RESEARCH ARTICLE OPEN ACCESS

The Case Study of an Early Warning Models for the Telecare Patients in Taiwan

Pay-Wen Yu*, Tian-Shyug Lee**
*(Graduate Institute of Business Administration, Fu Jen Catholic University, Taiwan
*(Graduate Institute of Business Administration, Fu Jen Catholic University, Taiwan

ABSTRACT
To propose a practical early warning analysis model for the telecare patients, this study applied data mining technology as a basis to investigate the classification of patient groups by disease severity and incidence using data contained in a telecare database regarding the number of a clinic. The ultimate purpose of this study was to provide a new direction for telecare system planning and developing strategies.
The subject of this case study was a private clinic which is providing telecare system to patients in Taiwan, and we used three data mining techniques including discriminant analysis, logistic regression and artificial neural network to construct an early warning analysis model based on several factors such as: Demographic variables, pathological signals, health management index, diagnosis and treatment records, emergency notification signal.
According the results, the telecare system can build stronger physician-patient relationship in advance through previously paying attention to patients' physiological conditions, reminding them to do self-management, even taking them to the hospital for observation. A comparison of discriminative rates showed that the artificial neural network model had the highest overall correct classification rate, 85.52%, and thus is a tool worthy of recommendation.
Keywords:- data mining, Telecare, early warning system, disease severity, artificial neural network

I. INTRODUCTION

More and more Information technologies are applied to healthcare industry such as medical affair operating (Electronic medical record system), medical service assistance (Picture archiving and communication system, Telecare system, or Doctor-surgery expert system), hospital administration supporting (Patient management system, Clinical system or Hospital resource planning system), and etc. Besides, the information management tools and theory are also implied in health fields. Especially discovering tacit or unknown knowledge from medical records data, using data mining to make decision support, or enhancing physical-patients relationship applied CRM [17].

Most Data mining algorithms can learn from past examples in clinical data and model the oftentimes non-linear relationships between the independent and dependent variables. The resulting model represents formalized knowledge, which can often provide a good diagnostic opinion [17][14]. It is important to extend the integration medical care service to upgrade the information technology applied to health industry.

We focus on telecare system, a service providing health and social care directly to people generally in their own homes, and study how the telecare system provides early-warning service to people through large amount data. The telecare instruments can build stronger physician-patient relationship (both service quality and satisfaction) in advance through previously paying attention to patients' physiological conditions, reminding them to do self-management, even taking them to the hospital for observation [15].

Therefore, to develop a practical early warning analysis model for the telecare patients, we conduct a case study to obtain data contained in a telecare database from a case clinic in Taiwan and use three data mining techniques (discriminant analysis, logistic regression, and artificial neural network) as a basis to investigate the classification of patient groups by disease severity and incidence.

The contribution of this study is to provide a new direction for telecare system planning and developing strategies.

II. LITERATURE REVIEW

1. Data mining in healthcare

As data sets grow to massive sizes, the need for automated processing becomes clear. Knowledge Discovery in Databases (KDD) is the process of getting high-level knowledge from low-level data. Data mining plays an important role in the KDD [16]. Data mining is the term used to describe the process of extracting value from a database. In today's competitive world, more and more companies store every piece of data they have collected, while others are more ruthless in what they deem to be "important". And with these large amount of data, the ability to extract useful knowledge hidden in these large amount of data and to act on the knowledge is becoming increasingly important [10].

The data mining technology is grounded by disciplines such as machine learning, artificial intelligence, probability and statistics theories. It can be divided into predictive model and descriptive model [9]. The predictive model often apply supervised learning functions to predict unknown or future values of other variables of interest. The descriptive model on the other hand, often apply the unsupervised learning functions in finding patterns describing the data that can be interpreted humans [10]. The predictive models are more commonly used in the healthcare field because there is a need of efficient analytical methodology for detecting unknown and valuable information in health data [2][9][11][12][19].

Data mining applications in healthcare can be grouped as the evaluation into: [1][3][7]

(1) Treatment effectiveness: Data mining can evaluate the effectiveness of medical treatments and prove effective by comparing and contrasting causes, symptoms, and courses of treatments.
(2) Healthcare management: Data mining can better identify and track chronic disease states and high-risk patients, design appropriate interventions, and reduce the number of hospital admissions and claims to aid healthcare management.
(3) Physician-patient relationship management: Through data mining application, physician-patient interactions may occur through call centers, physicians' offices, billing departments, inpatient settings, and ambulatory care settings.
(4) Medical fraud and abuse: Data mining applications can highlight inappropriate prescriptions or referrals and fraudulent insurance and medical claims through abnormally analysis.
(5) Medical device industry: Data mining techniques such as light weight, one-pass data stream mining algorithms can perform real-time analysis on-board small/mobile devices while considering available resources such as battery charge and available memory.
(6) Hospital Management: Through data mining application, service of hospital, medical staff, and patients treatment can be supporting effective and efficiency.

2. Telecare system

Telecare can be defined as a service providing health and social care directly to people, generally in their own homes, supported by information and communication technology. Telecare may be considered as an example of a complex socio-technical system. It provides safety and security monitoring, physiological and activity monitoring and information. McLean et al. describe three components of telecare [13]: (1) The patient is providing data, such as video, electrocardiography or oxygen saturation that gives information about the health status, (2) Information is transferred electronically to a health care professional at a second location, (3) The healthcare professional uses clinical skills and judgment to provide personalized feedback to the individual.

Telecare systems can support the independence and professional healthcare of older or disabled people. The functions of telecare system can be classified into seven items including teleported physiological signal, calling center, personal database, home care service, timely diagnosis, remote positioning, and health management consulting.

Many of the problems in implementing telecare systems relate to the organizational context within which it is being implemented, rather than the technology itself, and in the difficulties in apportioning the investment costs between different agencies [4][12]. As telecare is a care service delivered by means of technology, quality standards need to cover the organization of services and equipment. The following trends in the field of telecare are expected including increasing use of mobile health, increasing use of personal digital assistants, monitoring devices or applications, increased interaction between existing devices and more functions in one device, more user-friendly and economic device, enable used for prediction and prevention, and more extending use of telehealth into other fields.

3. Early Warning model in telecare system

Telecare systems offer the opportunity to react to hazardous events and also to prevent the further deterioration of the people ability to care for themselves or their health [18][20]. There are four core components of a telecare service which are physiological information monitoring, GPS and emergency rescuing, health counseling and interpersonal assistance, and healthcare liaison and coordination service [8][12]. According to the function of emergency notification and rescuing, telecare system offers continuous monitoring and observation of people's environment, daily activities at home, or behavior. Once system accepts abnormal signal, it will alert carers to do the emergency treatment. Because of the large amount of signal data, telecare system needs advance technology to process and analysis data. Both integrating hardware and software.

This study focus on the technique of date processing. We believe a telecare system not only provides alert model, but also needs to find potential emergency event. If telecare system can predict what kind of people might be onset soon, the probability of successful treatment will be [4][8]. The data mining techniques constructing predictive models are more commonly used in the healthcare field such as discriminant analysis, logistic regression and artificial neural network. We then apply the three methods to empirical data and compare the results.

III. RESEARCH METHODOLOGY

1. Research Framework

This study focused on the case of a private clinic which provides telecare service in Taiwan. And we used its medical care data to construct an early warning model to classify disease severity and incidence of the patients. The main purposes of the study were to sort the patients and work with the clinic telecare system to early-warning analyze of various related variables that may affect the disease severity and incidence of the patients.

Based on a summary of the literature and an analysis of the database before. Then, the focus shifted to ordering and filing the obtained independent variables that might affect patients' disease severity and incidence, including data on demographic variables, pathological signals, health management index, diagnosis and treatment records, emergency notification signal, and related variables. Next, using a dimensions of "whether the patient was onset or not onset," the patients in the database were divided into two groups and these two groups were the dependent variables used in the study.

To verify the applicability of the discrimination model, this study constructed the model using a two-stage mode that comprised a model training phase and a model testing phase, as shown in Figure 1. First, the patients' medical data from the case study clinic was randomly divided into training samples and testing samples according to an 80:20 ratio. In the first phase, model training, three early warning analysis models were established according to the three classification techniques used on the training samples.

In the second stage, model testing, the sample reserved for testing was used to assess the model's applicability and the test results were employed to predict the model's level of accuracy in discriminating among various groups.

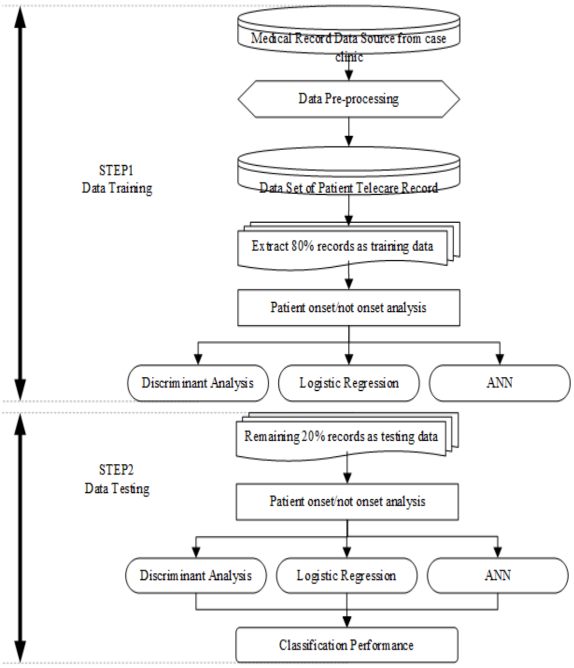

Fig.1: The Research Framework

2. Empirical Study

In this experiment the medical data related to Heart diseases is considered. This dataset was obtained from the case clinic's telecare patient database. The clinic's medical dataset concerns classification of patients into normal and abnormal person regarding dependent variable as heart disease severity and incidence.

The researchers entered into discussions with the owner of the case study clinic in March 2013, providing detailed explanations of the study's research motive, purpose, methodology, and expected results in an effort to strong physician - patient relationship in this endeavor. With the agreement of the clinic owner, patients' medical record database were acquired and underwent pre-processing, including data extraction, data aggregation, and data filtering.

After the dataset underwent preliminary consolidation, purification, and conversion, a total of 442 pieces of data on patients were obtained. The data fields included 23 related variables including patients' demographic variables, pathological signals, health management index, diagnosis and treatment records, emergency notification signal and etc. as Independent variables which may affect the patients' disease severity and incidence. The dependent variable is the record of emergency notification of the patients who have ever been on-set or not.

To enhance the multi-category characteristic discrimination results of the medium (small) samples of empirical data, this study employed separate tools including discriminant analysis, logistic regression analysis, and artificial neural network as sorting techniques to identify patients' onset or not. The process allowed us to construct, compare, and select the most suitable early warning scoring model.

IV. DATA ANALYSIS

1. Data Descriptive Statistics

The gender distribution among alumni was fairly evenly split between men and women, with men accounting for 40.72% and women accounting for 59.28%. Descriptive statistics regarding the patients' health status including sleeping quality, drink or not, take medicine or not are shown Table1.

Table1: Descriptive Statistics of Telecare Patients' health status

Variables	N	MIN	MAX	AVE	SD
Average age	442	27	75	42.79	12.12
Sleeping time (hours)	442	3	17	6.23	1.46

Variables	N	Percent
Gender	442	100%
Male	180	40.72%
Female	262	59.28%
Did you sleep well last night?	442	100%
Good	140	31.67%
Neither good nor bad	231	52.26%
Bad	71	10.06%
Did you drink last night?	442	100%
Yes	140	31.67%
No	231	52.26%
Did you take your medicine last night?	442	100%
Yes	46	10.41%
No	396	89.59%
Bad	71	10.06%
Emergency notification	442	100%
Onset	262	59.28%
Not-onset	180	40.72%

The patients' pathological signals are measured by Heart rate variability (HRV) analysis and other related equipment including systolic and diastolic blood pressure, body temperature, SNS (sympathetic nervous system), PNS (parasympathetic nervous system), Histogram, mean heart rate, SDNN (standard deviation of all normal to normal intervals), RMSSD, PSI (pressure index), VLF (very low frequency), LF (low frequency), and HF (high frequency). Descriptive statistics regarding are shown in Table 2. And health management index including wave type, emotional state, ABD (autonomic balance diagram), RRV (RR- interval variability) is assessed by telecare system as shown in Table 3.

Table2: Descriptive Statistics of Telecare Patients' pathological signals

Variables	N	MIN	MAX	AVE	SD
Systolic blood pressure	442	86	157	116.1	12.9
Diastolic blood pressure	442	59	97	77.9	8.8
Body temperature	442	36.1	363.4	39.4	31.0
SNS	442	1	4	2.4	1.0
PNS	442	1	4	2.4	.94
Histogram	442	1	5	3.3	1.0
Mean heart rate	442	53	647	80.5	55.8
SDNN	442	9	123	39.7	21.3
RMSSD	442	5.6	186	33.4	31.6
PSI	442	4.0	700	77.8	107.6
VLF	442	9.8	2672	455.9	471.5
LF	442	9.7	4598	381.4	577.2
HF	442	9.5	3086	457.6	666.5

Table3: Descriptive Statistics of Telecare Patients' health management index

Variables	N	Percent
Wave type	442	100%
Normal	99	22.40%
Decreased a little	37	8.37%
Decreased	15	3.39%
Decreased a lot	54	12.22%
Abnormal	237	53.62%
Emotional State	442	100%
Blue	178	40.27%
Red	151	34.16%
Green	113	25.57%
ABD	442	100%
Strong focus	31	7.01%
High sensitivity	55	12.44%
Low vitality	55	12.44%
Anxiety and impatience	37	8.37%
Poor immunity	264	59.73%
RRV	442	100%
concentrated	284	64.25%
dispersion	114	25.79%
Hollow	32	7.24%
Fault	12	2.71%

Results

This study was conducted by discrimination analysis, logistic regression, and artificial neural network using SPSS 17, and Weka software. The Independent variables were different chosen by three analysis as shown as Table 4, 5, 6. And dependent variable was patient is onset or not onset.

The overall correct classification rates of the three early warning discrimination models employed, namely, linear discriminant analysis, logistic regression analysis, and artificial neural network analysis, were 75.11%, 81.00%, and 85.52%, respectively, indicating that the model employing multivariate artificial neural network analysis generated the best results.

Table 4: Linear Discriminant Analysis (LDA) Results

		Predicted		Total
		onset	**not-onset**	
Original	onset	94	23	117
	not-onset	32	72	104
Total		126	95	221
The whole correct classification rate: 75.11%				

*Independent variables:

SDNN (standard deviation of all normal to normal intervals)

VLF (very low frequency)

Diastolic blood pressure

PNS (Parasympathetic nervous system)

Emotion State

Table 5: Logistic Regression Analysis (LRA) Results

		Predicted		Total
		onset	**not-onset**	
Original	onset	102	15	117
	not-onset	27	77	104
Total		126	129	92
The whole correct classification rate: 81.00%				

*Independent variables:

Diastolic blood pressure

SDNN (standard deviation of all normal to normal intervals)

VLF (very low frequency)

Table 6: Artificial Neural Networks (ANNs) Results

		Predicted		Total
		onset	**not-onset**	
Original	onset	106	11	117
	not-onset	21	83	104
Total		126	127	94
The whole correct classification rate: 85.52%				

*Independent variables:

SDNN (standard deviation of all normal to normal intervals)

ABD (autonomic balance diagram)

Histogram

HF (high frequency)

VLF (very low frequency)

PSI (pressure index)

Emotion State

V. Discussion and Conclusion

Due to rapid advances and ever-changing in information technology, it is easy to use high-speed computer to perform high-dimension and complex analysis. But it is difficult to analyze what themes and to find out all data relevant to that themes. Therefore, the starting of this paper was not only for the novelty and fun of research theme, but also to provide a solution for “which patient will onset?” that general existed in telecare system.

This research was focus on early-warning model of the telecare system. We conducted the empirical study form a private clinic which is providing telecare system to patients and used three data mining techniques including discriminant analysis, logistic regression and artificial neural network to construct an early warning analysis model based on several factors such as: Demographic variables, pathological signals, health management index, diagnose and treatment records, emergency notification signal.

According the results, the telecare system can build stronger physician-patient relationship in advance through previously paying attention to patients’ physiological conditions, reminding them to do self-management, even taking them to the hospital for observation. Empirical results indicate that out of the three discrimination models used in this study, the artificial neural network model yielded the best classification results and thus is a tool worthy of recommendation. We can make the conclusion as bellowing: (1) For telecare system, this study had the ability of data mining in classification and equipped with generalized construct procedures, so not only it could reach the precise forecasting, but also provide a reasonable solution for how to predict the patients whether if they onset or not exactly and rapidly. (2) For medical institutions, it was recommended that early warning model should be incorporated in their analytical process in order to decrease the onset risk.

REFERENCE

[1]. Ahmad, P., Qamar, S. and Rizvi, S.Q.A., “Techniques of Data Mining In Healthcare: A Review,” *International Journal of Computer Applications*, 120(15), June 2015, 38-50.

[2]. Bellazzi, R. and Zupan, B., “Predictive

Data Mining in Clinical Medicine: Current Issues and Guidelines," *International Journal of Medical Informatics*, 77, 2008, 81–97.

[3]. Dansky, K.H., Yant, B., Jenkins D. and Dellasega, C., "Qualitative Analysis of Telehomecare Nursing Activities," *The Journal of Nursing Administration*, 33, 2003, 372-375.

[4]. Desikan, P., Hsu, K.W. and Srivastava, J., "Data Mining For Healthcare Management," *2011 SIAM International Conference on Data Mining*, April, 2011.

[5]. Durairaj, M. and Ranjani, V., "Data Mining Applications in Healthcare Sector: A Study," *International Journal of Scientific & Technology Research*, 2(10), 2013, 29-35.

[6]. Ertas, A., Ganguly, S., Juric, R., Kataria, P. and Murat, M. T., "Sharing Information and Data Across Heterogeneous e-Health Systems," *Telemedicine and e-Health*, 15(5), 2009, 454-464.

[7]. Hirano, S. and Tsumoto, S., "Temporal Data Mining in Hospital Information Systems: Analysis of Clinical Courses of Chronic Hepatitis," *International Journal of Intelligent Computing in Medical Sciences and Image Processing*, 1(1):1, 2007, 11-19.

[8]. Hsu, C.H. and Tan, Y.J., "The Present and Future of Telehomecare," *Taiwan Geriatrics & Gerontology*, 3(4), 2008, 34-47.

[9]. Jothi, N., Aini, N., Rashid, A. and Husain, W., "Data Mining in Healthcare - A Review," *Procedia Computer Science*, 72(1), 2015, 306-313.

[10]. Kantardzic, M., *Data Mining: Concepts, Models, Methods, and Algorithms*, 2nd ed. Wiley-IEEE Press, 2011.

[11]. Koh, H.C. and Tan, G., "Data Mining Application in Healthcare," *Journal of Healthcare Information Management*, 19(2), 2005, 64-72.

[12]. Martin, S., *Research into the Future Housing and Support Needs of Older People, Electronic Assistive Technology – Supporting Older People within Local Communities*, Northern Ireland Housing Executive, 2010.

[13]. McLean, S., Protti, D. and Sheikh, A., "Telehealthcare for Long Term Conditions," *British Medical Journal*, 342, 2011, 374-378.

[14]. Milovic, B. and Milovic, M., "Prediction and Decision Making in Health Care using Data Mining," *International Journal of Public Health Science*, 1(2), December 2012, 69-78.

[15]. Poole, T., "Telecare and old people," U*K: King's Fund*, 2006, 1-26.

[16]. Rani, K.U., "Analysis of Heart Diseases Dataset Using Neural Network Approach," *International Journal of Data Mining & Knowledge Management Process*, 1(5), September 2011.

[17]. Silver, M., Sakata, T., Su, H.C., Herman, C., Dolins, S.B. and O'Shea, M.J., "Case study: How to Apply Data Mining Techniques in a Healthcare Data Warehouse," *Journal of Healthcare Information Management*, 15(2), 2001, 155-164.

[18]. Stanley, K.G. and Osgood, N.D., "The Potential of Sensor-Based Monitoring as a Tool for Health Care, Health Promotion, and Research," *Annals of Family Medicine*, 9(4), 2011, 296-298.

[19]. Tomar, D. and Agarwal, S., "A survey on Data Mining approaches for Healthcare," *International Journal of Bio-Science and Bio-Technology*, 5(5), 2013, 241-266.

[20]. Yoo, H.J., Yoo, J. and Long, Y., "Wireless fabric patch sensors for wearable healthcare," *in Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, 2010, 5254–5257.