discharge summaries data. The prototyping application is used and the prototyping shows the result to classify cardiovascula disease pattern. The system uses the social variable code such as age, gender, occupation, blood, region and patient's data factors as input vectors including symptoms, signals, clinic, etc, for pre-processing to be possible to provide how to get the clinical and healthcare recommending service with efficiency. The proposing system makes clusters with neighborhood patient-group using a new method reflected new multi-parametric features , that is, it classified by the classification as the classifier, which is generated via a SVM learning and patient's cardiovascula disease code in patient information based on clinical historical data. We use the whole clinical data sets. After that, the proposing system using SVM learning[6], provides how to get the clinical treatment information by improved method of performance of the cardiovascula disease code.

2.1. Application for Predictive Pattern Analysis using SVM to Segment the Clinical Historical Data

We use SVM learning to segment the clinical historical data, to join patients' clinical test data with input vectors of multi-parametric features in order to classify cardiovascula disease pattern for cardiovascula patients. A SVM classifier can be learned from training

data of relevance feaures of causations and irrelevance feaures of causations marked by users. Using the classifier, the system can retrieve more images relevant to the query in the database efficiently. It shows that the interactive learning and retrieval process can find correct feaures of causations increasingly. It also shows the procedural step for generalization ability of SVM under the condition of limited training samples of application.

Table 1. The Procedural Steps for Generalization Ability of SVM

Step 1. Retrieve by a traditional method.
Step 2. Mark top feaures of causations into two classes: relevance set I+ and irrelevance set IO.
Step 3. Prepare for SVM the training data (xi, yi).
Step 4. Construct classification function using SVM algorithm.

$$\text{Max} \quad \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \quad k(x_i, x_j) \qquad (3)$$

$$\text{s.t} \quad \alpha_i \geq 0 \quad i=1,\cdots,n, \quad \sum_{i=1}^{n} \alpha_i y_i = 0$$

Note: In order to output the similarity distance to the query, we ignored the function sign in the classifier f(x).
Step 5. Calculate the score for each image Ii in the database
Step 6. Sort all feaures of causations by score and return new result.

First, a large set of prototyping for clustering the clinical historical data, to join patients' clinical test data. From the result of clinical historical data, we find 6 level based on survival rate(%) of patient to recommend the services[4]. We show the result with statistics of percent output for possession of patients based on clinical historical data for the weight. We apply to make clustering of cardiovascula disease code to clinical data sets to join patients' clinical test data with input vectors of multi-parametric features using the electronic discharge summaries data to classify cardiovascula disease pattern for clinical treatment service of cardiovascula patients. We make the prototyping application and the result of application is divided to create clusters to be classified cardiovascula disease pattern. The system uses the input vectors as patient's clinical factors such as patient-id(varchar), disease code category(varchar), outpatient cure date(date), gender(boolean), age(numeric), hyper blood pressure(boolean), diabets millitus(boolean), smoking(boolean), old myocardial infarction(boolean), ejection fraction(numeric), blood glucose(numeric), total glucose(numeric), triglyceride(numeric), systoric blood pressure(numeric), diastolic blood pressure(numeric), and hyperlipidemia(boolean) and also, it is including symtoms, signals, system review, ergies, surgical history and family history, etc., to provide how to have clinical treatment service of cardiovascula patients efficiently. The system creates clusters with neighborhood patient-group using a new method of cardiovascula disease code, that is classified by SVM classifier as the classification and patient's disease code using electronic discharge summary. The system takes the preprocessing task that is able to use the whole clinical data sets by preferred curing clinical rate of the disease code and then makes cluster of clinical data sets sorted by category of cardiovascula disease code, joined cluster of patient data called by patient DB, neighborhood patient group. As a matter of course, the system uses the whole clinical data sets (medi_rd: called by clinical_record). After that, the system using SVM algorithm, provides how to get the clinical and healthcare

recommending service by improved method of performance of clustering of the cardiovascula disease code reflected by new multi-parametric features as the classifier.

2.2. Clinical and Healthcare Recommending Service for Medical Treatment Service

The recommending service for medical treatment service is the purpose of the function of help desk in clinical center. It suggests how to cure the cardiovascula disease or how to make the medical treatment service to recover their health for patient's clinical treatment based on successful clinical treatment history records. The system searches patient's cardiovascula disease code in patients' information. It scans the preference of curing cardiovascula disease, that is improved method of performance of the cardiovascula disease code, in cluster, suggest the preferred curing clinical treatment of cardiovascula disease code with the highest improved method of performance of cardiovascula disease code selected by the highest improved method of performance of cardiovascula disease code as the average of improved method of performance of cardiovascula disease code. This system provides recommending service for clinical treatment service by improved method of performance of the cardiovascula disease code predictively. This system generates recommending service for clinical treatment service efficiently through clustering method using cardiovascula disease code in SVM algorithm based on successful clinical treatment history records. It provides the associated recommending service for clinical treatment service to the best recommending service with clinical treatment list using electronic discharge summary for clinical treatment service of cardiovascula patients.

3. Experimental Result

3.1. Experimental Data for Evaluation

We have the experimental data of 250 patients who have the experience to have had the clinic for cardiovascula disease treatment in clinical center, the data of 7 clinical doctors, the data of 29 category of disease codes about disease category including cardiovascular and heart disease codes used in clinical center. The results of 497 clinical records for clinic is used to do the experiment of the proposal system to evaluate proposing system based on cardiovascula disease. It could be evaluated by MAE in clusters by improved method of performance of the cardiovascula disease code based on successful clinical history records for the clinical and healthcare recommending service. We report some of the experimental results through the experiment with learning data set for 9 months and testing data set for 3 months in a clinical center. We try to carry out the experiments in the same condition of the previous system.

3.2. Experiment and Evaluation

The initial classification of these examples of application had been performed by human inspection. For the support vector machine experiments, for each class of the examples of application were selected randomly for the training set and the remaining in total for the test set. We carry out experiments using MAE(mean absolute error) to measure its performance of proposing system for clinical and healthcare recommending service.

Table 2. The result of MAE by Comparing Proposal(SVM) with other System

Cluster	Proposal(SVM)	SOM	Existing
C1	5.6	7.26	58.29
C2	3.33	13.33	46.67
C3	6.2	9.69	55.79
C4	8.93	13.69	55.36

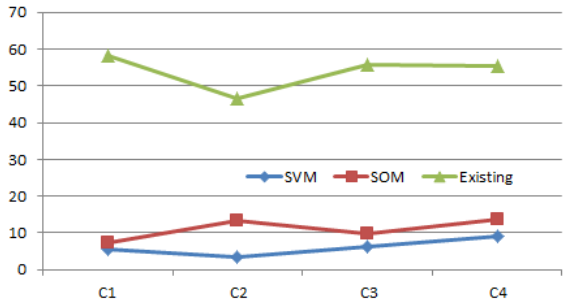

Figure 1. The Result of MAE by Comparing Proposal(SVM) with other System

The proposing system's overall performance evaluation presents above the result of evaluation of recommending service for clinical treatment service of cardiovascula patients with clinical MAE rate on the Table 2. The proposed is improved better than the previous systems. We use SVM learning to segment the clinical historical data. We report some of the experimental results as follows. The MAE rate of proposing system is 48.01 %, which is lower than existing system. As a result, we obtained recommender system with improved method of performance the for clinical treatment service of cardiovascula patients. The result of the proposal using cardiovascula disease code is improved, which is better than the previous system in the performance. As a result of that, we expect to give help in preventing cardiovascula disease and estimating prognosis by discovering useful knowledge, that is, we can extract new multi-parametric features in clinical data sets in order to enhance the accuracy and efficiency of the experimental results obtained via a new method reflected by new multi-parametric features as the classifier for the clinical and healthcare recommending service.

4. Conclusion

Recently almost hospitals have adopted the form based on EMR system that computerizes existing medical records that have been written on a paper without any loss of process structure, scope and content of information. The EMR system plays a key role in providing information for connection between all of systems managed in a hospital as well as gathering information for clinical research or strategic business[1]. Nowadays clinical and healthcare recommending service is required as an application based on EMR system to segment the clinical historical data. We reported the result of proposing experiment with SVM was improved more than the existing system. It could make the clinical and healthcare recommending service for each patient's cardiovascula disease based on clinical history data using electronic discharge summary in EMR in HL7 environment. We could simulate the application of SVM to classify cardiovascula disease pattern, generate clinical and healthcare recommending service to be possible to measure

the performance of clinical and healthcare recommending service by improved method of performance of the cardiovascula disease code based on successful clinical treatment history records. Thus, we could make clusters with the focus of accuracy and efficiency, and validate the system by our results. Then we could suggest an efficient recommending method for medical treatment service using electronic discharge summary, As a result, we could have the cardiovascula disease pattern analysis using electronic discharge summary for medical treatment service. We carried out experiments with data set of clinical center to measure its performance. Our proposing method was improved, which was better than the previous system in the performance, which was enhanced to the accuracy and efficiency of the experimental results obtained via a new method reflected by new multi-parametric features as the classifier for the clinical and healthcare recommending service. As a result, we reported some of the experimental results. It is meaningful to present a method of cardiovascula disease pattern analysis for clinical and healthcare recommending service.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (No.2013R1A2A2A01068923).

This paper is a revised and expanded version of a paper entitled “A Study on Cardiovascula Disease Pattern Analysis for Clinical and Healthcare Recommending Service” presented at [ksryu, December 16th 2015 at Jeju National University International Center, Jeju, Korea]."

References

[1] Y.S. Cho, K.H. Ryu, “Predictive pattern analysis using SOM in clinical data sets for clinical treatment service”, IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, **(2014)**, pp. 1-5.
[2] S. Kaur and R.K.Bawa, “Future Trends of Data Mining in Predicting the Various Diseases in Medical Healthcare System”, International Journal of Energy, Information and Communications, vol.6, no. 4 **(2015)**, pp.17-34.
[3] D.J. Hand, “Construction and assessment of classification rules”, Chicester, Wiley, **(1997)**.
[4] D. Michie, D.J. Spiegelhalter, C.C. Taylor, “Machine learning, neural and statistical classification”, Ellis Horwood, New York, **(1994).**
[5] S.M. Weiss, C.A. Kulikowski, “Computer systems that learn”, Morgan Kaufmann, San Mateo, **(1991)**.
[6] C. Cortes and V. Vapnik, “Support vector networks”, Machine Learning, vol. 20, **(1995)**, pp. 273--297.
[7] C. C. Chang and C. J. Lin, “LIBSVM: a library for support vector machines”, **(2001)**.
[8] K.J. Cios, G.M. Moore, “Uniqueness of clinical data mining: Uniqueness of clinical data mining”, Artif Intell Med, vol. 26, **(2002)**, pp. 1-24.
[9] D. Hand, H. Mannila, P. Smyth, “Principles of Data Mining”, The MIT Press, **(2001)**.
[10] K. Collier, B. Carey, E. Grusy, C. Marjaniemi, D. Sautter, “A Perspective on Data Mining”, Northern Arizona University, Flagstaff, **(1998)**.

Authors

Young Sung Cho, He joined a Professor of Computer Science at School of Dong Yang MiRae University, as well as hold a positions concurrently with Head of R&D Lab. in Comtree ltd., in Korea. His research interests include ime/Space DB, Machine Learning, Data mining, e-commerce, Web Service, Ubiquitous computing and Bio-Infometics. He recieved a MS in Computer Engineering from Yonsei University and a PhD in Computer Science from Chungbuk National University. email: youngscho@empal.com.

Song Chul Moon, He joined a professor of Dept. of Computer Science at Namseoul University in Korea. His research interests include Information Management, Information Security, e-commerce, and Ubiquitous computing. He recieved a MS in Management Information Engineering from the Korea. Advanced Institute of Science and Technology and a PhD in Information Management from Kukmin. email: moon@nsu.ac.kr.

Kwang Sun Ryu, He recieved a MS in Computer Science from Chungbuk National University. His research interests include Bio-Infometics, Dataming, Ubiquitous computing. He is a doctoral candidate in Computer Science from Chungbuk National University. email: ksryu@dblab.chungbuk.ac.kr

Keun Ho Ryu, He is a professor of Dept. of Computer Science at Chungbuk National University in Korea. His research interests include Time/Space DB, Temporal GIS, Knowledge based Information Retrieval, Data mining, Ubiquitous computing, Stream Data Processing, DB Security, Bio-Infomatics. He has a BS in Computer Engineering from SungSil University, MS and PhD from Yonsei University, email: khryu@dblab.chungbuk.ac.kr