Expert Systems with Applications 36 (2009) 3378–3387

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Improving scheduling of emergency physicians using data mining analysis

C.C. Yang [a], W.T. Lin [b], H.M. Chen [a,c,*], Y.H. Shi [a]

[a] Department of Industrial Engineering, Chung-Yuan Christian University, Taiwan
[b] Department of Industrial Engineering and Management, National Chin-Yi University of Technology, Taiwan
[c] Office of Medical Information Management, National Taiwan University Hospital, No. 7, Chung San South Road, Taipei 100, Taiwan

ARTICLE INFO

Keywords:
Data mining
Emergency department
Classification
6-Sigma
Performance evaluation matrix

ABSTRACT

Emergency departments are the first line in hospitals to face emergency patients. As a major function of emergency medicine, when a patient comes to the emergency department, the emergency medical personnel will first perform a triage procedure and then transfer the patient to associated departments for treatment. Due to the utilization pattern of the Taiwanese people in medicine, the emergency departments in most major hospitals are always overcrowded. The arrangement of manpower or the distribution of resources to handle patients' demands can affect disease outcomes and quality of medical treatment. Therefore, the prediction of demands of physician manpower certainly will affect the quality and cost in medical treatment, and has significant impact on patients' life and satisfaction.

This study used data mining, classification and a decision tree to analyze the prediction model of patients' demand in the Emergency department from real treatment situations. The result was the accuracy of shift anticipation improved from 22% to 50%. This study also used anticipant performance evaluation matrix integrated with loss function to evaluate the performance between the anticipation of demand established by mining and the original arrangement. It helped to save the cost of the medical personnel by 37%. In the end it combined the DMAIC action procedure from 6-Sigma and developed an anticipation model that can be suitable in different departments to dispatch medical personnel. It provided a reference of the decision maker of the hospital.

1. Introduction

Along with the development of the Internet and the maturation of database techniques, the methods of data and information storage in different businesses and professions have become more diverse and simpler today. In order to combine the applications of the Internet, many people are promoting data digitalization. Large quantity data collection techniques, high performance multi-processor computer structure and the maturation of data mining mathematical algorithms are the three most important elements for the prosperous development of data mining today. They have been extensively used in different business and professions (Kdnuggets, 2007). Generally medical institutes only use general statistics skill instead of data mining techniques to properly utilize this information. We need to think how to collect useful information from this enormous database and find valued knowledge for medical or hospital management; how to utilize this information effectively to improve the efficiency of management, quality and costs in the hospital; understand patients' demand for medical treatment and make a proper medical service strategy. But in fact, most of hospitals are not doing this at present.

Currently the dispatch of emergency manpower is always made by each department from experiences and emergency triage. However, patients' sickness distribution and variation in the patient numbers in peak and non-peak hours are not fully considered, so the emergency department might do jobs in a hasty and disorderly manner or might be over-crowded. While emergency medical procedures are heavy and complicated and emergency departments are always filled with patients, how can we establish a medical manpower deployment system to enhance medical quality and reduce medical disputes? (Academia Sinica, 1998; Yang & Yang, 2004) The main purpose of manpower resources planning is to reduce uncertainties. By clarifying the environmental uncertainties and planning before it happens, we expect to reduce the impact (Dessler, 2006; Richard, 1988). If medical manpower can be appropriately deployed, hospitals not only can provide medical care that is cost effective and also satisfying to the patient (Chou, 2003; LaMar, Jacoby, Meyer, & Potter, 1997; Yeh, Wang, Chen, & Li, 2003).

In the past most studies in physicians' shift arrangement emphasized outpatient physicians' shift arrangement (Liu & Wang, 2005); however characteristics of the demand on emergency physicians are a little bit different. There are two major differences between emergency services and outpatient services. The first is

---

* Corresponding author. Address: Office of Medical Information Management, National Taiwan University Hospital, No. 7, Chung San South Road, Taipei 100, Taiwan. Tel.: +886 2 2356 2904; fax: +886 2 2341 0239.
E-mail addresses: admchmz@ntuh.gov.tw, admchmz@ntu.edu.tw (H.M. Chen).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.02.069

the urgency of physicians' diagnosis. Generally patients in the emergency departments need urgent treatment in a very short time. Therefore, emergency departments have a triage system in accordance with patients' condition (Tsan, 2006). The second major difference is that the emergency service is provided 24 h a day, while the outpatient physicians have demand-led shifts.

In this research, hospitals divide outpatients into 6 departments, including internal medicine, surgery, traumatology, pediatrics, obstetrics and gynecology and dental departments. The rule of department assignment does not depend on patient numbers but because of the hospital's policy and the payment system of the National Health Insurance. Also because of hospital organizational structure and demands from medical education, emergency physicians do not take all the shifts in the emergency department; generally the above mentioned six departments assign a resident physician to perform the first line services in the emergency department each month. Therefore, the emergency department is like an epitome of a small hospital. The study takes the largest department in the emergency department, the internal medicine department, as the research object and makes some suggestions to improve the arrangement of the physicians' shift schedule.

In this research, the objective of human resources utilization is that the numbers of physicians should be considered by patients' requirement in order to promote the quality of emergency service and arrange proper number of personnel. It compares current internal medicine physicians' shift arrangement in the emergency department and selects the variable. In the emergency medicine of the case hospital, demand of the patients' side is to triage and provide emergency treatment to the patient in the first place. On the supply side it should include the allocation of personnel from different fields such as physicians, nursing staff, technicians, social workers, first aid technicians, administrative staff, janitors and volunteers. This study focuses on direct manpower supply of emergency physicians; accredited physicians from other departments and other indirect personnel are not considered.

To conclude from the above, according to actual demands for physicians, it is planned to use triage technique as a preliminary analysis tool to analyze the emergency patient inspection and shift arrangement data. After developing a forecast model, a performance evaluation matrix combining Taguchi's Loss Function is the final performance evaluation method. This procedure is also used to build a 6-Sigma's DMAIC action procedure (George, Rowlands, Price, & Maxey, 2005) to achieve a continuing improvement administrative circulation for future references, shown in Fig. 1.

2. Definition of the forecast performance evaluation matrix and control line

2.1. Definition of the forecasted performance evaluation matrix index

The performance evaluation matrix in the past took the Performance Evaluation Matrix that was mentioned by Lambert and Sharma (1990) as a reference for performance improvement strategy. Lin, Chen, and Chen (2005), however, modified the concept of Lambert and Sharma's Performance Evaluation Matrix. First we used the matrix composed by the level of important simple as an example explanation. According to the concept, the index of importance level is the vertical coordinate and the index of simple level is the horizontal coordinate. The range of both index values is between 0 and 1. In this study, it is modified by the moderate performance evaluation matrix. The forecast value is the vertical coordinate and the real value is the horizontal coordinate. The new area is re-defined as shown in Fig. 2, which follows.

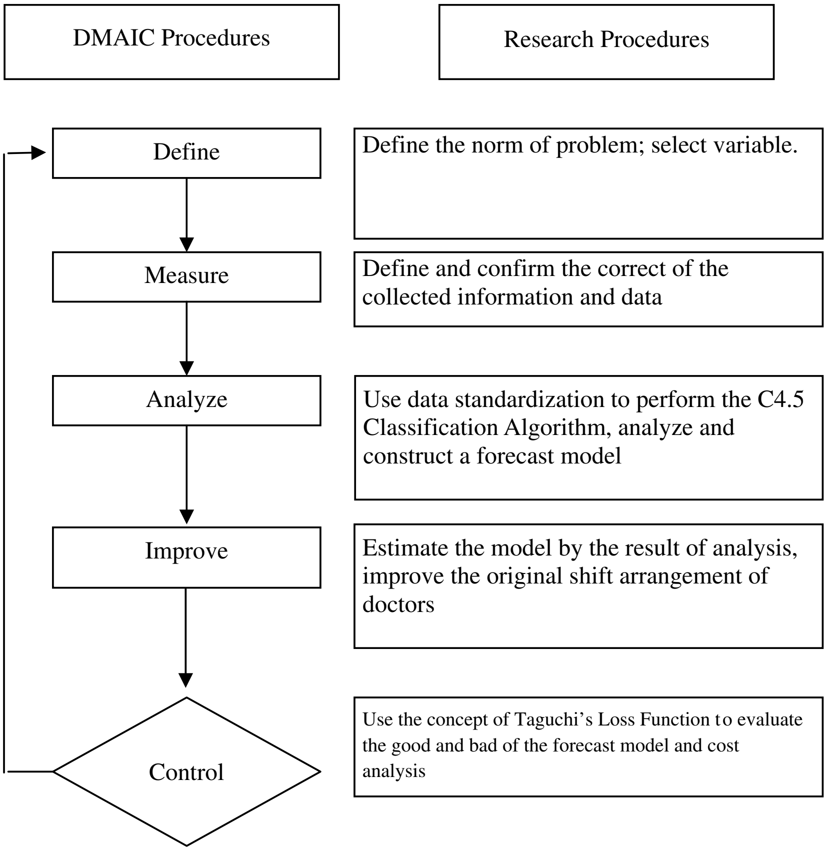

Fig. 1. The DMAIC research procedure.

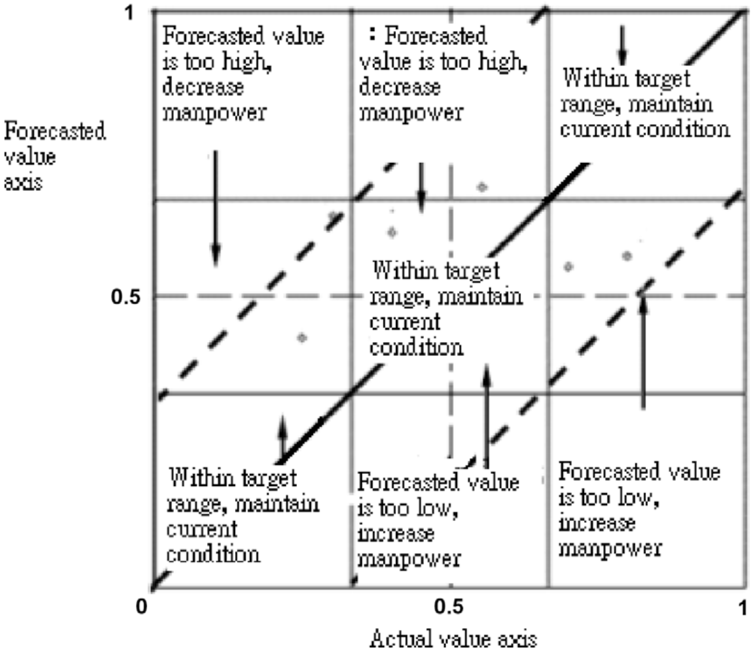

Fig. 2. Forecasted performance evaluation matrix area.

Then this study refers to the concepts of Parasuraman, Zeithaml, and Berry (1991) to utilize random variable T as the real number of physicians needed and F as the forecast or real number of physicians listed in the shift schedule. The index formula is modified as below:

T_{ij} is the actual physicians demanded on the J Shift at Day i (actual patient numbers/the weighted value of every physician to serve patient), F_{ji} is the forecasted or actual scheduled number of physicians in Shift J at Day i, PT_{ij} is the actual value index:

$$\mathrm{PT}_{ij} = \frac{T_{ij} - \min}{R} \qquad \text{(formula 1)}$$

PF_i is the forecasted or actual shift index:

$$\mathrm{PF}_{ij} = \frac{F_{ij} - \min}{R} \qquad \text{(formula 2)}$$

i is the 1 (day shift), 2 (night shift), $j = 1,\ldots,n$, min is the minimum value of the real number of physicians needed, forecast number of physicians, real number of physicians listed in shift schedule, R is the range of the real number of physicians needed, forecast number of physicians, real number of physicians listed in shift schedule.

2.2. Definition of the control line

Estimating the model shift schedule from data mining to achieve our object is having the best human resources arrangement by our estimate. We use forecast performance evaluation matrix to develop strategies to improve and manage physicians' shift arrangement. We used the Shewart Control Figure to define the performance control line and set up the target value at 0. According to the Central Limit Theorem, if the number of effective samples is more than 30 then the discrete distribution will tend to become normal distribution. According to the experience, the probability of falling into the range of ±3 times standard deviation is 99.73%, meaning the un-qualification rate will be 0.27%; the probability of falling into the range of ±2 times of standard deviation is 95.44, meaning that the un-qualification rate will be 4.56%; the probability of falling into the range of ±1 time of standard deviation is 68.26%, meaning the un-qualification rate will be 31.74. In this study, we chose ±3 times standard deviation to set the upper control line (UCL) and lower control line (LCL). The equation is shown below:Upper control line $\mathrm{UCL} = T + 3\sigma = 3\sigma$
Central line target value $T = 0$
Lower control line $\mathrm{LCL} = T - 3\sigma = -3\sigma$

And the Mean of the Population μ_ρ and Standard Deviation of the Population σ_ρ are used to find the UCL and LCL. We hereby assume the 6-Sigma Action Procedures comply with normal distribution, therefore we can find the μ_ρ value and σ_ρ value from the following equations:

$$\mu_\rho = \frac{\sum_{i=1,j=1}^{n} (y_j - x_i)^2}{n} \qquad \text{(formula 3)}$$

$$\sigma_\rho = \sqrt{\frac{\sum_{i=1,j=1}^{n} (y_j - x_i)^4}{n} - \mu_\rho^2} \qquad \text{(formula 4)}$$

From the above UCL and LCL definitions, we can get formulas (formula 5) and (formula 6) as follows:

$$\text{Upper Performance Control Line} : \mathrm{UCL} = 3 * \sqrt{\frac{\sum_{i=1,j=1}^{n} (y_j - x_i)^4}{n} - \mu_\rho^2} \qquad \text{(formula 5)}$$

Central line target value: $T = 0$

$$\text{Lower Performance Control Line} : \mathrm{LCL} = -3 * \sqrt{\frac{\sum_{i=1,j=1}^{n} (y_j - x_i)^4}{n} - \mu_\rho^2} \qquad \text{(formula 6)}$$

We use the concept of control line to divide the forecast performance evaluation matrix into four performance zones. Each represents the level of the difference between the actual daily manpower needed and our forecast value. Zone One is the area above the UCL as shown in Fig. 3. Any forecast point falling in Zone One indicates the actual value is less than the forecast value, meaning that day physicians are arranged more than actually needed. Because of the idle manpower; this area is called "Zone of excess manpower". Zone Two is the area between the UCL and central

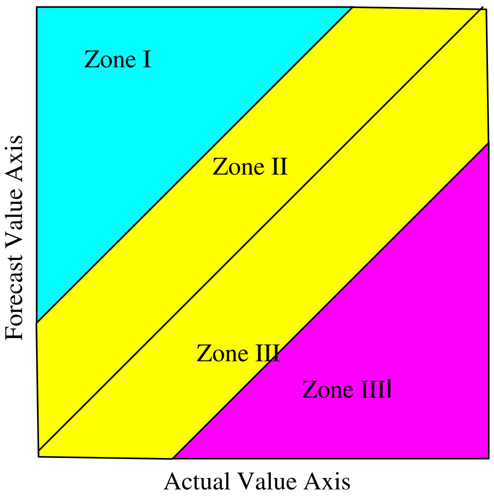

Fig. 3. Forecast performance control line.

line. Any point falling in this area indicates the actual value is less than the forecast value but still can be accepted. The area is called "Zone of moderate manpower". Zone Three is the area between the LCL and the central line. Any point falling in this area indicates the actual manpower needed is more than the forecast but still can be accepted. This area is called "Zone of moderate manpower II". Zone Four is the area below the LCL. Any point falling in Zone Four indicates the actual manpower needed is much more than the forecast. It means the on-duty physicians are less than actually needed on that day so the manpower is deficient. This area is called "Manpower deficiency zone".

3. The measure of data collection and analyzing the data mining

3.1. The measure of data collection

Variables are selected after discussion with hospital administration executive. We use the existing columns in emergency department's database including patients' attributes such as Date of diagnosis, Triage, Department. to be the input predictor of our daily shift according to day/night shift, week and national holiday or not. To ensure the compliance of our study samples and all kinds of data mining algorithms, we set up a measure for judging the collected sample:

(1) *Handling the missing values:* Delete patients that have registered but do not check with physicians and also delete patient information with incomplete case histories, to assure physicians do perform services for those patients.
(2) *Duration of samples:* Data is drawn from Emergency Medicine Department database from 2004 to 2005.
(3) *Sample subjects:* Patients in the database of emergency department database and physicians in the medical department of the case hospital from 2004 to 2005.

3.2. To analyze the data mining

The study plans to use the Decision Tree from the data mining method to do analysis procedures. Hopefully through decision tree, the final constructed decision rules are able to find characteristics of patients' demand and predict physicians' shift arrangement rules to be followed in the future. In emergency medicine, because different conditions of different patients require different treatment demands; we have to consider the actual situation to appoint the weighted value, while converting numbers of patients into numbers of physicians needed.

Decision tree is the most common fundamental classification in all classification methods. There are many kinds of algorithms in the decision tree method; the most common methods are ID3, PRISM and Gini, etc. (Adriaans & Zantinge, 1996; Berry & Linoff, 2004; Giudici, 2003; Hand, Mannila, & Smyth, 2001; Kantardzic, 2002; Roiger & Geatz, 2003; Soukup & Dabifdon, 2002). The ID3 was a decision tree algorithm presented by Quinlan (1986). The C4.5 decision algorithm used in this study was also presented by Quinlan (1993) by modifying ID3 algorithm, adding in missing values predictors and continuous value predictors, and performing the pruning process from bottom to top to delete bad pruning and construct a decision tree for historical information.

Both ID3 and C4.5 use Shannon's (1948) Information Theory and Entropy as a basis to test the attributes; its basic calculation principles are stated in the following.

3.2.1. Branches of the decision tree

If we assume there will be n kinds of results in this study (n kinds of shift arranges, or n different classifications) and the probability of occurrence for these n kinds of results are $P(X_1),\ldots,P(X_n)$, respectively, then it is defined that the information volume obtained will be:

$$I(P(X_1),\ldots,P(X_n)) = \sum_{i=1}^{n} -P(X_i)\log_2 P(X_i) \qquad \text{(formula 7)}$$

Furthermore, if it is assumed that there are $P_1,\ldots,P_j$ classification marks, while A represents certain attribute, V represents the sample population before testing the attributes, $V_1,\ldots,V_k$ represent sample sub-groups after testing the attributes, X_1 represents numbers of P_1 in V, X_2 represents numbers of P_2 in V and X_n represents numbers of P_j in V, then it is defined that when V is divided into $V_1,\ldots,V_k$ according to the value of attribute A, the information gained is:

$$\text{Gain}(V) = I(P_1,\ldots,P_j) - E(V) \qquad \text{(formula 8)}$$

$$I(P_1,\ldots,P_j) = -\frac{P_1}{P_1+\cdots+P_j}\log_2\frac{P_1}{P_1+\cdots+P_j} - \cdots - \frac{P_j}{P_1+\cdots+P_j}\log_2\frac{P_j}{P_1+\cdots+P_j} \qquad \text{(formula 9)}$$

$$E(V) = \sum_{i=1}^{k} \frac{P_{1i}+\cdots+P_{ji}}{P_1+\cdots+P_j} I(P_1,\ldots,P_j) \qquad \text{(formula 10)}$$

From formula 8, we obtain all kinds of possible branch predictor information gain, because the smaller the information volume, the smaller value of the Entropy is. And it is hoped through classification, the training samples can be classified into sub-groups with the smallest Entropy value. Therefore, in ID3 Algorithm it mainly selects the attribute with the smallest information volume after testing, i.e. to choose the attribute with the largest information gain.

3.2.2. Pruning of the decision tree

In C4.5 Algorithm, an Expected error rate pruning method was presented. By using the possibility concept of Binomial Distributions to calculate the predicted acceptable numbers of wrong classification, so that to inspect if the branched decision tree is still within the tolerable value, so that to proceed with the pruning procedure. The pruning procedures and formula are as follows:

(1) From bottom of the decision tree to inspect every sub-group; however sub-groups with one node are not taken into consideration.

(2) Suppose numbers of samples that are classified in the sub-group are N and numbers of samples that are not classified in that sub-group are E, then CF parameter is set.

(3) It can use Binomial Distribution to calculate the sub-group's $U_{\mathrm{CF}}(E;N)$ value, i.e. the upper tolerable mistake value. The formula is listed below.

Let CF possibility of the appointed parameter (in between 1 and 0)
E sample number of the sub-group in correct decision
N total sample number of the sub-group

$$\mathrm{CF} = B(E; N, U_{\mathrm{CF}}(E; N)) \qquad \text{(formula 11)}$$

(4) Add all $U_{\mathrm{CF}}(E;N)$ values from all sub-groups and compare them with the $U_{\mathrm{CF}}(E;N)$ of the previous classification node. Select the branch with smaller tolerable mistake value.

(5) Repeat (a–d), calculation and obtain the smallest value of $U_{\mathrm{CF}}(E;N)$ for all sub-group branches and then complete the pruning procedures.

3.2.3. Set up associate weighted value

3.2.3.1. Weighted value for patients with different triage categories. We set up the weighted value of patients in different degrees of triage according to the time to give emergency treatment. In the case hospital, the first degree should be treated within three minutes; second degree needs to be treated within 10 min, third degree should be treated within 30 min and the fourth degree can wait and be treated at a later time. This study set up the fourth degree treatment time within 60 min. According to this treatment time and calculate the weighted values (formula 12) for all patients in all levels by inverse calculation and modify (formula 13) to obtain the accurate percentage.

When we put those weighted values into the decision tree, we find a problem. Because the numbers of patients from first to fourth degree are all different, if we substitute the weighted value in the case study, the final patient number is significantly increased from 54,493 to 59,458. This actually affects the decision tree's final result, receiving slightly higher physician numbers than our forecast. Therefore, a second time weighed value modification (formula 14) is executed again. Finally the weighted value for first-degree patient we obtained is 1.186, 2nd degree patients is 1.103, 3rd degree is 0.866 and 4th degree is 0.51. We substitute the most appropriate weighted rate with numbers of patients in all degrees and receive the total patient number of 54,492.65, which very close to our requests. The weighted rate after calculations are stated in the following Table 1.

$$B_i = \frac{\sum_{i=1}^{n} Ai - A_i}{\sum_{i=1}^{n} Ai}, \quad i = 1 \ldots 4 \qquad \text{(formula 12)}$$

$$C_i = (B_i) * \left(\frac{4}{\sum_{i=1}^{4} B_i}\right), \quad i = 1 \ldots 4 \qquad \text{(formula 13)}$$

$$W_i = (C_i) * \left(\frac{\text{total actual numbers of people needed per day}}{\text{total number of people after first time weighted value calculation}}\right) \qquad \text{(formula 14)}$$

3.2.3.2. Physicians service patient's weighted value problem. The last output predictor in this decision tree study is how many patients should a physician see per day, while we are trying to find out numbers of physicians needed according to current patient numbers. According to The Triage and Work Standard for Hospitals with Emergency Services, the Regulation Ten (III) states that there should be more than 7 physicians with critical emergency services profession for Emergency Medicine Department, among that at least 50% are certified Emergency Medicine physicians. If the emergency patient numbers are more than 35,000 people in the previous year, then one emergency medicine physician should be added for every excess 4000 patients (Department of Health, Executive Yuan, 2006). Under this standard, the case hospital in this study had 108,480 patients in 2004, so 25 physicians are needed and each physician should see 11.88 patients per day. However, the actual situation always conflicts with the theory. In this study's shift arrangement principle, a physician serves at most 12 persons per day. According to the total patient number in the Internal Medicine Department of 54,493, a physician needs to serve 12.44 patients a day. If we set the regulation standard to be with the best diagnosis satisfaction, this study has achieved 95.5% (% of the best service quality divides with actual services quality) services satisfaction. In order to implement the shift arrangements that are decided by our decision tree, in this study we take 12.44 as the weighted value to be the serviced patients for a physician per day.

3.3. Result of the decision tree

After deciding the weighted value for patients in all degrees and physicians' services quality, we substitute those values into this study, totaling 732 samples if we divide 2004 into 2 shifts (day and night) per day. The input predictors are week, shift and holiday and the output predictor is numbers of physicians needed (formula 15). Before inputting and analyzing the decision tree, in order to prevent any in-effective factors or interfering factors in our selected predictors, we should inspect the input predictor and output predictor to see if they have related factors. From SPSS software, which runs out the Pearson correlation factors, stated in the following Table 2, each predictors all achieve $\alpha = 0.01$ standard, showing us that this study can be analyzed by decision tree.

$$\text{Numbers of physicians needed} = \frac{Q_i * W_i}{F} \qquad \text{(formula 15)}$$

Q_i is the number of patient in the i degree; W_i is the weighted value in the i degree; F is averaged physician's services quality.

It selects four layers as the maximum inputting branches. After our testing, any parameter above four receives the same results with the result coming out from the input value at four. And input that is less than the 4th layer, like the 3rd layer, the 2nd layer or the 1st layer has far larger forecast error rate than the 4th layer input. Therefore, we adopt 4 as our parameter for the maximum branch numbers in this study. Training samples and testing samples are randomly selected by computer, which selects 33.33% as the testing samples. In this decision tree structure, accuracy rate of the training samples is 50% and passing rate for the testing samples is 46.77%, as shown in Tables 3 and 4 below. The constructed tree shaped diagram is shown in Fig. 4. From Table 3 the accuracy rate of the testing samples, it uses 484 samples as the testing samples and the forecast results are divided into 5 forecasted circumstances, which are 5 physicians needed, 6 physicians needed, 7 physicians needed, 8 physicians needed and 11 physicians needed.

Table 1
Weighted rate operational table for patients in all levels

	1st Degree	2nd Degree	3rd Degree	4th Degree	Total
Treatment time (min) (A_i)	3	10	30	60	103
Inverse ratio (B_i)	0.971	0.903	0.709	0.417	3
Weighted rate after modification (C_i)	1.294	1.204	0.945	0.557	4
The most appropriate weighted rate (W_i)	1.186	1.103	0.866	0.510	–

Table 2
Manpower shift arrangement predictors' correlation factors examination table

Predictors	Relationship	Pearson correlation
Week	Week * numbers of physician needed	0.126[a]
Shift	Shift * numbers of physician needed	0.690[a]
Holiday	If holidays (Y/N) * numbers of physicians needed	0.244[a]
Numbers of physician needed	Numbers of patient * numbers of physicians	1

[a] Means has achieved $\alpha = 0.01$ standards.

Table 3
Accuracy rate table of the decision tree training samples

Training samples = 484 (shifts)		Decision forecasted numbers of physicians needed					
		5	6	7	8	11	Total
Numbers of physicians needed in reality	3	4	0	0	0	0	4
	4	51	1	1	0	0	53
	5	112	15	12	0	0	139
	6	40	52	38	1	0	131
	7	4	22	63	10	0	99
	8	0	8	19	17	0	44
	9	0	0	3	5	0	8
	10	0	0	1	2	0	3
	11	0	0	0	2	1	3
Total		211	98	137	37	1	484 (shift)

Table 4
Accuracy rate table of the decision tree testing samples

Training Samples = 248 (shifts)		Decision forecasted numbers of physicians needed					
		5	6	7	8	11	Total
Numbers of physicians needed in reality	3	2	0	0	0	0	2
	4	23	0	1	0	0	24
	5	57	13	3	0	0	73
	6	15	21	25	2	0	63
	7	2	18	32	8	0	60
	8	0	6	6	6	0	18
	9	0	0	5	1	0	6
	10	0	0	0	2	0	2
Total		99	58	72	19	0	248

And there are nine possibilities in reality, which are 3 physicians needed, 4 physicians needed, 5 physicians needed, 6 physicians needed, 7 physicians needed, 8 physicians needed, 9 physicians needed, 10 physicians needed and 11 physicians needed. In Table 4, the accuracy rate of the training samples, uses 248 samples as the testing samples and the forecast results are divided into 5 forecasted circumstances, which are 5 physicians needed, 6 physicians needed, 7 physicians needed, 8 physicians needed and 11 physicians needed. And there are eight possibilities in the reality, which are 3 physicians needed, 4 physicians needed, 5 physicians needed, 6 physicians needed, 7 physicians needed, 8 physicians needed, 9 physicians needed and 10 physicians needed. In this case study, because sample characteristics are put in such classification routes, the final result and sample tree for each node obviously has normal distributed demand features. Besides, in this study the final results all fall in the range of the sample's mode, mean and error variables of the shortening forecast. However, if the decision tree does not come out a complete accurate forecast, the result then is calculated as the forecast error, therefore this study has a very low accuracy rate. And the accuracy rate of the original shift arrangement only reaches about 22% as well, far lower than accuracy rates of the training samples and testing samples, which are 50% and 46.77%, respectively.

From Fig. 4 above, we can see that we approximately use day/night shift as the first classification point. Day shift distribution locates the areas of 7 physicians needed and night shift distribution locates the area of 5 physicians needed. Then we classify in accordance with the days in a week, from Monday to Sunday. For the 3rd layer, we classify in accordance with holiday/non-holiday, however we have a little bit lower support rate in certain areas. When we consider the execution possibility, we take week as our shift arranging rules, as shown in Table 5, which follows. And the holiday-classified information are mostly classifications that are coming out from the large numbers of patients during the Chinese New Year Holidays; therefore, we put it on record. On holidays, we will add one more physician in accordance with the following shift arrangement in Table 5.

4. Improvement of the shift arrangement

We use the decision tree forecast in the previous procedure to predict the 2005 physicians shift arrangement demands in the case hospital. We transform the total 2005 patients into number of demanded physicians, taking it as the *X*-axis. The decision tree forecasted numbers of physicians on the shifted schedule and the 2005 actual numbers of physicians on the shifted schedule as the *Y*-axis. We use those values and finish the following Performance evaluation matrix. In the decision tree forecasted shift arrangement performance evaluation matrix (Fig. 5), we calculate the doubled SD, using formula (formula 4), is 0.0866 and the tripled SD is 0.2598. The upper control line is made up by coordinates [0, 0.2589] and [0.7402, 1] in the performance evaluation matrix (Fig. 5) and the lower control line is made up by coordinates [0.2589, 0] and [1, 0.7402]. Moreover, in 2005 our Case Hospital's Shift Arrangement Performance evaluation matrix, still using formula (formula 4) we calculate the doubled SD is 0.913 and tripled SD is 0.2739. And the upper control line is made up by coordinates [0, 0.2739] and [0.7261, 1] and the lower control line is made up by coordinates [0.2739, 0] and [1, 0.7261]. The UCL, LCL and coordinate figure are shown in Table 6.

From Fig. 6, performance evaluation matrix for the Decision Tree Forecasts, we can see most forecast points fall evenly within the control center line and control line regardless of the day shift coordinates or the night shift coordinates. There is only two forecast points above the UCL and four points below the LCL. In Fig. 5, performance evaluation matrix for 2005 actual shift arrangement, we can clearly see that for the night shifts, points locate more at the right side so the manpower arrangement is inadequate and three points are dropping out of LCL. For the day shifts, there are three forecast points above the UCL and 5 points below the LCL. Those coordinates are shown in the following Table 7.

Then, we list the out-of-control line coordinates and date, as showed in the upper Table 8, to look for the reasons that may cause

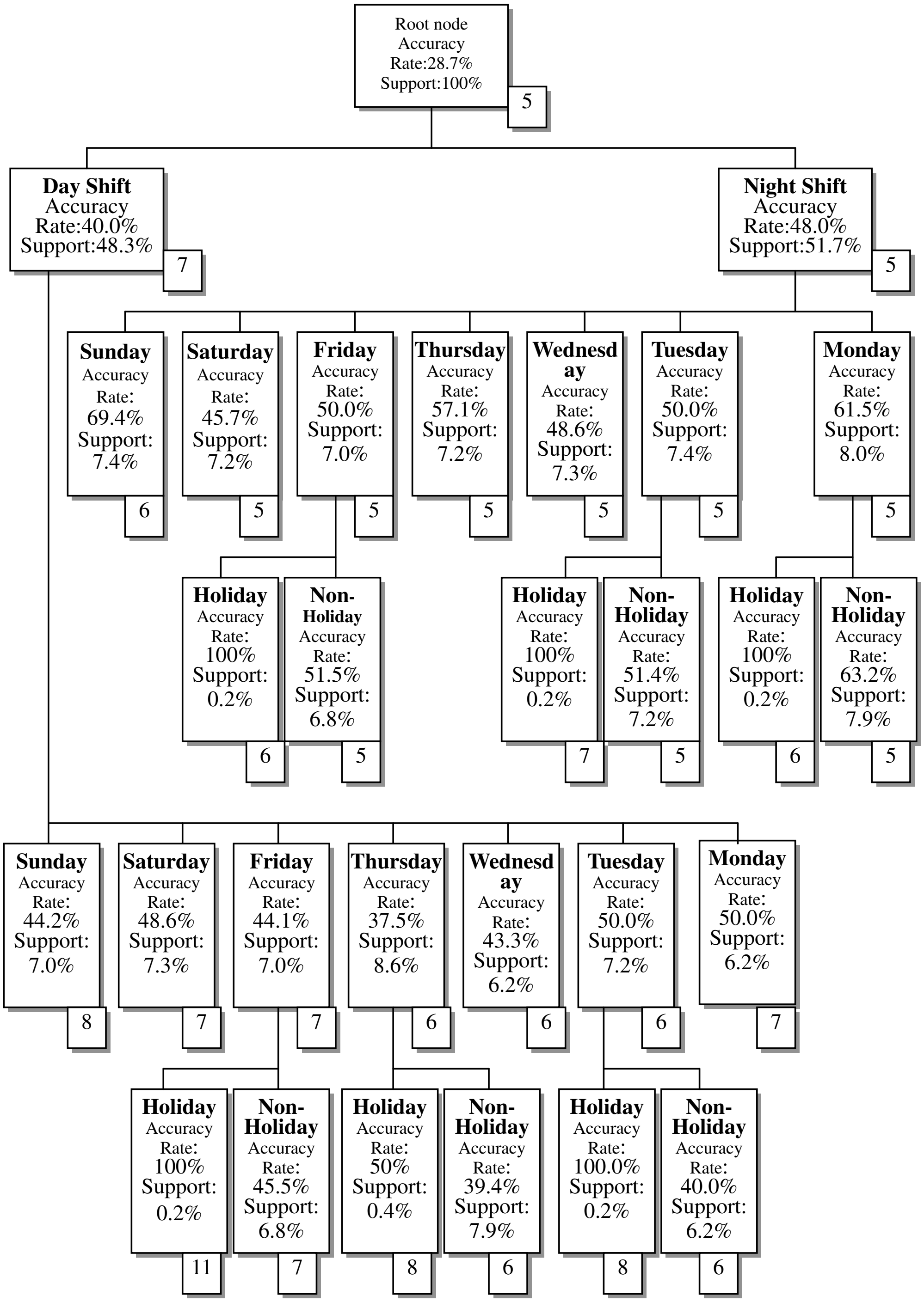

Fig. 4. Decision tree.

the coordinates locating out of control lines. We can see several phenomena from Table 8.

Points that are out of control lines, regardless in the day shifts or night shifts, mainly come from February 10th to February 13th. That was the Chinese New Year holidays and the number of patients reached the peak. Although this decision tree forecast model has increased numbers of physicians and actually reached the maximum improvement, this kind of situation is not controllable. Among those points at the day shift on February 12th in our forecast model, although they are out of lower control lines, but because we have stricter control line range, it happens in the out of control line circumstances. Besides, the forecast model is the same with the original shift arrangement that day and the performance evaluation matrix of the original shift arrangement does not have the out of control line points. So, if we use original shift arrangement, those points should be falling within the control lines. Therefore, those points can be treated as tolerable forecast error. And for the night shift at the lower control line, because this study's forecast model adjusts some day shift manpower to the night shift, there is no point out of the control line. However, the original shift arrangement during the Chinese New Year holidays represents a manpower shortage. So generally in the lower control line, this study's forecast model has better control than the original shift arrangement.

Table 5
Table of manpower arrangement forecast rules

Rule #	Rule	Decision
1	IF Week = 1 and Shift = Day Shift	7 Physicians
2	IF Week = 2 and Shift = Day Shift	6 Physicians
3	IF Week = 3 and Shift = Day Shift	6 Physicians
4	IF Week = 4 and Shift = Day Shift	6 Physicians
5	IF Week = 5 and Shift = Day Shift	7 Physicians
6	IF Week = 6 and Shift = Day Shift	7 Physicians
7	IF Week = 7 and Shift = Day Shift	8 Physicians
8	IF Week = 1 and Shift = Night Shift	5 Physicians
9	IF Week = 2 and Shift = Night Shift	5 Physicians
10	IF Week = 3 and Shift = Night Shift	5 Physicians
11	IF Week = 4 and Shift = Night Shift	5 Physicians
12	IF Week = 5 and Shift = Night Shift	5 Physicians
13	IF Week = 6 and Shift = Night Shift	5 Physicians
14	IF Week = 7 and Shift = Night Shift	6 Physicians
15	If Rule 14 and National Holiday = Yes	Add one more

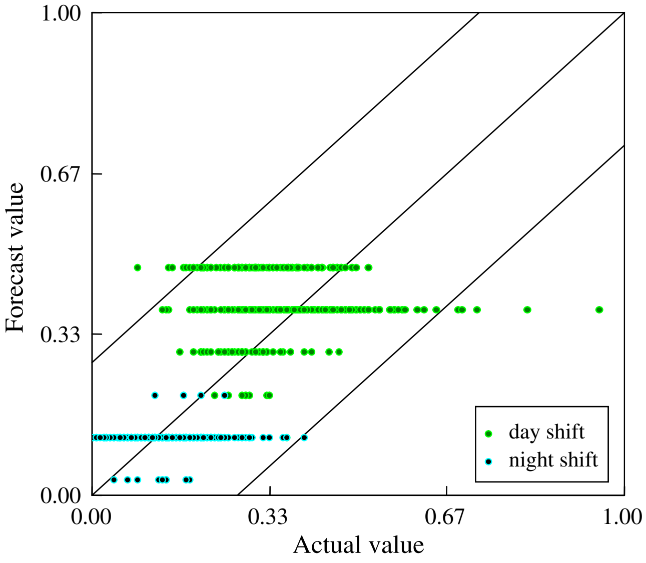

Fig. 5. Performance evaluation matrix for 2005 actual shift arrangement (before).

In points above the upper control line, the study's forecast model has excess manpower during the day shift on October 2nd. October 2nd was Sunday, however, and very few patients showed up. According to past experiences, it actually was a non-occasional case. If we apply it to the original shift arrangement's control lines, the point still falls in the control range. Therefore, this forecast error is acceptable. August 19th, December 22nd and December 14 were among the top four days (October 2nd was the second place) with the least patients; therefore the results of both the original shift arrangement and this study's forecast model are all above the upper control line. But the forecast model has one physician less than the original shift schedule. The improvement has been achieved.

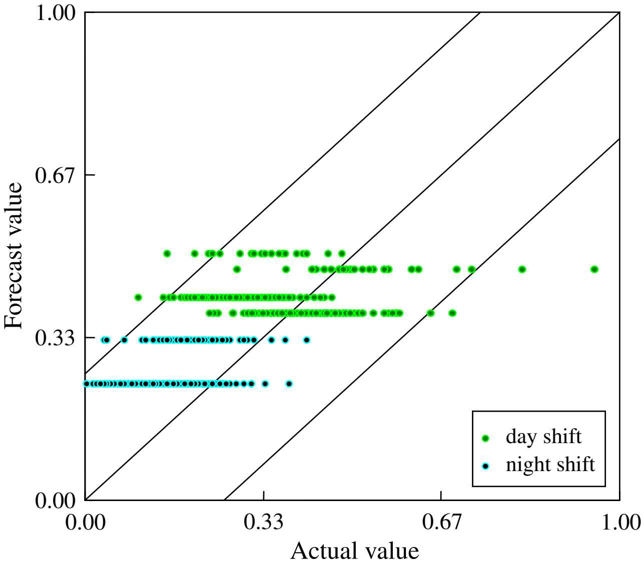

Fig. 6. Performance evaluation matrix for the decision tree forecasts (after).

Table 7
Performance evaluation matrix control Table 1

	Shift	Points inside the control lines	Points out of the UCL	Points out of the LCL	Total
Performance evaluation matrix for the decision tree forecasts	Day shift	359	2	4	365
	Night shift	365	0	0	365
Performance evaluation matrix for 2005 actual shift arrangement	Day shift	357	3	5	365
	Night shift	362	0	3	365

5. Cost evaluation and control

5.1. Non-symmetrical Taguchi's loss function

From current shift arrangement and the forecast shift arrangement by data mining, through non-symmetrical Taguchi's Loss Function the most appropriate and the most cost-efficient manpower demand can be found, and also costs can be improved. The concept is through the Loss Function to quantify and calculate the possible losses that manpower demand potentially can cause and bring to the hospital. The quality loss function brought up by Taguchi tells us that when the quality characters is always from the target value m, even a qualified product will still incur a square progressed increase quality loss (potential customers' loyalty loss). Concept of the Loss Function is stated in Fig. 7, which follows.

Table 6
Performance evaluation matrix corresponded coordinates

Index value and coordinate	Mean of the population, μ	SD of the population, σ	Tripled SD UCL	Tripled SD LCL	UCL coordinate (x,y)	LCL coordinate (x,y)
Performance evaluation matrix for the decision tree forecasts	0.0715	0.0866	0.2589	−0.2589	[0,0.2589] [0.2589,0]	[0.7402,1] [1,0.7402]
Performance evaluation matrix for 2005 actual shift arrangement	0.0858	0.0913	0.2739	−0.2739	[0,0.2739] [0.2739,0]	[0.7261,1] [1,0.7261]

Table 8
Performance evaluation matrix control Table 2

	Shift	Coordinates above UCL	Coordinates below LCL
Performance evaluation matrix for the decision tree forecasts	Day shift	Oct. 02 [0.178,0.471] Dec. 14 [0.121,0.383]	Feb. 05 [0.690,0.383] Feb. 10 [0.953,0.471] Feb. 11 [0.818,0.383] Feb. 12 [0.647,0.383]
	Night shift	N/A	N/A
Performance evaluation matrix for 2005 actual shift arrangement	Day shift	Aug. 19 [0.185,0.471] Dec. 14 [0.121,0.383] Dec. 22 [0.192,0.471]	Feb. 05 [0.690,0.383] Feb. 09 [0.725,0.383] Feb. 10 [0.953,0.383] Feb. 11 [0.818,0.383] Feb. 13 [0.697,0.383]
	Night shift	N/A	Feb. 10 [0.420,0.118] Feb. 11 [0.462,0118] Feb. 12 [0.427,0.118]

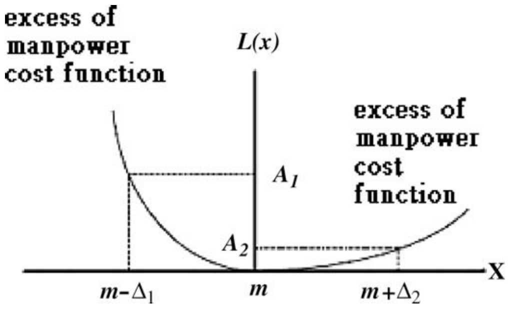

Fig. 7. Manpower forecast cost loss function diagram.

And in this study, manpower's cost function is defined as the following: X_{ij} is the distance of the manpower cost performance index and the target function; $TC_{\text{the } n \text{ saure}}$ is the total cost of the nth method; $i = 1,2,\ldots,365$ (actual variables); $j = 1,2,\ldots,365$ (predicted variables); $n = 1,2$ (1 is data mining forecast model; 2 is original shift arrangement); Δ_1 is the distance of the UCL and the target function; Δ_2 is the distance of the LCL and the target function; A_1 is the manpower cost loss when cost of excess of manpower is in the UCL; A_2 is the manpower cost loss when cost of deficient manpower is in the LCL.

$$k_1 = \frac{A_1}{\Delta_1^2}$$
$$k_2 = \frac{A_2}{\Delta_2^2} \qquad \text{(formula 16)}$$
$$L(x) = \begin{cases} k_1(x-m)^2 & \text{if} \quad x \leqq m \\ k_2(x-m)^2 & \text{if} \quad x \geqq m \end{cases}$$

$$\mathrm{TC}_{\text{then the method}} = \sum_{i=1,j=1}^{i>j} k_1(x_{ij}-m)^2 + \sum_{i=1,j=1}^{i<j} k_2(x_{ij}-m)^2 \quad \text{(formula 17)}$$

5.2. Results of cost control

Using the concept of non-symmetrical Taguchi's Loss Function, we calculate the cost loss with the distance that is away from the center value. For cost loss that is caused by manpower shortage, the deficient manpower might result in bad service quality or medical treatment disputes. According to the 653rd Weekly Report from Academia Sinica, total costs in time spent on handling the medical treatment disputes and the consulting services costs for appeals, litigations or rejections in 1997 reached NT$370,000,000; each physician had to spend NT$18,535 on average. And money that actually is spent on medical indemnification is NT$230,000 per case. In the past year, total cost for medical indemnification was around NT$270,000,000; each physician paid NT$13,690 in average, i.e. in every dollar expenditure, 40 cents are spent for indemnification and the remaining 60 cents are the so-called transaction cost. In average each physician has to spend NT$32,235 dollars each year for dealing with medical disputes. There are totaled 4064 medical man-hours in this case study, which accounts for 11.13 physicians; therefore total costs spent in medical dispute in Internal Medicine Department is converted and expected to be NT$358,775 ($\mathrm{TC_a}$). The total distance for manpower shortage in actual shift arrangement in 2005 is 7.13 ($\mathrm{D_a}$). Cost of manpower shortage can be calculated (through formula 18) to be NT$50,319.07. In the viewpoint of excess of manpower, physicians in the Medicine Department in average are paid NT3000 ($\mathrm{TC_b}$) per shift. From the formula, the matrix distant for every excess manpower or every deficient manpower is 0.0728 ($\mathrm{D_b}$), which can be used to calculate (by formula 19) the cost of excess manpower, which is NT$41,208.79

$$W1 = \frac{\mathrm{TC_a}}{\mathrm{D_a}} \qquad \text{(formula 18)}$$
$$W2 = \frac{\mathrm{TC_b}}{\mathrm{D_b}} \qquad \text{(formula 19)}$$

Finally, we substitute $W1$ and $W2$ into formula 16 and obtain formula 20, while the total cost is stated in formula 25. We substitute points in the forecast performance evaluation matrix, which is constructed by actual shift arrangement and actual physicians demanded, before improvement and also points in the forecast performance evaluation matrix, which is constructed by forecast physicians needed and actual physicians needed, after improvement into (formula 21) and obtain the final estimated total loss in

Table 9
Estimated total cost calculation table

Performance evaluation matrix	Manpower shortage matrix distance (A1)	Excess of manpower matrix distance ($A2$)	Total distance for points away from the Matrix's center point ($A1 + A2$)	Cost of manpower shortage matrix unit ($W1$)	Cost of excess of manpower matrix unit ($W2$)	The estimated total cost loss (TC)
Decision tree forecast performance evaluation matrix	2.92	5.55	7.84	50319.07	41208.79	341880
2005 Actual shift arrangement performance evaluation matrix	7.13	4.51	11.64	50319.07	41208.79	540065

cost (TG). The above-mentioned costs and distance of the matrixes are shown in Table 9, which follows:

$$L(x) = \begin{cases} (41208.79)(\mathrm{PF}_{ij} - \mathrm{PT}_{ij})^2 & \text{if} \quad \mathrm{PF}_{ij} < \mathrm{PT}_{ij} \\ (50319.07)(\mathrm{PF}_{ij} - \mathrm{PT}_{ij})^2 & \text{if} \quad \mathrm{PF}_{ij} \geqq \mathrm{PT}_{ij} \end{cases} \qquad \text{(formula 20)}$$

$$\mathrm{TC} = \sum_{i=1,j=1}^{i>j} (41208.79)(\mathrm{PF}_{ij} - \mathrm{PT}_{ij})^2 + \sum_{i=1,j=1}^{i<j} (50319.07)(\mathrm{PF}_{ij} - \mathrm{PT}_{ij})^2 \qquad \text{(formula 21)}$$

In the problem of physician's shift arrangement, using the decision tree analysis to forecast the shift arrangement problem can also achieve good results. For costs of the original shift arrangement through the performance evaluation matrix analysis, it is estimated that the total costs from excess of manpower and manpower shortage is about NT$540,065, while the total forecast cost from decision tree is NT$341,880, showing a pretty good achievement by this forecast model.

6. Conclusion

First of all, this study uses the 6-Sigma SD action procedures to establish the research steps, then through DMAIC measure to define and clarify our problems, causes, sample drawing measurement standard and measures to standardize predictors. By using numbers of physicians needed, which is converted from number of patients in the case hospital in 2004, we forecast number of emergency physicians needed in 2005 for the case hospital through data mining decision tree classification. We also take realistic conditions into consideration and make appropriate solutions. Besides, we also construct the forecast performance evaluation matrix with ideas from Taguchi's Loss Function to estimate the cost loss and also use the performance evaluation matrix to control and continuously improve the shift arrangement situation. By data mining decision tree classification, we not only forecast the physician demand in the case hospital, but also allocate emergency medicine manpower, obtain the decision tree model for physician demands and forecast, and present a new shift arrangement method. The accuracy rate is increased to 50% from the original 22%. Finally we use the forecast performance evaluation matrix to evaluate and quantify the performance of the data mining classification forecast model and the performance of the original shift arrangement; the total cost is reduced from the original's NT$540,065 dollars to NT$341,880 dollars, showing a 37% significant decrease in cost loss from shift arrangement.

The adopted data mining forecast model presents smaller values in both total estimated distance and total forecast costs. By combining the control line diagram ideas and the performance evaluation matrix, we focus on samples that are not in the range of control lines and obtain an understanding on each single sample. Therefore the forecast model, even if a few sample points are out of control line range, has better shift arrangement than the original shift arrangement, and actually obtains improvements. It also uses non-symmetrical Taguchi's Loss Function to estimate and evaluate the possible cost loss. It successfully constructs a performance evaluation model to evaluate our forecast. Hopefully through the completion of this study, it cannot only solve the existing problem in this case hospital, but also through the action procedures in this study to introduce in a different medical treatment environment.

References

Academia Sinica (1998). Influences to medical costs that are caused by medical disputes: The real study for physicians in Taiwan. Weekly Report from the Academia Sinica, (pp. 653, 111).

Adriaans, P., & Zantinge, D. (1996). *Data mining*. New York: Addison Wesley.

Berry, M. J. A., & Linoff, G. S. (2004). *Data mining techniques: For marketing, sales, and customer relationship management*. New York: Barnesandnoble.

Chou, C. H. (2003). The influence of staffing change on quality of care in emergency room-an example of three hospitals [Master's Thesis]. National Sun Yat-Sen University Taiwan, Taiwan: Kaohsiung.

Dessler, G. (2006). *Human resource management*. Englewood Cliffs, NJ: Prentice-Hall.

Department of Health, Executive Yuan, (2006). The triage and work standard for hospitals with emergency services. Department of Health, Executive Yuan in Taiwan.

Giudici, P. (2003). *Applied data mining: Statistical methods for business and industry*. New York: Wiley.

Hand, D., Mannila, H., & Smyth, P. (2001). *Principles of data mining*. Cambridge, MA: MIT Press.

Kantardzic, M. (2002). *Data mining concepts, models, methods, and algorithms*. Piscataway, NJ: Wiley-IEEE.

Kdnuggets. (2007). Rexer analytics data miner survey summary report. Kdnuggets News, URL: http://www.kdnuggets.com/news/2007/n16/6i.html, Visor, 2007.8.10.

LaMar, J. E., Jacoby, I., Meyer, G. S., & Potter, A. L. (1997). Provider workforce model for regional TRICARE networks. *Military Medicine, 162*(2), 590–596.

Lambert, D. M., & Sharma, A. (1990). A customer-based competitive analysis for logistics decisions. *International Journal of Physical Distribution and Logistics Management, 20*(1), 23.

Lin, W. T., Chen, S. C., & Chen, K. S. (2005). Evaluation of performance in introducing CE marking on the european market to the machinery industry in Taiwan. *International Journal of Quality Reliability Management, 22*(5), 503–517.

Liu, L. L., & Wang, J. Y. (2005). Attending physicians' preferences for shift-scheduling. *Taipei City Medical Journal, 2*(2), 193–202.

George, M. L., Rowlands, D., Price, M., & Maxey, J. (2005). *The lean six sigma pocket toolbook*. New Orleans: McGraw-Hill.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1991). Understanding customer expectation of service. *Sloan Management Review, 32*(3), 39–48.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*(1), 81–106.

Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. San Mateo, CA: Morgan Kaufmann.

Richard, J. N. (1988). Models for human resource decisions. *Human Resource Planning, 11*(2), 95–107.

Roiger, R. J., & Geatz, M. W. (2003). *Data mining a tutorial-based primer*. New York: Addison-Wesley.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technology Journal, 27*, 379–423.

Soukup, T., & Dabifdon, I. (2002). *Visual data mining: Techniques and tools for data visualization and mining*. New York: Wiley.

Tsan, C. Y. (2006). Emergency triage: Past, current and future. *National Taiwan University Hospital Journal of Nursing, 4*(3), 24–30.

Yang, C. Y., & Yang, H. I. (2004). The difference of medical malpractice disputes between traditional chinese medicine and western medicine in Taiwan district. *Journal Chinese Medicine, 15*(1), 1–15.

Yeh, J. Y., Wang, S. F., Chen, W. Z., & Li, Y. D. (2003). Using a genetic algorithm for optimizing nurse schedule in a hospital emergency department. In *Proceedings of the Chinese institute of industrial engineers national conference*. Taiwan: Changhua.